Ontology Outreach Advisory Requirements

Mustafa Jarrar, Werner Ceusters (VU Brussel), Diana Maynard (U Sheffield), Rob Shearer (U Manchester), Lyndon Nixon, Elena Paslaru, Malgorzata Mochol (FU Berlin)

Abstract.
EU-IST Network of Excellence (NoE) IST-2004-507482 KWEB
Deliverable D1.3.3 (WP1.3)
This deliverable presents the requirements for setting up the Ontology Outreach Advisory OOA.

Document Identifier:	KWEB/2005/D1.3.3./v2.3
Class Deliverable	KWEB EU-IST-2004-507482
Version:	2.3
Date:	January 31, 2006
State:	Final Draft
Distribution:	Public

Knowledge Web Consortium

This document is part of a research project funded by the IST Program of the Commission of the European Communities as project number IST-2004-507482.

University of Innsbruck (UIBK) – Coordinator
Institute of Computer Science,
Technikerstrasse 13
A-6020 Innsbruck
Austria
Contact person: Dieter Fensel
E-mail address: dieter.fensel@uibk.ac.at

École Polythechnique Fédérale de Lausanne (EPFL)
Computer Science Department
Swiss Federal Institute of Technology
IN (Ecublens), CH-1015 Lausanne.
Switzerland
Contact person: Boi Faltings
E-mail address: boi.faltings@epfl.ch

France Telecom (FT)
4 Rue du Clos Courtel
35512 Cesson Sévigné
France. PO Box 91226
Contact person : Alain Leger
E-mail address: alain.leger@rd.francetelecom.com

Freie Universität Berlin (FU Berlin)
Takustrasse, 9
14195 Berlin
Germany
Contact person: Robert Tolksdorf
E-mail address: tolk@inf.fu-berlin.de

Free University of Bozen-Bolzano (FUB)
Piazza Domenicani 3
39100 Bolzano
Italy
Contact person: Enrico Franconi
E-mail address: franconi@inf.unibz.it

Institut National de Recherche en Informatique et en Automatique (INRIA)
ZIRST - 655 avenue de l'Europe - Montbonnot
Saint Martin
38334 Saint-Ismier
France
Contact person : Jérôme Euzenat
E-mail address: Jerome.Euzenat@inrialpes.fr

Centre for Research and Technology Hellas / Informatics and Telematics Institute (ITI-CERTH)
1st km Thermi – Panorama road
57001 Thermi-Thessaloniki
Greece. Po Box 361
Contact person: Michael G. Strintzis
E-mail address: strintzi@iti.gr

Learning Lab Lower Saxony (L3S)
Expo Plaza 1
30539 Hannover
Germany
Contact person: Wolfgang Nejdl
E-mail address: nejdl@learninglab.de

National University of Ireland Galway (NUIG)
National University of Ireland
Science and Technology Building
University Road
Galway
Ireland
Contact person: Christoph Bussler
E-mail address: chris.bussler@deri.ie

The Open University (OU)
Knowledge Media Institute
The Open University
Milton Keynes, MK7 6AA
United Kingdom.
Contact person: Enrico Motta
E-mail address: e.motta@open.ac.uk

Universidad Politécnica de Madrid (UPM)
Campus de Montegancedo sn
28660 Boadilla del Monte
Spain
Contact person: Asunción Gómez Pérez
E-mail address: asun@fi.upm.es

University of Karlsruhe (UKARL)
Institut für Angewandte Informatik und Formale Beschreibungsverfahren – AIFB
Universität Karlsruhe
D-76128 Karlsruhe
Germany
Contact person: Rudi Studer
E-mail address: studer@aifb.uni-karlsruhe.de

University of Liverpool (UniLiv)
Chadwick Building, Peach Street
L697ZF Liverpool
United Kingdom
Contact person: Michael Wooldridge
E-mail address: M.J.Wooldridge@csc.liv.ac.uk

University of Manchester (UoM)
Room 2.32. Kilburn Building, Department of Computer Science, University of Manchester,
Oxford Road
Manchester, M13 9PL
United Kingdom
Contact person: Carole Goble
E-mail address: carole@cs.man.ac.uk

University of Sheffield (USFD)
Regent Court, 211 Portobello street
S14DP Sheffield
United Kingdom
Contact person: Hamish Cunningham
E-mail address: __obert@dcs.shef.ac.uk

University of Trento (UniTn)
Via Sommarive 14
38050 Trento
Italy
Contact person: Fausto Giunchiglia
E-mail address: fausto@dit.unitn.it

Vrije Universiteit Amsterdam (VUA)
De Boelelaan 1081a
1081HV. Amsterdam
The Netherlands
Contact person: Frank van Harmelen
E-mail address: Frank.van.Harmelen@cs.vu.nl

Vrije Universiteit Brussel (VUB)
Pleinlaan 2, Building G10
1050 Brussels
Belgium
Contact person: Robert Meersman
E-mail address: __obert.meersman@vub.ac.be

Work package participants

The following partners have taken an active part in the work leading to the elaboration of this document, even if they might not have directly contributed writing parts of this document:

FU Berlin	Lyndon Nixon, Elena Paslaru, Malgorzata Mochol
USFD	Diana Maynard
VUB	**Mustafa Jarrar, Werner Ceusters**
VUM	Rob Shearer

Changes

Version	Date	Author	Changes
0.1	1-09-2005	Mustafa Jarrar	First draft
0.2	25-10-2005	Werner Ceusters	Elaboration of the Life Sciences and Health Chapter Added type of OOA Members and advantages and Working groups Added Belgian legal requirements for not-for-profit organizations
0.9	03-11-2005	Rob Shearer	Ontology evaluation section updated
	11-11-2005	Elena Paslaru Bontas	Elaboration of Section 4.2
	11-11-2005	Diana Maynard	Added a new section on Lessons Learned from other initiatives, and edited the section on Other Domains.
	11-11-2005	Lyndon Nixon, Malgorzata Mochol	Rework of the Human Resource section
	12-11-2005	Mustafa Jarrar	Editing contributions
	12-11-2005	Werner Ceusters	Edited all contributions
1.0	15-11-2005	Diana Maynard, Rob Shearer	Ontology evaluation section updated
	16-11-2005	Lyndon Nixon	Added section on the XML Clearinghouse
		Diana Maynard	Language checking
	19-11-2005	Werner Ceusters	Overall integration
1.1	22-11-2005	Lyndon Nixon, Diana Maynard, Elena Paslaru Bontas, Werner Ceusters	Last minute additions and changes
	22-11-2005	Mustafa Jarrar	Review and Editing
	23-11-2005	Werner Ceusters	Final Integration and Post-editing
1.2	25-11-2005	Werner Ceusters, Diana Maynard	Final verification
1.9	25-11-2005	Werner Ceusters	Integration of the reviewers' comments, adding up to date information
2.1	24/01/2006	Mustafa Jarrar	Integration and edition of final contributions and comments.
2.1	25/01/2006	Werner Ceusters	Final edit
2.2	30/01/2006	Mustafa Jarrar	Final editing and review, submit to partners
2.3	31/01/2006	Mustafa Jarrar	Finalization, submit to Coordinator

Executive Summary

This deliverable is a report of the activities conducted in preparing the creation of the KnowledgeWeb Ontology Outreach Advisory (OOA).

To carry out the OOA's mission, which is the development of ontology standardization and recommendation strategies, thereby promoting and outreaching good quality ontological content (chapter 2), two requirements have been identified: (1) the availability of a suitable ontology evaluation framework and (2) a realistic operational work plan to keep the OOA self-sustainable (chapter 3).

With respect to the former, we have done an extensive study of the available literature, exchanged ideas with various stakeholders, and looked deeply into third party attempts to work in similar directions (chapter 4). From this we concluded that an ontology can be submitted for recommendation to the Ontology Outreach Advisory if it is provided with high quality documentation of the engineering process and an accurate documentation of internal evaluation results as generated by the engineering team. Evaluation by the OOA will then consist of assessing the adequacy of the internal evaluation according to the original requirements specification.

With respect to the latter, we focused on both the internal working principles of the OOA and their associated legal, financial, administrative and managerial requirements (chapter 5), and the external operations. The best option thus far seems to be the creation of the OOA as one of the four not-for-profit organisation types that are recognised in Belgium (chapter 6), but which one precisely, requires more insight in the pro and cons.

Approach of the market is carried out in line with broad business domains, and this sectorial approach is reflected in the structure of the OOA as it exists thus far: domain committees take care of outreach and evaluation in their particular domain, while a Steering Board is responsible for the overall functioning. So far, two sectorial committees have been set up: one for healthcare and life sciences (section 7.1), and one for human resources (7.2). These two domains have been identified as being most promising for early success on the basis of the wealth and variety of ontology-related activities and stakeholders that they exhibit. Meanwhile, other domains such as legal, financial, e-Government and e-Learning are also being investigated (chapter 8).

The present deliverable, we feel, forms a sound basis to work towards the next milestone, which is the actual creation of the OOA as an independent legal body, and which is scheduled for M36.

Contents

1 Introduction

This deliverable describes primarily the requirements that have been identified to create and maintain the KnowledgeWeb Ontology Outreach Advisory (OOA) as an independent organisation whose main objective is to become a trusted source in the certification of claims made about particular ontologies, and as such is intended to be indirectly seen as the recommending body for ontologies. In addition, this deliverable describes the activities conducted thus far in preparing the creation of the OOA which at this moment is in an incipient stage.

The Ontology Outreach Advisory (OOA) was originally conceived to be an Ontology Outreach ***Authority***, a meeting place for interacting with interested industrial parties to take advantage of the last research results, including tools. This outreach action would be based on the dozens of industries that had already expressed their interest for the FP5 OntoWeb Thematic Network. In order to realize this objective, KnowledgeWeb set up an alliance with several industry bodies, called the Industrial Board [1]. Discussions with this Board, as with individual industries, revealed that it would be more appropriate to set up an ***Advisory***, rather than an Authority. This resulted in a slightly modified way of moving forward. While at the level of the Industrial Board agreements on using initial results and tools are taken care of, To achieve this, a new and better role of the OOA was defined, its main working principles being the organization of industry into domains/sectors and setting up a committee ("chapter") for each sector where each committee will consist of the most active (industrial + research + (maybe) government) members with respect to ontology development, use, or education in that sector. Each committee is expected to play a leading role in ontology development, validation, and recommendation activities within its sector.

The planned deliverables foreseen in the Project plan M132-M48, consist of:

D 1.3.3 [Planned: month 24] Report on requirements of OOA.

D 1.3.6 [Planned: month 36] Set up of OOA.

D 1.3.7 [Planned: month 48] Standards on metadata for ontology and interoperability.

D 1.3.8 [Planned: month 48] Report on OOA activities.

2 OOA Mission

The OOA is devoted to develop and extend ontology standardization and recommendation strategies, thereby promoting and outreaching good quality ontological content.

In the present state of the Semantic Web and of its adoption by industry, ontological content is difficult to standardize by the methods used by typical standardization bodies. Content is typically subjective and application-dependent, and general evaluation criteria are lacking. Therefore a form of "standardization lite" by recommending ontologies will be adopted. The OOA is intended to play the role of an ontology recommendation body. The focus and operation of the OOA therefore will be (mainly) content-oriented rather than (only) technology-oriented. Good quality ontologies will be recommended and promoted to industry. The OOA will consist of the most active (industrial + governmental + research) members with respect to ontology development, use, or education.

3 OOA Goals and requirements

The main and strategic goals of the OOA, in achieving its mission, are:

1. *Ontology Standardization.* Evaluation and recommendation strategies of ontologies.
2. *Ontology Outreaching.* Promotion of ontology content.

To achieve these goals, there are two important requirements that the OOA should provide:

1. a suitable ontology evaluation framework that can be generalized for ontology recommendation;
2. a realistic operational workplan that keeps the OOA self-sustainable.

In the following we elaborate on these issues and discuss the possible difficulties and opportunities.

3.1 Ontology evaluation criteria

Ontology evaluation, i.e. assessing the quality of an ontology, is interpreted in various ways in the Semantic Web community. In addition to the difficulties encountered when attempting to clearly define what constitutes an ontology, every approach to ontology evaluation is confronted with the following dilemmas:

- Can a good and commonly agreed data schema be considered as a good ontology, or even as an ontology at all?
- Should ontologies only account for a particular meaning of a domain vocabulary?
- Should ontologies only capture distinguishing characteristics and intrinsic properties of the concepts being modelled?
- Should ontologies commit to reality (i.e. should axioms be valid/true in reality)?
- Can a "true/correct" ontology be considered as a good ontology even if it is rarely used in practice?
- Can an ontology be considered as a good ontology simply because it is popular (i.e. has been used in many applications)?
- Should a good ontology be application independent?

The answers provided by the Semantic Web research to such fundamental issues are still controversial and require further investigation. One approach was presented in the call-for-papers of the Workshop on Core Ontologies in Ontology Engineering [2]:

> "*…By "successful ontology" here we mean (at least one of) the following settings:*
> *a) the core ontology has been used to reach an agreement on the types of entities (and their relationships) needed in a community of practice.*

b) the core ontology is being used to dynamically negotiate the intended meaning across a distributed community
c) the core ontology has been used to align/ integrate/ merge several sources of metadata/ ontologies/ terminologies
d) the core ontology has been used to build more than one application or service
e) the core ontology has been adopted as a template for specifying the content in some domain)…".

More details about existing approaches to ontology evaluation can be found in Deliverable D1.2.3. However, as reported in this deliverable, our analysis revealed that given the present state of the art, achieving an agreement or generalizing about ontology evaluation criteria is still a very difficult task.

The overall value of an ontology is clearly dependent on a number of different factors; an ontology which excels in one respect may provide little to no value in real applications if it does not meet minimal standards in other respects. The following have been identified as aspects of ontology value significant enough to merit attention in any evaluation. While there is some attempt to isolate orthogonal evaluation axes, some overlap between evaluation criteria is inevitable.

3.1.1 Ontological evaluation

This type of evaluation focuses on the relationship between an ontology (as a logical theory) and its underpinning conceptualization. The research on this kind of evaluation is mainly carried out by the LOA's research group [3].The main idea is that given a set of intended models (a conceptualization in mind) and a logical theory that accounts for it, the ontological quality of this theory is the ability to specify *(only and all)* these intended models. Non-intended models and situations should be excluded. To evaluate the quality of an ontological content, one can determine a list of situations that the ontology is supposed to cover; document these situations by means of illustrations (annotated multimedia documents) showing the agreed intended use of ontology terms; make sure that for each term multiple examples and counter-examples are given, and establish the correspondence between situations and ontology models.

Although this is indeed a fundamental and rigorous approach, which leads to high accuracy ontological content, it is difficult to apply in practice as it is too philosophical and requires intensive work. For example, according to this approach, an ontology that might be successful in practice (i.e. used in many applications) can be seen as a bad ontology if it is does not capture all (or only) the intended models. Such things are difficult to measure.

3.1.2 Logical evaluation

Logical evaluation is the main concern of description logics. Basic "satisfiability reasoning" can be performed to find out whether an ontology is *consistent* as a whole (free of contradictory axioms) as well as whether it declares any unusable vocabulary (unsatisfiable concepts). In other words, each vocabulary term in an ontology is satisfiable if an instance of that term could be instantiated. Whatever the underlying conceptualization, whatever the application/applicability, the ontology (as a logical

theory) is evaluated as to whether it is logically sound. For ontologies specified using DL-based languages, basic satisfiability evaluation can be performed automatically (e.g. using FaCT or Racer reasoners).

Logical evaluation can be extended to include analysis of the implications of the declared semantics of an ontology; lack of contradiction is an important (and easily checked) aspect of logical evaluation, but it is insufficient to determine whether the formal interpretation of the ontology is compatible with the intended semantics. Inferred classification hierarchies can be analyzed and axiomatic interactions identified; such consequences can be compared with the expectations and intentions of the modelers.

The richness, scope, and power of such logical implications are another important sub-axis of evaluation: some ontologies may offer sufficient value with little inferential power (providing little more than a collection of recommended vocabulary), while others may be designed specifically for the value of their ramifications with the formal framework (offering extensive formal semantics for the provided vocabulary and the opportunity for significant automated classification and reasoning).

3.1.3 Tractability and computability evaluation

While logical evaluation measures the correctness of formal interpretation, tractability evaluation is concerned with the ability of automated reasoning systems to discover and make use of that interpretation. Expressive ontology languages (such as OWL) are known to require reasoning algorithms with poor worst-case performance. Reasoner technology for such languages is quite advanced, however, and a large number of specialized optimization techniques are available which make true worst-case performance extremely rare in practice. Certain ontological structures are known to negatively impact the performance of such systems, however it is frequently possible to regain excellent reasoner performance with small modeling changes to the underlying ontology. Furthermore, experience has shown that poor reasoner performance is frequently an indicator of errors in logical modeling.

As automated processing has become a pervasive feature of ontology-driven applications, an understanding of an ontology's impact on reasoning systems is vital to evaluation of an ontology's overall value. Ontologies should be evaluated with respect to available tools, most importantly those with which they are intended to be used.

3.1.4 Standards compliance

The representation used to encode a particular ontology is a crucial element of that ontology's potential for reuse. Several ontology languages have been standardized, including RDF and three varieties of OWL, and these standards continue to evolve. For any given ontology, it is important to identify those standard(s) to which it conforms as well as any non-standard extensions used.

It is also common practice for modelers to make use of existing ontologies, many of which have achieved some level of standardization themselves. An appreciation for the standardization status of any referenced ontologies is thus another important element of evaluation and indicator of interoperability.

3.1.5 Applicability evaluation

Given an ontology, (and given particular application needs), ontology applicability evaluation is concerned with "how much the ontology fulfils the application needs". Notice that a good ontology (i.e. with high ontological and logical quality) might *not* be usable for applications, i.e. from a business/usefulness perspective. An ontology might be usable (useful/relevant) for an application but not for another. Some issues that make an ontology useful for an application but not for another -- we call them *usability perspectives* [4]-- for example are:

- coverage and scope boundaries
- granularity level
- knowledge structure (i.e. epistemology)
- specification language
- formality level (i.e. light-weight vs. expressiveness).

See [4] for more details about ontology usability vs. ontology reusability.

Another approach to evaluate ontology usability (or called suitability) is OntoMetric [5]. OntoMetric defines a set of processes that the user should carry out to obtain the measures of suitability of existing ontologies, regarding the requirements of a particular application. The basic decision criteria of this method are called dimensions, which are the fundamental aspects to be considered by the user before choosing an ontology. The dimensions specify the following features: the content represented in the ontology, the language in which the ontology is specified, the methodology followed to develop the ontology, the software environments used for building the ontology, and the costs of using the ontology in the system.

Notice that this kind of evaluation alone is not suitable for the OOA purposes, because ontology recommendation cannot be issued only based on/for one particular application. Evidence that an ontology provides value to one concrete application must be balanced against the other axes of evaluation (which tend to measure the ontology's potential for reuse).

3.2 Ontology evaluation framework

The aforementioned approaches identify an inventory of valuable criteria for performing ontology evaluation. However, given the different foci and the limited coverage of the particular evaluation methods, it is still unclear exactly how so many different (and sometimes contradictory) measures can be combined into a single meaningful measure reflecting the value of ontologies.

One can see from the evaluation criteria enumerated above that they reflect diversity of focus and understanding rather than agreement, while none of them alone has proven satisfactory applicability in practice yet. The question of the relative importance of the various criteria in the evaluation and recommendation of ontologies (especially at the domain level) is therefore still open, and indeed difficult to answer, due not only to the fuzziness of the task itself, but also because of its domain specificity. Each specific domain or even each application scenario may have its own evaluation priorities and best practices. Ontologies that are not good for certain applications or domains may be good for others. As a consequence, a possible

solution might be to strive for a more generic and extensible evaluation approach, which on one hand preserves domain and application independence (but can be customized along these dimensions) and on the other hand abstracts from particular, usually technically focused evaluation methods (as those described above) for applicability purposes. A first draft of this approach is elaborated in the next section.

3.2.1 Evaluate an ontology w.r.t. its requirements specification.

In this approach the ontology is evaluated w.r.t. its usability in its intended context, as foreseen by the majority of ontology engineering methodologies which have emerged in the last decades. According to these methodologies, preliminary to the conceptualization of the application domain and its implementation in a specific representation language, a team of ontology engineers and domain experts agrees on a set of requirements which should be fulfilled by the final ontology. Depending on the type of the requirements and the form in which they are specified (informal vs. formal, pre-defined template vs. unstructured document etc.) methodologies introduce different methods by which the implemented ontology can be evaluated against the requirements.

The activities of the OOA are aimed to complement this internal reviewing step. An ontology can be submitted for recommendation to the Ontology Outreach Advisory if it is provided with high quality documentation of the engineering process, in particular a complete requirements specification (including ontology goals, scope, domain coverage, additional information sources etc.) and an accurate documentation of internal evaluation results as generated by the engineering team. If the evaluation is positive, the ontology gets a recommendation stamp by the OOA as a certification of the evaluation results.

The proposed evaluation framework provides a generic description of the OOA evaluation process. In particular it identifies an inventory of documents, which are required by the organization in order for the evaluation process to be performed. As each domain sector has its own specificities and community of practice, a committee of evaluators (mainly representatives of industrial organizations) will be selected for pre-defined industrial sectors, called the *OOA chapters*. Each chapter will decide upon its own domain-specific evaluation method and may extend the set of minimal requirements introduced by the present framework.

After analyzing the current state of the art wrt. ontology evaluation approaches (as described above and illustrated in the Knowledge Web Deliverable 1.2.3) we have compiled a preliminary set of information items and associated documents which are minimally required for an OOA recommendation:

- **Ontology domain**: the domain which is modelled in the ontology. This information is required in order to assign the submitted ontology appropriately to the corresponding OOA chapter. In particular, the developers should classify the ontology according to its generality level (upper-level ontologies, domain ontologies, task ontologies etc.)
- **Information sources**: information sources an ontology relies on. Ontology developers should provide a list of documents which have been used for knowledge acquisition purposes during the building of the ontology. Among relevant information sources in this category they should mention the

dictionaries, thesauri, user interviews and glossaries from which an ontology was derived, but also directives, official documents and standards, laws or regulations in a particular domain with which an ontology might be compliant. These information sources provide additional input for the description of the ontology domain and might trigger a specific evaluation procedure.

- **Application scenario**: the type of application the submitted ontology was originally designed for, the role of the ontology in this setting and the application requirements. For instance, an ontology might be designed for a content management scenario in which it aims at semantically retrieving information. Knowledge about the application scenario is important for the evaluators as this has a direct impact on the application requirements the ontology is expected to fulfil (in the example above, such requirements might be related to precision and recall values)
- **Evaluation**: the methods, tools and results of the internal evaluation. In particular, the authoring organization should provide information about the syntactic validity (representation language and syntax, validation tools, results and documentation), the logical consistency (representation language, reasoning tool, results and documentation), application-specific evaluation (evaluation method, results, documentation) and/or general-purpose evaluation activities (such as expert reviewing, ontological evaluation etc.).
- **Documentation**: the documentation of the engineering process, including comments and definitions of the ontology primitives, information about the engineering methodology, design guidelines and patterns eventually applied to build the ontology.

The enumerated aspects are subject to further refinements. In particular they are to be extended with domain-specific requirements by the OOA chapters. In terms of their formulation they could be specified by means of an evaluation template or an extension module of the core ontology metadata vocabulary introduced in the Knowledge Web Deliverable 1.3.2 (see Appendices **9.3** and **9.4**, respectively).

3.3 Promoting and outreaching recommended ontologies

The aim of the OOA is not only to perform evaluation and recommendation of ontologies, but also to play an outreaching role. For example, all recommended ontologies will be promoted through a public demonstration event, newsletters, the OOA ontology repository, etc. Unlike the many conferences and events that are being organized within the Semantic Web community, which typically are *technology-oriented,* the OOA aims to organize a content-oriented annual event[1]. All recommended ontologies will be demonstrated and promoted in this event.

Furthermore, the OOA will investigate which of the many conferences and symposia that are organised in specific domains such as Life Sciences and Health, and Human Resources, two areas that have been identified as most promising for achieving early

[1] In fact, there only has been one workshop that has focus on ontology content (rather than technology) called "Core Ontologies in Ontology Engineering", at the EKAW 2004 conference. http://www.loa-cnr.it/core_onto.html (visited Sept. 2005)

successes (see section 7 p.30), are ideally suited for active participation. A list of forthcoming events in Health Informatics can be found in Table 3-1.

24-27 Jan 2006	Medicine meets Virtual Reality	California
6-7 Feb 2006	Putting EHR standards to work - implementing the e-Health Society	The Hague
5-7 Apr 2006	Med-e-Tel	Luxemburg
6-8 Apr 2006	EFMI Special Topic Conference & ROMEDINF : From Cell to e-Patient	Timisoara
18-20 Apr 2006	AINA-2006, the IEEE 20th international Conference on advanced Information Networking and Applications	Vienna
20-22 Apr 2006	ARES-2006, The First International Conference on Availability, Reliability and Security	Vienna
23-27 Apr 2006	ACM Special Interest Group on Applied Computing : Special Track on Computer Applications in Healthcare	Dijon
27 Apr 2006	11th Nursing Informatics	Mons
27-29 Apr 2006	IMIA WG 4 working conference on Secure e-Health	Dijon
24-25 Apr 2006	Visual Information Expert Workshop	Paris
3-5 May 2006	IST - Africa 2006, including a e-Health track	Pretoria
16-18 May 2006	AMIA Spring Congress : Informatics Across the Spectrum - From Clinical Care to Biomedical Research	Phoenix, Arizona
20-24 May 2006	TEPR'06, 22th Annual Conference and Exhibition	Baltimore
22-23 June 2006	19th IEEE, CBMS-2006 : special track The use of Open Source Software and Open Document Formats in Health and Medical Systems	Salt Lake City
31 Aug - 1 Sept 2006	6th Nordic Conference on eHealth and Telemedicine : From Tools to Services	Helsinki
13-20 Oct 2006	MEDNET-2006, 11th World Congress on Internet in Medicine	Toronto
7-8 Dec 2006	Telematics@health.be	Brussels
20-24 Aug 2007	Medinfo-2007	Brisbane
19-23 May 2008	MIE-2008	Göteborg
30 Aug - 3 Sept 2009	MIE-2009	Sarajevo

Table 3-1: Forthcoming Conferences and Congresses in Healthcare Informatics

In the Human Resources and 'lifelong learning' standardisation areas, the OOA will continue its collaboration with the IEEE-LTSC learning technology standards committee and the HR-XML Consortium Europe, with the mission to promote semantics to the XML standard specifications, starting with competencies and CVs (resume, HR-XML SEP). This will lead to a number of events, and contributions to conferences in this field:

ePortfolio & eLearning

ePortfolio.org is a student-centred platform. Students can create and customize portfolios for academic, career, or personal uses; maintain their plan of study; and share their work, goals, and achievements with advisors, career counselors, and employers. Faculty, departments, and institutions can create portfolio assignments linked to scoring rubrics. Assessment committees can randomly select portfolios, score them with rubrics, and generate assessment reports.

Key-Pal

Key-PAL (Key Skills Portfolio Assisted Learning) is an EU project that aims to establish the relevance and potential impact of the ePortfolio on the development and assessment of key-skills (basic skills / key competencies), This is a critical issue if we want to make Europe become the most competitive and dynamic knowledge-based economy in the world, capable of sustained economic growth with more and better jobs and greater social cohesion.

Co-Drive

Co-Drive is an EU project (Leonardo da Vinci program) that creates a methodology for developing and maintaining vocational competency ontologies, contributing to a European reference model for competency-driven education.

EIfEL (European Institute for E-Learning)

EIfEL is an independent, not-for-profit European professional association whose mission is to support organisations, communities and individuals in building a knowledge economy and a learning society through innovative and reflective practice, continuing professional development and the use of knowledge, information and learning technologies. EIfEL is currently promoting the development and take-up of the electronic portfolio across Europe and worldwide, through its establishment of the European Consortium for the ePortfolio (Europortfolio), leading and contributing to a range of European projects and other initiatives, including the organisation of many international ePortfolio conferences in Europe and beyond as well as the first ePortfolio plugfest to promote interoperability across all learning, and employment services.

TELCERT

TELCERT is a Technology Enhanced Learning research and development project under the European Union's 6th Framework program. Led by a consortium of eLearning providers, research and industry organizations, TELCERT will develop innovative software testing and conformance systems to assure interoperability in eLearning content and technology. Next event associated: LIFEFest in Paris June 2006.

Inflow

INFLOW (Informal Learning Opportunities in the Workplace) aims at promoting acquisition and recognition of key skills / core skills, especially in informal contacts and to develop sustainable partnerships of training, industry and accreditation bodies in the UK and transnationally, to create integrated approaches to recognition and validation.

GEARS

GEARS is an EU project that aims to improve competitiveness of SMEs by ensuring that senior / owner managers in engineering SMEs will be supported in developing strategic plans for their own companies that take account of management theories. This learning will be based on higher-level skill development but will be directly focused on immediate application to their own real business world and therefore will have an immediate direct effect on their competitiveness.

4 Related initiatives: best practices and lessons learned

In this section we describe briefly some related initiatives, focusing in particular on their methodology and modus operandi, and attempt to draw some conclusions about best practices and lessons we can learn from such initiatives which will help in setting up the OOA.

4.1 LIRICS

4.1.1 Overview

The EU e-content project LIRICS (Linguistic Infrastructure for Interoperable Resources and Systems) is concerned with providing ISO ratified standards for language technology to enable the exchange and reuse of multilingual language resources [6]. One of its main aims is to provide models for linguistic annotation and semantic content. LIRICS addresses the needs of today's information and communication society for new standardisation developments, by recognising existing de facto standards and transforming them into de jure International Standards. LIRICS thus aims to:

- Provide ISO ratified standards for language technology to enable the exchange and reuse of multilingual language resources;
- Facilitate the implementation of these standards for end-users by providing an open-source implementation platform, related web services and test suites building on legacy formats, tools and data;
- Gain full industry support and input to the standards development via the Industry Advisory group and demonstration workshops;
- Provide a pay-per-use business model for use by Industry and in particular SMEs validated during the project for the benefit of all actors in the content and language industries.

The LIRICS Consortium brings together leading experts in the field of NLP (Natural Language Processing) and related standards development via participation in ISO committee and National Standardisation committees. The Consortium has strong Industry support and involvement through the 21 members of the LIRICS Industry Advisory Group. LIRICS aims to increase awareness of language engineering standards and promote their take-up on a European scale.

4.1.2 Data Category Registry

The most relevant part of LIRICS for our purposes is the establishment of the Data Category Repository. Because differences in approach among different language resources and individual system objectives inevitably lead to variations in data category definitions and names, the DCR has been established in order to promote the use of uniform data category names and definitions within the same resource domain (e.g. among terminological resources, lexicographical resources, annotated textual corpora, etc.), at least at the interchange level. This contributes to system coherence and enhances the reusability of data. Procedures for defining data categories in a given resource domain also need to be uniform in order to ensure interoperability of individual data category registries.

The creation of a single global data category registry for all types of language resources treated within the TC37 environment provides a unified view on the various applications of such a reference resource. A universal registry should include traditional collections, such as the current ISO 12620:1999, as well as a wide range of other data category selections needed in conjunction with current or future standardization projects. ISO committee TC37 or any of its sub-committees can resolve at any time to officialise specific thematic domains to deal with the management of those selections when appropriate.

4.1.2.1 Requirements

Of particular interest to us are the requirements established by the LIRICS consortium for the implementation of the DCR, in order to fit the needs of the various activities related to the scope of standardization activities within TC37. These are as follows:

- Act as a reference for all the existing or future standards in TC37 related to data modelling or data interchange;
- Be available on-line, free of charge;
- Register existing practices by associating a data category with the way it is implemented in specific projects or initiatives. This may consist in registering various types of encodings, from basic codes ('f' for feminine in Eagles morpho-syntactic descriptions) to actual XML representations;
- Provide names and reference definitions in a variety of languages;
- Describe the usage of a data category in a variety of language settings. This may consist of a specific definition (for instance when the data category has a slightly different application scope), some usage notes, examples, or list of values (e.g. the conceptual domain of /gender/ is {/masculine/, /feminine/} in French, and {/masculine/, /feminine/, /neuter/} in German;
- Describe the usage of a data category in a variety of data processing environments; e.g. some data categories function somewhat differently in machine translation lexica from the way they function in terminological resources or in human-oriented lexicographical resources;
- Associate administrative information to each data category so that it is possible to trace the submission, acceptance or revision of the data category;
- Associate a data category with one or several profiles corresponding to the application domains where the category is relevant (for instance, /Part of speech/ is relevant for POS annotation and lexical representation;
- Provide a mechanism by which a working group in TC37 can submit a group of categories relevant to their scope of activities;
- Be updated on a regular basis by integrating, according to rules to be defined, proposals from experts in the field;
- Provide a personal working space within which experts can upload and publicize their data category proposals, even before they are submitted to the registry;

- Be implemented in accordance with the main principles stated in ISO 11179.

4.1.2.2 Implementation

The following set of rules have been established by the LIRICS consortium for dealing with the actual implementation of the DCR and the decision processes undergone in order to introduce, revise and validate new entries.

The DCR should be implemented as one central data category registry encompassing all its possible activities in the domain of data representation and coding. The registry is the place where data categories are maintained, whether they represent placeholders in a data structure (i.e. “complex data category”, such as /gender/), or values for them (“simple data category”, e.g. /masculine/).

Even if centralized, the data category registry is based on the hypothesis that it can be accessed through thematic views, i.e. domains of activities, which, in the scope of TC37, requires the identification of specialized subsets of the registry. For instance, such a view may correspond to the data categories that can be used in morpho-syntactic annotation, or also the various data categories involved in language coding. Not only do thematic views correspond to ways of accessing the registry, but they also correspond to a basis for submitting and maintaining new data categories. It is indeed anticipated that the management of the registry should not be fully centralized, but based on a structure that will both put together the right expertise within a subfield of linguistic resources and ensure coherence within the registry.

Accordingly, the decision process that leads to the introduction or revision of a data category into the registry shall occur in two steps:

- A *selection* process by which a thematic committee identifies those data categories that are relevant for a certain application field within TC37;
- A *harmonization* process, operated by a DCR board, which guarantees the coherence of new proposals with the scope of the registry and data categories it already contains.

A new data category proposal is then *validated* in the following way:

- The submitted data category shall be defined by the minimum necessary criteria for a data category as outlined in the terms and conditions.
- DC board level shall ballot the DC. If positive votes of more than 70% are received, the status of the data category shall be increased to 'board level standard'. If less than 70% is received the DC shall be given a 'rejected' status and reasons for rejection will be fed back to the proposer.
- Where a rejected DC is proposed again following modification, this shall follow the process for a new DC but shall contain notes regarding the previous submission.

4.1.3 Lessons to be learned

Although on first appearances LIRICS seems quite different from the OOA, there are several things in common and it is clear that there are useful things to be learnt from the modus operandi of LIRICS and from both the successes and difficulties that they have encountered. Where the OOA is seeking to offer a mechanism for validating

proposed ontologies, the LIRICS consortium is seeking rather to offer a mechanism for validating the proposed types of data and structures that resources can have. Where the two coincide is that they are both seeking to provide validation mechanisms in order to facilitate, amongst other things, the interoperability between different tools and resources. Essentially, this all boils down to the fact that as far as the OOA is concerned, ontology providers want to have their ontologies stamped with a recommendation seal and become de facto standards, while LIRICS partners want to progress a stage further and have their data categories and values stamped as ISO norms, which would transform them from de facto to de jure standards. Like the OOA, LIRICS consists of an Industry Advisory Group, thereby maintaining strong industrial links in order to ensure that this is not simply a research phenomenon but has real benefit in the outside world.

We should also take note of the establishment of the DCR and the processes for decision making about submitted entries. The DCR is essentially a list of values: what is interesting for us in the OOA is to look at how the committee agree on these values. LIRICS partners have set up a tool for building the values [7], which enables providers to create their entries via the site, discuss the values to be added, and then validate their choices. The tool can also be used to search for particular values. More information about how the tool operates can be found at [8]. The OOA will similarly require mechanisms for submission, discussion and validation of proposed ontologies, procedures for reaching consensus among the board of evaluators, and so on. It should be noted that LIRICS is very much ongoing work and that not all procedures have been finalized, so it is also possible that a two-way process can ensue and that LIRICS can learn from the OOA. It is currently unclear exactly how LIRICS and the OOA can best collaborate, but plans include joint attendance and participation in meetings and activities, discussion of ideas and best practices between the two groups, and mutual dissemination of information; especially as some Knowledge Web partners are participating in both the OOA and LIRICS.

4.2 XML Clearinghouse

4.2.1 Overview

The XML Clearinghouse for Berlin and Brandenburg realises knowledge transfer as a publicly available service. As instruments it uses conferences, colloquia and public reports.

By monitoring, processing and placement of the development of XML and new Semantic Web technologies and their applications, an exchange of experiences between the regionally active firms and academic institutions is provided. Some of these firms formed the regional hub <xmlcity:berlin>, which brought together and supported competency in XML technologies among companies in the Berlin-Brandenburg region. “Platform for Intelligent Collaboration Portals” (PINK) is the current follow-up activity in which these firms are developing an innovative portal infrastructure using Semantic Web and Web Service technologies.

The results of XML Clearinghouse’s activities are available publicly and therefore indirectly used after a knowledge transfer from research to application. The XML Clearinghouse is funded by the Free University of Berlin in co-operation with the German Ministry of Research (BMBF) from November 2004 until October 2007.

4.2.2 Test Labor

One workpackage within the XML Clearinghouse is the "Test Labor". It is foreseen that within this workpackage:

- a definition of interoperability requirements will be made, together with its operationalisation through the definition of a test suite;
- the testing and preparation of software in the areas of Semantic Web and XML standards for eBusiness will be performed.

The focus of the Test Labor in the XML Clearinghouse is to define a test suite for submitted software to test its interoperability with respect to XML or Semantic Web standards, giving submissions a stamp of approval if they are able to pass the developed tests. While XML Clearinghouse has a much narrower (regional) reach and has a focus on software rather than ontologies, there is some overlap in aims with the OOA. Both activities will determine an evaluation methodology and then apply that methodology to submissions in order to potentially be able to offer them a stamp of approval.

It is the intention of the Free University of Berlin (who is also an OOA partner) to continue to operate the Test Labor past the duration of the XML Clearinghouse, and to maintain it as a facility for the testing and recommendation of XML and Semantic Web software submitted to it by local business.

4.2.3 Conclusions

Given the similarities in both activities, we opened discussions with XML Clearinghouse regarding a possible future collaboration/partnership with the OOA. Our considerations were that the Test Labor could also draw from the results of the OOA in its determination of a Semantic Web ontology test suite and in return could function as part of the OOA as a resource for the performance of ontology evaluations. We spoke with the organisers of XML Clearinghouse and the core firms who are participants of the PINK initiative and who work with the XML Clearinghouse.

These conclusions could be reached from our discussions:

(1) The current group of regional firms in XML Clearinghouse have other specialisations/immediate-focus (e.g. chemistry) than those chosen initially by the OOA;

(2) XML Clearinghouse responds to the needs of the regional industry and currently *technology* evaluation is foreseen rather than ontology *content* evaluation.

(3) The regional firms we spoke with expressed concern about the overhead involved in ontology evaluation activities and questioned the added value of their participation, given the immature state of use of ontologies as present within their enterprises.

It seems that it is too early at this stage to be able to plan for a concrete partnership between both initiatives. We feel that the OOA has a responsibility here to act as a leader within the chosen domains. It is only through such leadership, in which ontologies are promoted within domains by the activities of the OOA Chapters, that

smaller, regional firms such as those which participate in XML Clearinghouse may become more involved in working with ontologies.

4.3 The US National Center for Ontological Research (NCOR)

This centre has been created October 27th, 2005 in Buffalo [9]. The goal of NCOR is to promote the application of scientific methods in ontology research in the United States:

- by establishing cross-disciplinary networks among those individuals and groups involved in ontological research and applications in such a way as to foster a high degree of interaction at the four levels of infrastructure, content, methodology and application
- by fostering, through challenge evaluations and other methodologies, objective measures for the quality (usefulness, usability, reliability ...) of ontologies
- by developing, testing and promoting best practices in ontology research and development, including conformity to reference ontologies and to top-level integration ontologies
- by developing partnerships with institutions in academia, industry and government designed to enable the sharing of expertise and to consolidate best practices
- by organizing and strengthening educational and training programs in ontology
- by organizing outreach programs designed to promote greater public awareness of the importance of high-level ontology research.

Within this centre, a number of Committees have been installed, three of which are relevant in the light of the OOA. Two are discussed hereafter, while the third one, because of its focus on Life Sciences and Health, is covered in section 7.1.1.8.2.3 p.42.

4.3.1 NCOR's Ontology Evaluation Committee

The Evaluation Committee is charged with the responsibility to design and conduct objective, empirical tests that measure the quality of ontologies. This will involve the establishment of an application experiment platform that supports rigorous intrinsic quality tests and experiments performed using ontologies provided by participants from academia, government, and industry that demonstrate the potential value of ontology-based technologies. It is intended for the results of these evaluations to inform the development of NCOR best practices for ontological engineering. For its tasks, the committee will leverage the collective experience of its members with prior evaluations of semantic technology through open competitions such as the Text Retrieval Conference (TREC), Message Understanding Conference (MUC), and Information Interpretation and Integration Conference (I3CON).

The Ontology Evaluation Committee has an Ontology Evaluation Working Area where resources, plans and activities are shared.

Convenors are Steven Ray (NIST) and Todd Hughes (Lockheed Martin Advanced Technology Laboratories).

Members are Susan Golden (Nervana), Inderjeet Mani (Mitre), Robin A. McEntire (GlaxoSmithKline), and Conor Shankey (Visual Knowledge Software Inc.)

4.3.2 Ontology Outreach Committee

The tasks of the Outreach Committee include:

- providing information to the ontology user and developer community of the importance of NCOR's goals of advancing the scientific method in ontology through media and events (writing and speaking for general audiences as well as technical and domain experts;
- advancing NCOR's research methodology by gathering information from organizations about their needs, activities, and investment decision-making processes;
- recruitment of partners with the goal of achieving a critical mass of major public sector agencies and academic institutions, private sector companies, system integrators, and consultancies as members of the NCOR consortium;
- establish relations with partners who will provide resources for funding of (joint) research, communications, events, public information, and receive in their turn predefined benefits such as:
    - certain publications or other research information from NCOR,
    - access to NCOR team expertise -- phone advice, workshops, training, on-site consultation or professional services, etc.
    - advisory participation in NCOR activities or events

Convenors are Werner Ceusters (European Centre for Ontological Research) and Mills Davis (TopQuadrant).

4.3.3 Conclusion

NCOR seems to have taken a real jump start with active participation from major research institutes and large companies in the US. It is definitely worthwhile to monitor its further activities and to learn from these in setting up the KnowledgeWeb OOA in Europe.

4.4 OASIS

Somehow relevant to the OOA is OASIS [10]. OASIS (Organization for the Advancement of Structured Information Standards) is a not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards. The consortium produces more Web services standards than any other organization along with standards for security, e-business, and standardization efforts in the public sector and for application-specific markets. Founded in 1993, OASIS has more than 5,000 participants representing over 600 organizations and individual members in 100 countries.

OASIS is distinguished by its transparent governance and operating procedures. Members themselves set the OASIS technical agenda, using a lightweight process expressly designed to promote industry consensus and unite disparate efforts. Completed work is ratified by open ballot. Governance is accountable and unrestricted. Officers of both the OASIS Board of Directors and Technical Advisory Board are chosen by democratic election to serve two-year terms. Consortium leadership is based on individual merit and is not tied to financial contribution, corporate standing, or special appointment.

The Consortium hosts two of the most widely respected information portals on XML and Web services standards, Cover Pages and XML.org. OASIS Member Sections include CGM Open, DCML, LegalXML, PKI, and UDDI.

Thus far, OASIS did not participate directly in ontology related business what is probably a mistake. Therefore, it might turn out to be one of the prominent players in the field to which the OOA should offer his expertise.

5 OOA operational workplan

Whereas Semantic Web *technology* is virtually sector-independent, ontologies themselves are usually very domain-dependent [11]. This motivates a division of labour focused upon specific market segments. Furthermore, in order for the OOA to be self-sustainable, it should account for the needs in practice. Therefore, industrial parties should play the steering role in the OOA. Looking at other standardization/recommendation bodies one can see that their main players are companies. For the OOA, we aim to involve industrial parties and organize their participation into committees of vertical sectors, such as legal, financial, human resources, healthcare, etc. Participants should be leaders in their industrial sector. Each committee will be responsible for ontology recommendation and outreaching in its sector. Notice that this somewhat simplifies the ontology evaluation process as each sector may decide its evaluation criteria and needs.

Before discussing how vertical sectors can be identified, in the following we overview the main added values of the OOA.

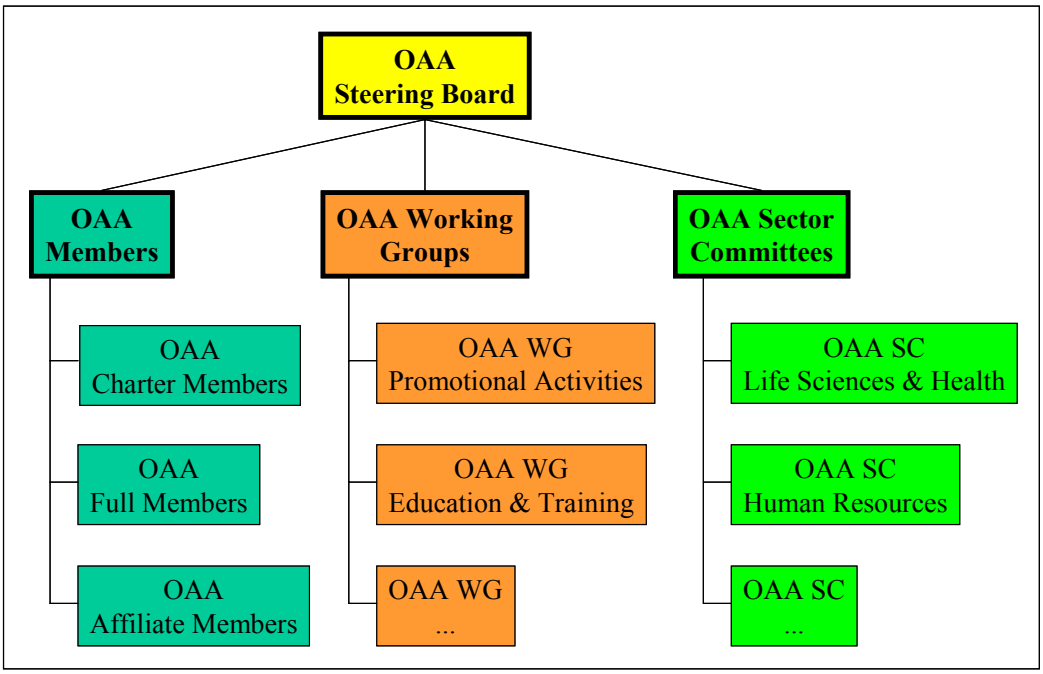


Figure 5-1: Envisaged organizational structure of the OOA.

5.1 OOA added values

5.1.1 Internal added values

The OOA will mainly be populated through the establishment of different types of memberships[2]. On the one hand, the type of membership depends on certain eligibility criteria, and on the other hand will give certain advantages to players enjoying a specific membership type.

5.1.1.1 Charter Members

Charter Members are those that will support the creating and defining of the organisation from its inception.

The benefits of Charter Membership include:

- The opportunity to nominate representatives to sit on the Steering Board.
- The opportunity to have the nominated attendees run as inaugural committee chairs.
- The waiving of the joining fee that will be applicable to Full Members, i.e. those joining after the formal constitution of the OOA.
- The ability to take a lead role in defining the creation of the organisation including:
    - its articles of constitution;
    - future investment plans;
    - product development plans;
    - implementation experience and development.

In addition to the opportunity to define and manage the OAA, Charter Members will automatically derive the benefits that will come to Full Members as they join the organisation after its creation.

5.1.1.2 Full members

These are members who join after the OOA's inception and who have an outstanding reputation of involvement in ontology research or application. Detailed eligibility criteria will be defined by the Steering Board

The benefits of Full Membership include:

- full access and unlimited usage rights to all developments of the OOA;
- equal rights in nominating members of the Steering Board twelve months after joining;
- equal rights in nominating domain sector committee members twelve months after joining;
- immediate and unlimited access and contribution to all working groups;

[2] *Membership* is here to be understood in general sense, and not in the legal sense as legal members of an association.

- equal voting rights at all levels of the organisation;
- equality in membership fee rates using metrics to be established by the Steering Board.

5.1.1.3 Affiliate Membership

Affiliate Members of the OOA will be individuals or organisations who do not qualify for any other membership type.

The benefits of Affiliate Membership will include:

- the personal or organisational use of OOA materials;
- full access to the Working Groups of the OOA;
- the opportunity to have Committee representation through chairing a Working Group.

5.1.2 External added values

5.1.2.1 OOA added values for ontology developers

The OOA can be seen as an analogy of the scientific publication system, where authors submit their papers to be reviewed by a conference committee, and accepted papers are published and promoted. In the same way, ontology developers will submit their ontologies to a certain sector committee for evaluation, and "good" ontologies will be recommended and promoted by the sector committee. Developers of "rejected" ontologies will also get the benefit of the evaluators' comments.

From a business viewpoint, as recommended ontologies will become de facto standards, their developers will also get business values and benefits as their ontologies will be adopted by other users in their sector.

5.1.2.2 OOA added values for the market

The community benefit of the OOA is that efforts in each market sector are organized. Ontologies recommended by sector experts become *de facto* standards for that sector. The OOA promotion and outreaching activities improve the community innovation and productivity at large.

5.2 Identifying vertical market sectors.

A vertical sector is a domain or market sector that has its own needs and specifics, such as finance, legal, Health and Life science, human resources, automobile, news, sports, telecommunication, translation, customer relationship management (CRM), etc. It is not necessary for the OOA to cover all sectors. The OOA, especially in its first steps, will focus on the sectors where the need for ontologies and the Semantic Web technology is recognized. In later phases, a committee of upper level ontologies (as a horizontal committee) is needed to serve the general ontological needs of all sectors.

The need for ontologies in some sectors is not necessarily obvious for all users. For example, we have found that the added-value of using ontologies for competency modeling and job descriptions is not well understood in the human resources sector.

Instead, this community has been trying to model and standardize competencies in XML, i.e. syntax vs. semantics. See for example the work done in HR-XML. Likewise, the financial sector has defined an XML standard for financial products.

To obtain feedback from industry about whether there is a demand for the OOA to exist, we have established two committee sectors:

1) Health and Life science, including bioinformatics, biomedicine, etc.

2) Human resources including employment, competencies, jobs, etc.

Our experience and promising results will be reported later in this document. Furthermore, we shall present a brief analysis of other sectors, and identify what ontologies are being used for.

5.3 OOA Committees

5.3.1 Set up a committee for each domain sector

Players active with respect to ontology development, use, or education in a sector will be invited to join the committee (10-20 members) for their respective sector. Both users and developers of ontologies in the domain will be solicited for active participation. They will be recruited from the most active industrial enterprises primarily, but also from research institutes that play a leading role in ontology development, as well as in the validation.

From participants, little or no technical knowledge is required. Instead they should bring a deep understanding of their own sector and its interoperability needs (different languages, meaning, and processes). They should also show willingness to promote and disseminate the potential (awareness, services, tools) made available through the KnowledgeWeb OOA, and, after being trained, are expected to help introduce the use of semantics and ontologies for use in their professional environment.

For each sector committee, the following tasks have to be performed:

- Appointment of a Committee Chair;
- Promotion of the committee's activities to potentially interested parties;
- Adjustment of the working principles to match better the sector's requirements;
- Management of the ontology certification procedures. This includes providing some training to the evaluators and might require the installation of a Technical Board which is responsible for teaching domain experts in ontology matters from a technical perspective, and to support them with technical consultancy;
- Promotion of recommended ontologies by means of public demonstration events, newsletters, the organization of ontology-content-oriented conferences/meetings, etc.

Ontology certification itself might be done by members of the committee, or subcontracted to third parties (e.g. reviewers/validators). When an ontology with its documentation is submitted to a certain sector committee, its goals, requirements,

scope, etc. should be clearly stated. The sector committee then verifies whether this ontology is in concordance with these goals and requirements as claimed.

5.3.2 Setup an OOA steering board

The Steering Board is responsible for the overall functioning of the OOA. It will consists of 5 to 10 members, at least one per vertical sector. One important issue to be determined is the status of incorporation and future financial means.

5.4 OOA Working groups

Working groups will be set up to study specific issues with the objective to prepare detailed plans that then can be voted upon by the Steering Board, or, if applicable, the Sectorial Committees.

Issues that might lead to the creation of a working group include:

1. promotional activities involving the use of standards-based tools to create, manage, and use knowledge sources, services, and metadata to assist research communities to develop ontologies that will allow researchers and clinicians in discovering, assimilating, and using new knowledge;
2. promotional activities towards the broad and open dissemination of software, computer-based knowledge content, ontologies and services developed under certain quality assurance principles, including the development of a Web portal for existing and future knowledge sources and metadata that may be distributed anywhere across the Internet, allowing scientists to annotate, query, and analyse data;
3. providing access to a laboratory for the ongoing demonstration and evolution of knowledge sources, metadata, ontologies and associated tools and applications in the context of driving problems;
4. monitoring cross-disciplinary educational opportunities for training domain experts and computer scientists in the best practices of ontology building;
5. the implementation of effective management and advisory structures.

6 OOA infrastructure and administration

6.1 Requirements

6.1.1 Legal requirements

It has been decided by the Knowledge Web partners that the OOA will be a non-profit organization. However the decision whether the OOA will be an independent organization (i.e. legal entity), or a part of an existing organization (such as OASIS.org, CEN/ISSS, etc.), is still open.

The first scenario gives more independence and freedom on the OOA policies and producers. The second scenario empowers the OOA with the reputation of the existing organization, however it is not easy to convene such organizations to extend their missions, especially at this early phase of the OOA.

Our approach, for both scenarios, is to establish a collaboration link with all related organizations, i.e. a bottom up approach. For example, in our first OOA meeting, (Crete June 2005) we invited 1) the IEEE chair on "competencies", 2) the chair of the HR-XML Europe, 3) a representative of the Dutch employment agency, 4) the Executive Director of the European Centre for Ontological Research. The OOA (human resources sector) in cooperation with the HR-XM Europe is organizes regular CEN/ISSS workshops on ontology based competency modelling.

In case the OOA is established as an independent organization, the OOA will be incorporated in legal form as a Belgian not-for-profit organization, known as an ASBL (Association sans but lucrative), on the legal basis of an international not-for-profit organization (INPA). This requires us to prepare some formal documents, such as the OOA work program and OOA bylaws to define the operation processes for voting, membership fees, and rights and obligations of members, etc. These documents will proposed to the Knowledge Web Board and members for approval, and it will form the basis of the OOA operations.

6.1.1.1 Requirements for the OOA as a Belgian not-for-profit organization (NPO).

With respect to domestic and international not-for-profit organizations, Belgium modified (by a law of 2 May 2002) the older acts of 27 June 1921 and 25 October 1919, and later supplemented the law by several royal decrees (15 May, 26 June and 19 December 2003). All these texts can be found on [12] in French and Dutch.

The new legislation provides for four categories of not-for-profit organizations, having their seat in Belgium: two for ***associations***, and two for ***foundations***.

The NPOs become a separate legal entity as soon as they have their charter registered within a public file (for an ASBL and private foundation), or after a royal decree (for an AISBL and foundation of public interest).

All NPOs have to publish annual accounts. They have to send these accounts to an official public file: either the tribunal (for small ASBL and small private foundations), or the ministry (for foundations of public interest), or the Central Bank of Belgium (for big ASBL and big private foundations). When the accounts have been published according to the latter system, they are very easily accessible (even via the internet).

There are two different accounting regimes, one for small NPOs, that have to record their cash expenses and cash receipts, but that have also to draft a so-called simplified balance of their assets and debts, and one for the big NPOs, that have to satisfy the accounting rules provided for by the Belgian company law, with some peculiarities.

A NPO becomes "big" when it exceeds, for one year, two of the three following criteria: 5 employees (in equivalent full time), 250,000 € of usual cash receipts, 1,000,000 € for the total of the balance sheet.

Some NPOs have to submit their accounts to a check by a legal auditor. This is the case for NPOs having over 100 employees, or exceeding two of the three following criteria: 50 employees, over 6,250,000 € of usual cash receipts, or over 3,125,000 € for the total of the balance sheet. The legal auditor is appointed by the general assembly of the members (ASBL) or by the Board (AISBL and foundations). He has not only to check the accounts, but also the legality of the expenses of the association in relation with the law and with the charter of the association.

In all kinds of NPOs, there is a principle of limited liability towards third parties, both in favour of the members (if any) and for the Board members (except if they commit a misdemeanour).

As opposed to the company Board members, association Board members are however not concerned by a set of special criminal rules. But of course, they are concerned by the general rules included in the Criminal Code of Belgium.

The judge has the power to liquidate associations that do not strictly respect the law.

A judge's decision is moreover the only way to liquidate a foundation.

6.1.1.1.1 Not for profit associations

There are two types of not for profit associations: national and international ones. The criteria to be fulfilled for setting up any of these are listed below.

Both types have some common working principles. Once functional, an association must have some kind of general meetings of the members, minimally once yearly.

Within an association, the rights of the members are very strictly enforced. Any appointment of a Board member has to be done by the general assembly of the members. Any exclusion of a member has to be decided by a special majority within the general assembly. The general assembly may change the charter of the association, without unanimous consent by the members. All members of an association have the right to read and investigate the minutes of the Board, as well as the financial documentation of the association.

6.1.1.1.1.1 Not for Profit Association (proper; VZW/ASBL)

The most common framework, a ***VZW/ASBL***, requires the association to fulfil mainly the following criteria:

- the purpose of the association is not a commercial activity,
- there is no will from the members to take some material profit from their participation to the association,
- there are at least three members (being Belgian or foreigners),

- in case of liquidation, the remaining balance has to be used for a not-for-profit goal.

The Board may not be a single person, and may not have as many members as the general assembly has: there must be at least one more member of the general assembly with regard to the membership of the Board. The association can be converted into a not-for-profit company when the need to set up commercial activities is identified but still without any will to give material benefit to the associates (or just a little interest on their personal capital contribution).

6.1.1.1.1.2 International Not for Profit Association (IVZW/AISBL)

Another framework, the ***IVZW/AISBL***, has many common features with the previous one, the differences being having the word "international" in its denomination, and the first criterion as stated above changed into: “the purpose of the association is not a commercial activity, but a not-for-profit goal with an international dimension.” This international dimension will be clearer if the membership is itself international, but this is not a condition. Conversely, there is no request that the management of the AISBL includes at least one Belgian member.

6.1.1.1.2 Foundations

The foundations may work only with a Board: foundations have no members. The charter of a foundation must include a procedure for solving any conflict of interests arising within the foundation.

6.1.1.1.2.1 Private foundation

A “***private foundation***”, the third framework, has to fulfil the following criteria:

- the foundation has received from its founder(s) a gift in consideration of a not-for-profit goal. There is no minimum value for these initial assets and the founder is not necessarily a natural person: another association, or even a company, can be a valid founder;
- there is no will from the founders or Board members, and not even from third parties, to take some material profit from the foundation. However, there can be a material interest for third parties, if this is required for achieving the not-for-profit goal of the foundation. For instance, a foundation established for helping a disabled child may give some support to this child;
- the foundation will need an act by a public notary;
- in case of liquidation, the remaining balance has to be used for a not-for-profit goal, but the founder or his legal successors may take back the gift or the value of the gift the founder has made while establishing the foundation.

A private foundation can be converted to a foundation of public interest.

6.1.1.1.2.2 Foundation of Public Interest

The fourth framework, a ***foundation of public interest***, has to fulfil the criteria set for a private foundation, and must have a philanthropical, philosophical, religious, scientific, artistic, educational or cultural aim.

6.1.2 Staff Requirements

To support the administration function of the OOA committees, the OOA should have a head office. This office will be located in the Brussels region. As minimum, one general secretariat should be employed to serve the administration and technical management of the OOA.

6.1.3 Financial requirements

At the foundation phase, the OOA will be supported financially by the Knowledge Web project. Later the OOA will operate depending on the membership fees, ontology evaluation and recommendation fees, income of organized activities, etc.

6.1.4 Technical requirements

To serve the OOA activities, an OOA website is required. This website will be used as the main portal of the OOA, for promoting the OOA activities, announcements, mailing lists, etc. Another functionality of the OOA website is an ontology repository, which will be used for publishing metadata about the ontologies that will be recommended by the OOA sector committees. For this, we shall use the ontology repository that we have developed in WP12.

6.2 Decisions taken

During the past year, the relevant consortium partners have discussed the requirements described above. Since many options exist, and the pros and cons have not all been assessed yet, formal decisions have thus far not been taken, except for the OOA website and repository. In the following, we overview the main functionality of the repository and its adaptation to suit the OOA needs.

6.2.1 Summary of the OOA website and repository specification

One specific key idea in the OOA is to host a repository of ontologies in so far this is in line with the IPR and distribution principles endorsed by the contributing authors. To maximise access and reusability of these ontologies, we decided to use a portal infrastructure that exploits a coherent set of metadata to describe these ontologies.

The metadata portal stores information according to a metadata vocabulary, OMV, which has been proposed as the Knowledge Web metadata "standard" based on discussions and agreement amongst Knowledge Web members. As an ontology metadata standard, OMV distinguishes between an ontology base and an ontology document and thus leads to an efficient mechanism for tracking several versions and evolvements of ontologies as well as for different representations of one knowledge model in different ontology languages [13, 14].

As the importance of metadata increases with the number of ontologies stored, the demand for supporting technologies like storage and access techniques, ontology sharing and interoperability become important as well.

For these reason, Oyster (a P2P ontology sharing software that has been developed within Knowledge Web and that supports the OMV above) will be offered for the OOA portal community.

Oyster's functionalities can be separated into two groups, based on their usage. Basic functionalities are provided to every user who accesses the OOA portal whereas sophisticated functionalities are provided for both reviewers and administrators.

A *visitor* is an anonymous user who is allowed to browse the public content of the portal. Search and export are available to any visitor without being registered to the repository.

Since providing new metadata is based on a certain community confidence, a visitor has to register at the portal to become a registered *user*. This is done by completing an application form on the website. Users can customise their portal, e.g. the content of their start-page or their bookmarks. If a user wants to submit metadata to the portal, this submission has to be reviewed before it is published.

7 First Steps: Establishing two OOA committee sectors for HR and HC

In order to collect feedback from industry on the OOA goals and requirements, we select two key industry sectors which are potential first adopters of ontology-based technologies: 1) Health and Life Sciences, including bioinformatics, biomedicine, etc. 2) Human Resources, including employment, competencies, jobs, etc. In this section we discuss these sectors, and report our activities to date, experiences and promising results. An analysis of other sectors which we aim to target in the future is presented in the next section.

7.1 Establishing the OOA-HC (Health and Life Sciences sector)

7.1.1 Sector analysis

Health and Life Sciences, as explained further down, is one of the first domains in which the need for ontologies to advance science on the one hand, and to make better applications on the other hand, has been understood. However, there are not that many success stories. The main reason is that this insight was limited to technology providers mainly, while they were not enough aware of the cultural and social rules that govern this domain. The following analysis intends to give a picture of the scene such that the OOA HC is well prepared to finally make a difference.

7.1.1.1 Introduction

Biomedical information systems are nowadays considered to be a cornerstone for the practice of medicine. That has been quite different in the past. Originally, and maybe still in some underdeveloped rural areas, medicine was a matter of pagan theology and rites, until philosophers in about 400 BC started to argue for the application of rigorous rules within any intellectual discipline [15]. The vision of medicine as an art was installed by Hippocrates. “What makes medicine an art”, he wrote, “and what makes the *technai* (derived from the Greek word for art) intellectually valuable to man, is that the practice of these disciplines is not left to chance but is indeed governed by specific principles” [16]. But also this vision, although still highly appreciated by many practising physicians today, did not survive as it became clear that deduction on the basis of a scientific theory is perhaps sufficient to come to a diagnosis, but that giving adequate treatment was, certainly at the time this insight prevailed, i.e. some 300 years ago, a matter of induction on the basis of carefully observed phenomena on the side of individual patients [17]. For example, the healing effect of certain herbs was statistically induced from the observations that patients with specific disorders improved when being treated with particular herbs, and not on the basis of a theory that explained on physiologic and biochemical grounds why healing could be expected.

It is the latter insight that became the basis for what in 1992 became introduced as *Evidence Based Medicine*. David Sackett, one of the proposers of the paradigm, defined Evidence Based Medicine as “the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical experience with the best available external clinical evidence from systematic research” [18]. Under this paradigm, evidence is not supposed just to be based on personal

experience, which is often bias-prone, but rather on the experience gained by the entire medical community.

The key question became then: how to bring the relevant evidence in a timely fashion to the attention of the practitioner? Originally, evidence was a scarce resource. With increasing technological development, however, new techniques in data collection, in addition to the exponential advance of science that results from analysing this data, causes quite a different sort of problem: that of information overload [19]. This problem can have severe consequences: an estimated 2,886,800 cases of cancer were diagnosed in Europe in 2004, with some 1,711,000 cancer deaths [20]. Many of these deaths could have been prevented if cancer had been diagnosed at an earlier stage, or if appropriate therapies had been available. A study in the UK showed that a significant number of patients referred non-urgently who are eventually diagnosed with cancer experience times for treatment which are much longer than would be necessary if those resources and knowledge which are already available could be applied in an efficient way [21].

This seems surprising in an era in which *Electronic Health Record systems* (EHR systems), i.e. computer systems used to store and retrieve data about an individual patient's symptoms, disorders and treatments throughout his lifetime, have become a routine instrument for patient management in many European Member States. The problem is that EHR systems function to a large degree in isolation, on individual physicians' desks or in individual healthcare institutions. It could be solved if these systems could be made *semantically interoperable.* And this does not hold just for EHR systems, but for all information systems that are used in the Healthcare and Life Sciences sector.

7.1.1.2 From *bio-informatics* and *medical informatics* to *biomedical informatics*

Biomedical information systems are information systems pertaining to information about entities in the biomedical domain. The term *biomedicine* appeared already in Dorland's Medical Dictionary in 1923, which described it as "*clinical medicine based on the principles of physiology and biochemistry*" [22]. Meanwhile, cellular biology, physiology and biochemistry have become so influential on the way medicine is practised, that biomedicine is no longer considered to be a discipline within medicine, but to be medicine as it is practiced today, at least in the West [23]. That is to say, the term 'biomedicine' underwent a shift in meaning: today it refers to the medical practice of the past in addition to 'biomedicine' as described in Dorland's Medical Dictionary.

Surprisingly, no similar shift in meaning occurred for the term *medical informatics* in contrast to *biomedical informatics*. Some consider medical informatics to be "the field concerned with the management and use of information in health and biomedicine" [24], hence adopting the switch from medicine to biomedicine, while others describe it as the discipline that "has been focused on the intersection between computer science and clinical medicine" [25], hence *not* including that which the biomolecular sciences bring to the field. This difference allows these authors then to describe 'biomedical informatics' as "the framework rising from the synergy between bio-informatics and medical informatics providing for developing and sharing new biomedical knowledge"; *bio-informatics* then being described as "the discipline developed to handle large amounts of data, mainly sequences, generated in the

laboratories" [26]. The latter view is the most appropriate, since it excludes from the applications of bioinformatics systems for electronic prescriptions, billing, appointment scheduling, and so on. Such uses do indeed belong to the realm of medical informatics, but they do not build further upon what was gained through the completion of the Human Genome Project, which triggered the closer cohesion of bioinformatics and medical informatics.

7.1.1.3 Challenges for Biomedical Informatics

The challenge for biomedical informatics in this context is thus threefold:

(1) It must develop methodologies that allow rapid and accurate analysis of genomic and proteomic data.

(2) It must create a unified framework for incorporating various types of data (such as gene expression and protein expression). Such a framework will enable clinicians and researchers to establish relevant relationships between the entities involved in their respective disciplines, thus facilitating communication between disciplines.

(3) It must bridge the gap between biological knowledge and clinical therapy by linking bioinformatics tools and resources to clinical patient data.

While a huge international effort has been made to realize (1) [27, 28], large portions of (2) and (3) remain unexplored [29]. To realize (2), software tools are needed which can integrate data on gene transcription, proteomics, interaction networks, and small chemical molecules. Integration of this data will make it possible to retrieve information from existing databases in ways which will help clinicians and researchers establish the relevant relationships between the entities involved. Biological data sources typically include information about basic cellular processes in the form of gene sequences and expression profiles, protein interactions, known polymorphisms, protein structures, or the results of ligand-binding experiments. Epidemiology oriented databases might contain information about the prevalence of certain diseases and statistical correlations with phenomena observed in patients exhibiting these diseases. Combining these databases might reveal the pathophysiological processes that are responsible for the observed correlations.

One main obstacle to (2) is that the sorts of data contained in biomedical information systems are every bit as varied as the subject matter of biomedicine itself, ranging from the smallest elements at molecular level (genes, proteins, biochemicals, and so forth) over subcellular elements, cells, organs, and the human body, to entire populations [30]. This variation has been dealt with by compartmentalizing the data in different databases (for example, a database of physiological pathways and a database of gene expressions). Another obstacle to (2) is inconsistencies in the language used to describe molecules, pathways and so forth.

If the OOA HC intends to become successful in orienting itself towards the biomedical informatics domain, it should understand these challenges deeply and focus part of its activities in identifying how it can assist in helping them to overcome.

7.1.1.4 Stakeholders in Biomedical Informatics

Healthcare and Life Sciences, the domain in which biomedical information systems are used, is a vast area with many stakeholders, most of them having already taken important steps in the direction of semantic interoperability through ontologies [31].

Global pharmaceutical R&D organizations generate large amounts of information of many different kinds which are typically securely stored in a database with some sort of query interface so that users can recall the information and use it to make decisions. However, several factors interfere with the ability of these organisations to get the most of their research investments. Semantic Web technologies could be used here for the effective communication of experimental protocols, descriptions of model systems, statistical criteria for data acceptability, and many other critical elements amongst knowledge bases. Another use of SW technologies is that of synthesising results from the various knowledge bases into a holistic picture of physiology as in *translational research*, defined as the "multidisciplinary scientific efforts directed at accelerating therapy development, i.e. moving basic discoveries into the clinic more efficiently".

Consumer health informatics is the emerging science at the crossroads of health informatics and public health which deals with investigating determinants, conditions, elements, models, and processes to design, implement, and maximise the effectiveness of computerised information as well as telecommunication and network systems for consumers. One of the central topics of consumer health informatics is how to guide consumers to quality health information. Technology for producing and distributing information is useless without some way to locate, filter, organise and summarise it.

Also governments, both at national and European level, are facing many health-related challenges:

- rising demand for health and social services, due to an ageing population and higher income and educational levels. In particular, by 2051, close to 40% of the EU's population will be more than 65 years old;
- the increasing expectations of citizens who want the best care available, and at the same time to experience a reduction in inequalities in access to good health care;
- increasing mobility of patients and health professionals within a better functioning internal market;
- the need to reduce the so-called 'disease burden', and to respond to emerging disease risks (for example, new communicable diseases like SARS);
- the difficulties experienced by public authorities in matching investment in technology with investment in the complex organisational changes needed to exploit its potential;
- the need to limit occupational accidents and diseases, to reinforce well-being at work and to address new forms of work-related diseases;

- management of huge amounts of health information that need to be available securely, accessibly, and in a timely manner at the point of need, processed efficiently for administrative purposes;
- the need to provide the best possible health care under limited budgetary conditions.

From this analysis, it is clear that for the OOA HC to be successful, the needs of all these stakeholders need to be taken into account appropriately. For this reason, as explained further down, initiatives have been taken to make these stakeholders participate in the design phase of the sector committee.

7.1.1.5 Ontologies in Life Science and Health

Ontologies are currently a hot research topic in Life Science and Health, their main purpose being, as it is hoped, to assure semantic interoperability of systems. More than in other domains, it seems, there is a divide between researchers approaching the issue from an information science and software engineering perspective, and those taking a philosophical stance.

The former group understand by 'ontology' a formal representation of a (partial) intensional definition of a conceptualization of an application domain [32], i.e. as a first order vocabulary with semantically precise and formally defined logical terms that stand for concepts and their inter-relationships of an application domain, and thus more as a knowledge representation (and as such not different from the old frame-based or semantic-network variations). This community works with minimalist "models" that then are used as templates to look at those parts of reality that fit the model (hence you can only see what the model allows you to see). The models are usually implemented by means of some form of description logic (DL).

The key characteristic features of description logics reside in the constructs for establishing relationships between *concepts* by means of *roles* [33]. Concepts are given a set-theoretic interpretation: a concept is interpreted as a set of individuals, and roles are interpreted as sets of pairs of individuals. The domain of interpretation can be chosen arbitrarily, and it can be infinite. In this context, it is important to understand, as stated in [34:pp.30-31] that '*Model-theoretic semantics does not pretend, and has no way to determine what certain words and statements "really" mean. (...) It [= model theoretic semantics] offers no help in making the connection between the model (the abstract structure) and the real world*'.

It is this lack of explicit reference that disturbs those who take an analytical-philosophical stance, and for whom the term ontology denotes rather a representation of reality. This community argues that an ontology should correspond to reality itself in a manner that maximises descriptive adequacy within the constraints of formal rigor and computational usefulness. By 'ontology' they mean: a representation of some pre-existing domain of reality which (1) reflects the properties of the objects within its domain in such a way that there is a systematic correlation between reality and the representation itself, (2) is intelligible to a domain expert, and (3) is formalised in a way that allows it to support automatic information processing. By 'terminology', they then mean a set of terms that within the linguistic and professional community by which they are used, are accepted designators for the entities represented in the ontology. This is a more precise notion than the one endorsed in

linguistic-terminological cycles in which terminologies are perceived as a class of systems, either in software or on paper, that contain the terms which specialists in a specific domain are supposed to use when exchanging information. Their purpose is twofold: to allow an unambiguous understanding of what is conveyed, and to stabilise as much as possible the terminology within a specific domain. In this notion, there is not the requirement that there ought to be for each term a referent in reality.

With respect to *patient data* for instance, an ontology enables explicit references to be made to the real instances (particular cases) to which the statements in the electronic health record may refer only implicitly, and to describe in a formally rigorous way the relationships that occur between these instances [35]. A good biomedical ontology thus reflects the most general categories in reality, i.e. those categories which contain all of the categories into which biomedical data is organised. This makes it possible to link together the general terms that are provided by clinical or biological terminologies.

Unfortunately, most ontologies in biomedicine are marked by a number of serious defects when assessed in light of their conformity to both terminological and ontological principles [36, 37, 38, 39, 40, 41, 42]. This means that much of the information formulated using such ontologies remains implicit to both human interpreters and software tools. Vital opportunities for enabling access to the information in such systems are thereby wasted. These defects manifest themselves in difficulties encountered when the underlying resources are used in biomedical research. Such defects are destined to raise increasingly serious obstacles to the automatic integration of biomedical information in the future, and thus they present an urgent challenge to research.

The major overarching challenge to be met by ontology is thus two-fold: (1) to bridge the gap between clinical research conclusions and the need to make personal decisions in healthcare, and (2) to bridge the gap between data models evolved separately in the two discrete worlds of healthcare and bioinformatics.

7.1.1.6 Main ontologies in Life Sciences and Health

The following sections give an overview of the most relevant ontologies (or ontology-like artefacts) in this domain. They are ideal candidates for being the first to be validated or certified once the OOA HC is up and running.

7.1.1.6.1 DL-supported Concept-based ontologies

7.1.1.6.1.1 SNOMED-CT

SNOMED-CT® [43] is developed by the College of American Pathologists [44], and grew out of the merger, expansion, and restructuring of SNOMED RT® (Reference Terminology) [45] and the United Kingdom National Health Service Clinical Terms (also known as the Read Codes) [46]. As of July 2005, SNOMED-CT contains over 366,170 health care concepts organised into hierarchies, with approximately 1.46 million semantic relationships between them, and more than 993,420 terms. It is available in English and Spanish language editions.

The main merits of SNOMED-CT for clinical documentation are its broad terminological coverage (this has been shown repeatedly in the course of its development and in various areas such as ophthalmology [47], diagnosis and problem

list entries [48, 49], nursing [50], therapeutics in general [51] or indications for drug treatment in oncology in particular [52]), combined with a fairly successful maintenance methodology to avoid multiple representations of the same entities and to provide some stability in meaning. Also, its efforts in mapping to other terminological systems are useful. Furthermore, sufficient resources are available to update the system in the forthcoming years.

From the above it should not be inferred that we endorse all aspects of SNOMED-CT. There are serious problems associated with using SNOMED-CT as an ontology instead of a terminology, i.e. for reasoning. SNOMED-CT organises terms according to a minimalist model and (during the design phase) lets a description logic compute whether statements are consistent with the model. This does not guarantee however that statements are consistent with reality nor is it a safeguard against semantic inadequacy of the labels: often, users reading a term (e.g. via a browser) attach to it a meaning that is not intended by the system (which can be verified by analysing in detail the formal statements through which the term is defined). In [39], mistakes are reported such as improper assignment of relationships, including the so important ISA-relationship, shifts in formal meaning from one version to another (but with preservation of the semantics at face value), lack of a sound mereology for anatomy, and so forth. For this reason, SNOMED-CT's relational organisation is still best conceived as a convenient mechanism for browsing through the terminology in order to find better descriptors, but not as a representation of how the corresponding instances are related together in reality [53].

7.1.1.6.1.2 National Cancer Institute's Thesaurus

The National Cancer Institute's Thesaurus (NCIT) is a cancer research nomenclature with features resembling those of an ontology in the sense in which this term is used in the current bioinformatics literature: a controlled vocabulary organised as a structured list of terms and definitions. It was created by the National Cancer Institute's Center for Bioinformatics and Office of Cancer Communications for use not only by the Institute's own researchers but also by the cancer research community as a whole. Its main goals are:

1) to provide a science-based terminology for cancer that is up-to-date, comprehensive, and reflective of the best current understanding;

2) to make use of current terminology "best practices" to relate relevant concepts to one another in a formal structure, so that computers as well as humans can use the Thesaurus for a variety of purposes, including the support of automatic reasoning;

3) to speed up the introduction of new concepts and new relationships in response to the emerging needs of basic researchers, clinical trials, information services and other users [54].

The NCIT serves several functions, including annotation of the data in the NCI's repositories and search and retrieval operations applied to these repositories. It is also linked to other information resources, including both internal NCI systems such as caCore, caBIO and MGED and also external systems such as the Gene Ontology and SNOMED-CT. It is part of the Open Biomedical Ontologies library [55] and is also available under Open Source License on the NCI download area [56]. This makes it

an important candidate for the delivery of vocabulary services in cancer-related biomedical informatics applications in the future.

NCIT is a thesaurus, and one can thus expect it to be of use to researchers engaged in biomedical database annotations. At the same time its ontological underpinnings are designed to open up the possibility of more complex uses in automatic indexing and bibliographic retrieval and in linking together heterogeneous resources created by institutions external to the NCI. It is this last potential application that is receiving most attention in the biomedical research community.

Recently, version 04.08b of the NCIT, released on August 2, 2004 and made publicly available through the NCI website [57], was studied along three lines with respect to quality [42]:

1) conformity with relevant terminological standards put forward by ISO;
2) ontological principles; and
3) appropriateness of OWL as a knowledge exchange format.

A large number of inconsistencies were found, and this from various perspectives:

- Many of ISO 1087-1:2000's requirements concerning definitions are frequently violated by the definitions in the NCIT. From the total of 37,261 classes in the Thesaurus, 33,720 were stipulated to be *primitive* in the Description Logic sense. This means that the majority of these classes are merely *described* rather than *defined*, with the consequence that only a small portion of the NCIT ontology can be used for purposes of automatic classification.
- A total of 16,711 verbal definitions supplied by the NCI itself, together with 5,368 definitions borrowed from elsewhere were found. The numerical mismatch arises in virtue of the fact that some classes in NCIT are assigned more than one verbal definition, whereas at least 55.2% of classes lack a definition.
- Several entries were found where NCIT defines *words*, rather than *concepts*.
- When the NCIT provides several definitions for the same concept, these sometimes contain conflicting information.
- The NCIT stretches the meaning of "synonym" in such a way that the claimed synonymy of numerous terms in the NCIT cannot be accounted for even under ISO's more relaxed definition of "quasi-synonym".
- The most fundamental problem for the NCIT is the unprincipled way in which its class hierarchy is built up. No motivation is given for the specific choice of its 21 top classes, and some of the choices made seem questionable, to say the least.
- An inconsistent use was noticed of the OWL-qualifiers *allValuesFrom* and *someValuesFrom*.

One of the reasons for the identified shortfalls lies in the way the NCIT was constructed: '*by bootstrapping the initialization of NCI Thesaurus from existing terminologies, the project gained the co-operation of diverse stakeholders and avoided pitfalls associated with trying to develop a science based terminology de novo'* [54, p. 36]. By selecting this route, the NCI has taken over some of the characteristic errors of the terminologies from which it draws, and especially some of the characteristic inconsistencies of the UMLS [58, 59].

We are confident that the NCIT is a useful tool for the internal purposes of the NCI, which must be given credit for trying to bridge the clinical and basic biological terminology realms in a single resource, for keeping track of updates, and for trying to harmonise with external ontology modeling practices. The NCI Thesaurus is a never-ending work in progress, the content of which is dictated by the needs of its users and customers. If, however, it wants to establish itself as a useful and trustworthy terminological resource and to play the role of a reference ontology in other contexts, then a considerable effort will have to be made in order to clean up its hierarchies and to correct the definitions and ambiguous terms which they contain.

7.1.1.6.1.3 GALEN

GALEN's research programme into medical terminology began in 1991 [60], while in 1999 OpenGALEN [61] was formed to provide an open source route both for disseminating the results of that programme and as a framework for its future development [62]. Currently available open source resources include an ontology development environment and an open source description logic-based ontology for the medical domain. GALEN is thus far primarily used to improve the usability of old-style coding and classification systems such as the International Classification of Diseases [63], or the French national procedure classification for surgery [64].

Despite the logically principled approach followed by GALEN [65], the spatial information included in this ontology was often found to be ambiguous and the possibilities for implementing consistent automatic reasoning within or across ontologies is thus limited [66]. This was discovered by using a formal theory of parthood and location relations among individuals, called Basic Inclusion Theory (BIT). Since biomedical ontologies are comprised of assertions about classes of individuals (rather than assertions about individuals), parthood and location relations are defined among classes in the extended theory Basic Inclusion Theory for Classes (BIT+Cl). It was then demonstrated that class-level spatial relations with different logical properties are not always explicitly distinguished.

7.1.1.6.2 Realism-based ontologies

These types of ontologies are based on the theory of Basic Formal Ontology (BFO) which gives a formal account of the distinctions between (a) universal and particular, (b) continuant and occurrent, (c) dependent and independent, and (d) formal and material. [67].

7.1.1.6.2.1 The Ontology of Biomedical Reality (OBR).

OBR is an ontology that provides a preliminary classification of organismal continuant entities [68]. Continuant entities are entities which endure self-identically through time while undergoing a variety of different sorts of changes of size, shape,

location, internal structure, and so forth [38]. The OBR classification distinguishes two high-level universals in the realm of organismal continuants: *material anatomical entity* and *material pathological entity*, which are disjoint in the sense that they share no instances in reality. In accordance with the classification schemes presupposed in standard treatises of pathology, OBR conceives the universal *material pathological entity* as comprehending subtypes such as *tumor*, *ulcer*, *portion of pus*, which have no equivalents in normal, healthy organisms. This work provides the foundations for an ontology of pathological continuants that can thus be subdivided into:

- *pathological formations*, for example a carcinoma, a blister, an ulcer, which are newly formed continuant entities evolving in some larger anatomical structure;
- *pathological anatomical structures*, for example a carcinomatous lung, a blistered thumb, an ulcerated colon;
- *portions of pathological body substance*, for example a portion of pus, a portion of amyloid [69].

7.1.1.6.2.2 The Foundational Model of Anatomy

The Foundational Model of Anatomy (FMA) is a structured representation of the anatomy of instances (particulars, individuals), whose constituent nodes are representations of those 'multiply located anatomical entities (i.e., universals) that exist in the instances (particulars) that they subsume'[70, 71]. The universal *anatomical structure* is defined by the FMA as "An anatomical structure is a material physical anatomical entity which has inherent 3D shape and is generated by coordinated expression of the organism's own structural genes". The particular entities which satisfy this definition, and which are thus instances of the corresponding universal, include cells and organs, as well as cardinal body parts such as the head and trunk. For reasons outlined in [71], the FMA is restricted to anatomical entities which are 'typical' in the sense that they can be conceived as belonging to an 'idealized', healthy male or female adult human being. (Such entities are identified in the literature of the FMA also as 'canonical' entities.).

7.1.1.7 Main Ontology-related standardization initiatives

7.1.1.7.1 Open Biomedical Ontologies

Open Biomedical Ontologies is an umbrella web address for well-structured controlled vocabularies for shared use across different biological and medical domains. It includes concept-based ontologies such as the Gene Ontology [72] and MGED [73].

Within the Open Biomedical Ontologies (OBO) framework [74], it has now been agreed upon that contributing ontologies are to be constructed in line with the OBO Relationships Ontology whose foundations are laid down in [75]. This ontology satisfies the principles of good, realist-oriented design.

7.1.1.7.2 caDSR

caDSR [76] is an initiative that uses the ISO/IEC 11179 metadata repository standard [77] in an attempt to standardise the way identical kinds of data are collected across different cancer research studies.

This standard, unfortunately, follows a rather confusing concept-oriented paradigm. The central element in this standard is a **data element**, which is a unit of data for which the definition, identification, representation and permissible values are specified. Such a data element represents, according to the standard, a **data element concept**. As an example, a patient's race is considered to be a data element concept, that can be represented in different ways. If one wants to store data about a patient's race in an EHR, a specific data element must be created for which permissible values are specified (e.g. 'white', 'black', 'Caucasian', …), as well as of what data type these values are. However, neither the data elements, data element concepts, or permissible values are currently defined using sound ontological principles. What is built does indeed conform to the ISO/IEC 11179 specifications, but these specifications alone are not sufficient to create data elements with precise and clear meanings.

7.1.1.8 Ontology outreach, evaluation and certification initiatives

7.1.1.8.1 European Initiatives

7.1.1.8.1.1 The European Q-Rec project

Q-REC is the acronym for "*European Quality Labelling and Certification of Electronic Health Record systems*", a project that will be launched January 2006. Q-REC is a Specific Support Action (SSA) which aims at complementing (bottom-up wise) the existing e-Health ERA Co-ordination Project "Towards the establishment of a European e Health Research Area", which main goal is to coordinate the planning of eHealth R&D and coherent national roadmaps in Europe [78].

The main objective of Q-REC is to create an efficient, credible and sustainable mechanism for the certification of EHR systems in Europe by addressing mainly:

- EHR Systems Quality Labelling and Certification Development, thereby:
    - producing a State of the Art Report on EHR-Certification Schemas as already implemented in at least three European countries;
    - performing a Pan-European Requirements Assay;
    - proposing a Labelling Terminology and Functional Profiles for EHRs to be certified;
    - comparing and harmonising the EHR-Certification Procedures at a European level;
    - drafting Model Certification Guidelines and Procedures;
    - planning the Validation of the Guidelines.
- Resources for EHR Interoperability, including:
    - the register of Conformance Criteria and Guidance Documents for obtaining EHR Certification;

- an inventory and guidelines for EHR Archetypes;
- the registration of Coding Schemes in Europe (as mandated by CEN/TC 251);
- an inventory of relevant EHR related standards;
- a register of XML Schemas and Open Source components for EHRs.

- Benchmarking Services:
  - Benchmarking Services Manual for Quality Labelling and Certification;
  - preparing the Business Plan for new EHR-Certification related Services.

The project is led by the Prorec/Eurorec Network [79]. The EUROREC Institute is a non-profit, independent organisation, promoting the use of high-quality Electronic Health Records in Europe. EUROREC's mission is to promote the adoption and extended use of Standardised Electronic Health Care records. The European Institute for Health Records represents a permanent network of National Prorec Centres and provides quality and affordable added value services to the European market, Healthcare providers, Governments and Patients.

7.1.1.8.1.2 The European RIDE-project

Also to be started in January 2006 is RIDE, a roadmap project for interoperability of eHealth systems leading to recommendations for actions and to preparatory actions at the European level [80]. This roadmap will prepare the ground for future actions as envisioned in the action plan of the eHealth Communication COM 356 by coordinating various efforts on eHealth interoperability in Member States and the associated states. Since it is not realistic to expect to have a single universally accepted clinical data model that will be adhered to all over the Europe and that the clinical practice, terminology systems and EHR systems are all a long way from such a complete harmonisation; the RIDE project will address the interoperability of eHealth systems with special emphasis on semantic interoperability. In order to create RIDE Roadmap, first the European best practices in providing semantic interoperability for eHealth domain will be assessed and the quantified requirements to create a valid roadmap will be identified. Based on these requirements, the goals and the economical, legal, financial and technological challenges of the industry for the 21st century for achieving interoperability in eHealth solutions will be elaborated. RIDE will also focus on the limitations of the policies and strategies currently used in deploying interoperable eHealth solutions. A research portal for sharing resources addressing semantic interoperability in the eHealth domain will be created and maintained; the key actors and stakeholders will be coordinated around RIDE special interest groups to create a wide consensus at the European level. Through eight RIDE workshops a shared vision for building a Europe-wide semantically interoperable eHealth infrastructure will be created. After assessing the gaps, the emerging trends and opportunities to achieve the vision statement, the required advances in the state of the art research, technology and standards will be identified.

7.1.1.8.2 Non-European Initiatives

7.1.1.8.2.1 The US National Center for Biomedical Ontology

The National Center for Biomedical Ontology, created in October 2005 following a $18.8 million grant from the NIH [81], is a consortium of leading biologists, clinicians, informaticians, and ontologists who develop innovative technology and methods that allow scientists to create, disseminate, and manage biomedical information and knowledge in machine-processable form. The Center's resources include the Open Biomedical Ontologies (OBO) library, the Open Biomedical Data (OBD) repositories, and tools for accessing and using this biomedical information in research.

An important part of the center's mission is ontology outreach, which consists of [82]:

- offering training workshops during which experts in ontology development from the Center work directly with biomedical scientists in using the Center's methods and technologies, allowing them to gain hands-on experience in developing specific biomedical ontologies and in evaluating the results of their work.
- maintaining active relationships with professional societies, government laboratories, and academic groups who are developing biomedical ontologies.
- providing Internet resources for discussion, critique, and improvement of existing biological ontologies, ontology tools, and access to the Center's methodology for ontology evaluation.

7.1.1.8.2.2 FQHC Electronic Health Records certification

The Health Resources and Services Administration, Bureau of Primary Health Care (BPHC) [83] has developed a specification that describes the functional requirements for EHRS in Federally Qualified Health Centers (FQHC) [84] and invited providers to self-certify [85]. Unfortunately, the specifications do not take into account that many EHRSs need to be configured within a specific institute, and that it depends on this configuration whether or not the quality criteria are satisfied. Nevertheless, it is a very good resource to assess what functionalities have been considered appropriate by a group of relevant experts.

7.1.1.8.2.3 NCOR's Committee on Ontology for Health Informatics

NCOR (see section 4.3 p.17) has created a *Committee on Ontology for Health Informatics* which has the task of advancing ways in which scientific methods in ontology can bring benefits to healthcare. They will work with public and private sector institutions to help advance high-quality work on the ontologies used in conjunction with computer-based applications in areas such as the electronic health record, hospital management, and public health.

One purpose of these ontologies is to support the integration of data from different sources, and they believe that such integration will become of increasing importance as ever larger amounts of data become ever more critical to the care of individual patients. The application of ontologies in healthcare can bring practical advantages also in controlling healthcare costs, for example by reducing costs of record keeping

by allowing the same data to be used simultaneously for a variety of purposes, ranging from billing and cost-control to clinical trials and diagnostic decision support.

NCOR's Committee on Ontology for Health Informatics will examine how ontology can contribute to the accomplishment of these goals by contributing to the improvement of existing resources in ways designed to advance their harmonization and interoperability.

The Committee has a collaborative relationship with the National Center for Biomedical Ontology (7.1.1.8.2.1).

Convenors are Rex Brooks (OASIS) and Barry Smith.

Members are ,Thomas Beale, Werner Ceusters, Christopher Chute, Mark Musen, and Bob Smith.

7.1.2 Activities of the OOA-Healthcare and Life Sciences

7.1.2.1 Recruitment

A first list of potential interested parties has been compiled on the basis of the Innomed Platform and the network around IFOMIS:

> *InnoMed* is being led by the European Federation of Pharmaceutical Industry and Associations (EFPIA) which guarantees a commitment from all the stakeholders needed to change the process of drug development in Europe. The course for addressing the necessary changes is to first develop a Strategic Research Agenda (SRA) that will encompass the whole path from discovery of a new drug target to the validation and approval stages of a new drug compound. This will be agreed by all the relevant stakeholders via meetings and workshops. Four key bottlenecks in the drug development process will be addressed: Safety, Efficacy, Knowledge Management, and Training and Education [86].
>
> The *Institute for Formal Ontology and Medical Information Science* was founded in April 2002 utilizing a grant of the Alexander von Humboldt Foundation. IFOMIS comprehends an interdisciplinary research group, with members from Philosophy, Computer and Information Science, Logic, Medicine, and Medical Informatics, focusing on theoretically grounded research in both formal and applied ontology. Its goal is to develop a formal ontology that will be applied and tested in the domain of medical and biomedical information science [87].

These two initiatives are currently the most prominent in the Life Sciences and Health sector with respect to ontology matters, what has been our prime motivation for selecting contacts amongst their midst.

Invitations were sent to the following addresses (see the invitation letter in the appendix):

Name	Email	Country
Werner Ceusters	werner.ceusters@ecor.uni-saarland.de	EU
Andreas Persidis	biovista@ath.forthnet.gr	Greece
Andre Michel	andre.michel@aureus-pharma.com	France
Bernard Vatant	bernard.vatant@mondeca.com	France
	dfrench@consider.biz	Canada
Frederique Lisacek	frederique.lisacek@genebio.com	Switzerland
Ferran Sanz Carreras	fsanz@imim.es	Spain
Geoffrey Pegler	geoffrey.pegler@insightful.com	
Georges demoor	georges.demoor@ugent.be	Belgium
Jean-Christophe Mestres	jc.mestres@fr.ibm.com	France
Julie Barnes	Julie.Barnes@biowisdom.com	UK
Karen Strandgaard	karenstrandgaard@efpia.org	Switzerland
Karima Boubekeur	karima.boubekeur@roche.com	Switzerland
Nicolas Grandjean	Nicolas.grandjean@novartis.com	Switzerland
Pascal Coupet	pascal.coupet@temis-group.com	France
Scott Boyer	Scott.Boyer@astrazeneca.com	Sweden
	srooney@consider.biz	Canada
Steve Gardner	Steve.Gardner@biowisdom.com	UK
Thierry Lave	thierry.lave@roche.com	Switzerland
Patrick Mallea	mallea.p@chu-nice.fr	France
Carole Moquin-Pattey	cmoquin-pattey@esf.org	
Yannick Legré	legre@clermont.in2p3.fr	France
Vincent Breton	Vincent.Breton@clermont.in2p3.fr	France
Robin Smith	robin.smith@synthematix.com	USA
Alex L. Bangs	bangs@entelos.com	USA
Karl Stroetmann	karl.stroetmann@empirica.com	Germany
Dirk Colaert	dirk.colaert@agfa.com	Belgium

Responses thus far:

- *Positive*: ECOR: Prof. Werner Ceusters, BioVista: Dr. Andreas Persidis, AGFA Healthcare: Dr. Dirk Colaert, BioWisdom: Dr. Steve Gardner, HealthGrid: Dr. Yannick Legré, Empirica: Dr. Karl Stroetmann.
- *Checking right person involvement*: IBM Research: Amnon Shabo
- *Negative*: Entelos: Alex Bangs (Computer-based disease simulation systems)

The key idea of the recruitment policy is that the composition should be balanced and should reflect market segments such as Biomedical and bio-informatics, Pharmaceutical (therapeutics, diagnostics), Biomedical equipment, and Healthcare services (therapeutic, diagnostic, health economy, insurance). It should also reflect horizontal forces within the domain such as standardisation bodies (CEN/TC251, ISO/TC215), active communities (OpenEHR, OpenClinical, …), Networks of Excellence (Semantic Mining), and several 'ontology' builders and distributors.

7.1.2.2 Participation in ongoing activities

Werner Ceusters, Chair of the KnowledgeWeb OOA - Chapter for Life Sciences and Health, is involved already in all major initiatives that have been listed above: the

European RIDE and Q-REC projects, and the US-based National Center for Ontological Research.

Relevant outreach activities conducted include:

- *Overview of European Developments in Ontology*. UB Buffalo, North Campus, Park 280, NY, September 18, 2004.
- *Using realist ontology to link patient records with terminologies*. EUROREC-2004 satellite workshop on Ontology and semantic interoperability of systems sharing bio-medical information, Brussels, Belgium, November 25, 2004. [88]
- *Realist Ontology for Electronic Health Records*. Eurorec-2004, Brussels, Belgium, November 26, 2004. [89]
- *Realist Ontology for the Semantic Web: Applications in Biomedical Informatics*. Guest lecture at the University of Diepenbeek, Belgium, December 2, 2004 [90].
- *Terminology and Ontology in Semantic Interoperability of Electronic Health Records*. Joint EU-WHO workshop on Semantic Interoperability Prerequisites for Efficient eHealth Systems. Brussels, Belgium, February 14-15, 2005 [91, 92].
- *Ontology for Emergency Medicine*. DICOEMS Workshop, Medetel, Luxemburg, April 8, 2005 [93].
- *Introduction to Ontology. Tutorial for Semantic Mining in Biomedicine 2005* (SMBM2005), EBI, Hinxton, UK. April 10, 2005 [94].
- *Ontology Outreach Advisory: Healthcare & Life Sciences Chapter*. KnowledgeWeb Project Conference, Knossos, Crete, June 3, 2005 [95].
- *Practical implementations of realism-based ontologies*. MIE 2005 Tutorial #35 on Ontology Design, Geneva, Switzerland, August 28, 2005 [96].
- *Ontology: the need for international coordination*. Inaugural meeting of the National Center for Ontological Research (NCOR), Buffalo, NY, October 27, 2005 [97].
- *Ontology for indexing electronic healthcare records*. GdR Stic-Santé, Paris, France, December 8, 2005 [98].

7.1.3 Planned activities for the OOA-HC

The concrete action plan of the OOA-HC covers short and mid-term objectives. Short term objectives include making an inventory of biomedical and bioinformatics ontologies, selecting evaluators, defining evaluation fee principles (in & out), inviting developers for submission, and drafting a targeted dissemination and promotion plan. Longer term objectives aim to move towards quality improvement of ontologies and to agree on terminology for types of ontology.

Crucial for its success is to elaborate a strategy for dealing with the many ongoing initiatives as discussed before. Already at this stage, a connection between the OOA HC and many of the other initiatives is established by means of a number of key

people that are involved in several of these initiatives. Given the huge challenge that each of these initiatives is facing, it might be appropriate to come to a close form of collaboration. This might be initiated by organising a joint workshop.

7.2 Establishing the OOA-HR (Human Resources)

7.2.1 Sector analysis

7.2.1.1 Recruitment techniques and procedures

There is an increasing body of evidence to demonstrate the importance of investment in human capital for economic development. Although most businesses rely on recruiting channels such as newspaper advertisements, online job exchange services, trade fairs, co-worker recommendations and human resources advisers, online personnel marketing is increasingly used with cost cutting results and efficacy. This means that Human Resources Management (HRM) has also discovered the Internet as an effective communication medium, which through intelligent IT solutions expedites the recruitment procedure. Reliable information regarding the availability of employment opportunities, qualifications and individual procurement capability is crucial for effective employment procurement. Due to the employment and consultation market's vast array of suppliers, services and channels on the Internet, visibility is rather limited. Instead of simplifying and facilitating the job search, the Internet has made navigation even more complex and difficult. The information flow in the online labour market is therefore still far from optimal.

From an organization's viewpoint, a typical recruitment process can be divided into three main phases, in which relevant problems and "opportunities for reform" can be discovered.

1. **describing the requirements of the job position**: Nowadays the job postings are written in the form of free text using uncontrolled vocabularies, which makes the search process harder and limits the machine processability of postings in the later phases of the recruitment process.
2. **publishing the job posting**: a large number of online job portals have sprung up, dividing the online labour market into information islands. There are websites and online portals financed by publishing fees (Jobpilot [99]), various business websites and portals set up by state job centres (German Federal Employment Office [100] or the Swedish National Labour Market Administration [101]). Publishing postings on a corporate website is cheap but reaches only a very limited audience, because the indexing capabilities of current search engines like Google are too imprecise to support directed searches for open positions. Another problem is that job portals differ substantially according to their positioning regarding geographical areas, specific industries or occupation groups. Further differentiation criteria include the range of online functionalities and the scale of integration of jobs. This is why it is close to impossible for a job seeker to get an overview of all relevant open positions, since visiting every single job exchange site would be next to impossible. Furthermore, the meta-search engines available conduct searches on a full text basis and accordingly they are limited in their ability to provide offers that match the precise needs of their clients. Because of this, defects in

quality of data and query results have been observed. To cover as wide a part of the labour market as possible and to reach more potential applicants, companies would have to publish their offers on multiple sites and portals. However, the strong market position of the job portal as the prime starting point for job seekers allows them to charge employers high fees for publishing open positions. Due to these costs, employers publish their job postings only on a small number of portals, which prevents the offers from reaching all qualified applicants. For this reason, employers are very interested in solutions which help to decrease transaction costs for publishing job postings and at the same time do not reduce the number of applicants that can be reached.

3. **receiving applications and decision making**: Employers often receive a large number of applications for an open position, due to the strained situation of the labour market. The costs of manually pre-selecting potential candidates have risen and employers are searching for means to automate the pre-selection of candidates.

7.2.1.2 Knowledge-based approaches

Generally, the field of Human Resources (HR) is a generic domain into which a great deal of effort in terms of knowledge management tends to be placed, because every company, organisation and business unit must encounter it. HR departments often have an eye open for knowledge management in order to monitor their environment in the best way, and many recruitment consultancy companies have watchdogs to monitor and alert them to changes. Among the multiplicity of online portals there exist a variety of job search engines (portals) which already use knowledge management extensively to link employees and employers, e.g. JobSearch [104] and Job Portals [102]. The growing pervasiveness of Knowledge Management (KM) in industry marks an important new watershed. KM has become embedded in the strategy, policy and implementation processes of institutions and organisations worldwide. The global KM market has doubled in size since 1991 and is projected to exceed US$8.8 billion in 2005. KM applications are expected to save Fortune 500 companies around $31 billion, and the broader application cost has similar projected forecasts.

The HR domain has many facets. But one particular ‘view’ on HR is fast becoming the motor for serious change. Competency-centric HR is not only reaching a great number of traditional HR processes, it is also the start for a new wave of change in the labour market as a whole. Programs such as ‘Matching on Competencies’ (MoC, replacing the traditional Job-CV matching), Mobility@Work and competency driven qualifications form the basis for competencies to readily become the currency of the European labour market.

Clearly, therefore, the combination of KM and Human Resources has enormous implications for the growth and dispersion of such new technologies to industry as a whole. Tools and resources such as next generation Knowledge Management platforms pave the way for such developments, by leading to interesting and useful acquisitions of knowledge that save time and money and benefit real users in industry. Examples of such systems are the h-TechSight Knowledge Management Platform [103] and Ontotext’s JOCI [104] (Jobs and Contacts Intelligence – Recruitment Intelligence through Semantic Web Technologies). Companies such as Innovantage

[105] have recently been established which make use of such technology to provide information such as vacancies, contact information and biographies harvested direct from company, academic and government websites via tangible business intelligence tools for the recruitment market.

A "conditio sine qua non" for such projects, tools and systems to make a real impact, however, is a "meaningful" way to exchange competency data between industry, education and public & private employment services. Europe therefore needs a 'semantic' standard topping the existing internationally accepted HR and Learning syntactic standards. HR is far from alone in this respect.

7.2.1.3 Lifelong learning and Competencies, in need of a semantic boost?

With the increased focus on lifelong learning and the development and assessment of competencies in the labour market, the learning & knowledge domain is slowly beginning to provide functional bridges between education, industry and public employment. This societal realignment of interoperability requirements poses new challenges for the design and implementation of technology standards relating the learning and human resources domains. As computing environments evolve from self-enclosed, proprietary, monolithic systems toward a service-oriented architecture (SOA), the challenge involves developing XML standards to support these functional bridges.

One of the most prominent areas in need of standardization for such cross-domain communication and functional synergy is probably 'competencies'. There are currently several standard specifications which each originated in their own community of standards:

The IMS (www.imsglobal.org) Reusable Definition of Competency or Educational Objective (RDCEO) specification was based on a draft from the IEEE Learning Technology Standards Committee (LTSC). Released in 2001, this specification was aimed primarily at the learning management domain. In a completely different arena, the HR world, the HR-XML consortium (www.hr-xml.org) developed a specification for competency records, aimed primarily at the recruiting and employee selection domain.

Fast forward to 2005. The RDCEO specification is back in the accredited IEEE standardization process, where the LTSC is using it as the basis for the Reusable Competency Definitions (RCD) standard draft P1484.20.1 [106]. The HR-XML consortium has established a liaison with the eLearning community and is developing new specifications for XML encoding of competency information that incorporate the Reusable Competency Definitions concept and that can be extended to the areas of assessment and learning management. Various national and European initiatives have worked on profiles for the RDCEO model, or built competency inventories that are compatible with this model.

There are also many 'out-of-band' competency approaches implemented by the European Public Employment services, which currently are slowly moving towards HR-XML compliance and pressuring the HR-XML group in addressing their requirements.

Being a semantic rich environment, the RCD and related standards have been hampered by the lack of a semantic underpinning. When it is considered in a

perspective of lifelong learning, and the attendant requirements for a lifelong competency framework, this situation is likely to get worse.

Take for instance the concept of a Curriculum Vitae (CV). In Learning technology, an embodiment of this concept can be found in the IMS ePortfolio specification. In the HR world, HR-XML has two different specs that use CV-like concept (or "resume"), but even within that consortium those specifications are not aligned properly.

What semantic technology could offer is a conceptual layer that bridges these related standards specs, and to which they could all ontologically 'commit'. As such the semantic specification model would add meaningful interoperability to the other specifications that are based on simpler data models and instances that focus on the syntax of the data for exchange.

7.2.1.4 Main ontologies in the HR sector

- COKE, a three-level ontology containing a top-level Human Resources ontology, a middle-level Business Process ontology and a lower-level Knowledge Objects ontology [107];
- TOVE (Toronto Virtual Enterprise Ontologies), a set of integrated ontologies for the modelling of commercial and public enterprises [108];
- PROTON (PROTO-Ontology), a 4-level ontology which specializes in coverage of concrete and/or named entities (i.e. people, organizations, numbers) and is used for HR applications [109].

7.2.1.5 Main standardization initiatives in the HR sector

- The HR-XML consortium has built up a library of more than 75 interdependent XML schemas which define the data elements for particular HR transactions, as well as options and constraints governing the use of those elements [110]:
  - HR-BA-XML was developed by the German Federal Employment Office and is a German extension of the international HR-XML standard. The categories defined in HR-XML were supplemented on the basis of German employer requirements.
  - HR-XML-SE is a Swedish standard which consists of the original HR-XML parts (transformed from DTD's into schemas), to which schemas with Swedish extensions are added.

7.2.2 Activities of the OOA- Human Resources

7.2.2.1 Activities carried out

The following companies have been invited thus far (see the invitation letter in appendix). These companies have been selected on the basis of their overall interest in the use of semantics in HR and e-learning standards, especially with the emergance of

cross-domain standards as needed in anything "lifelong": lifelong learning, lifelong competency framework, CVs (resume, etc.).

Synergetics: Luk Vervenne,
IEEE- LTSC: Claude Ostyn
HR-XML Europe
SAP - R&D: Torsten Leidig
EDS: Frank Huynen, Client Industry Executive, Government Sector - Belgium
Nederlands Bakkerij Centrum: Hans Koster director
Kenteq.nl: Dries Pruis
Ordina.nl: Theo Hensen
British Telcom: Joep Dejong
City of Antwerp: Marc van Praet
Klett: Jurgen Schmitt
IMC Volker Zimmermann
Le Forem: .Jeanclaude Xheneumont
IDS-Scheer: Julia Wagner
Catalysoft: Adil Hameed

Most of the partners reconfirmed their interest in the idea of semantic underpinning of xml standards, but also stressed the need of concrete examples and tests.

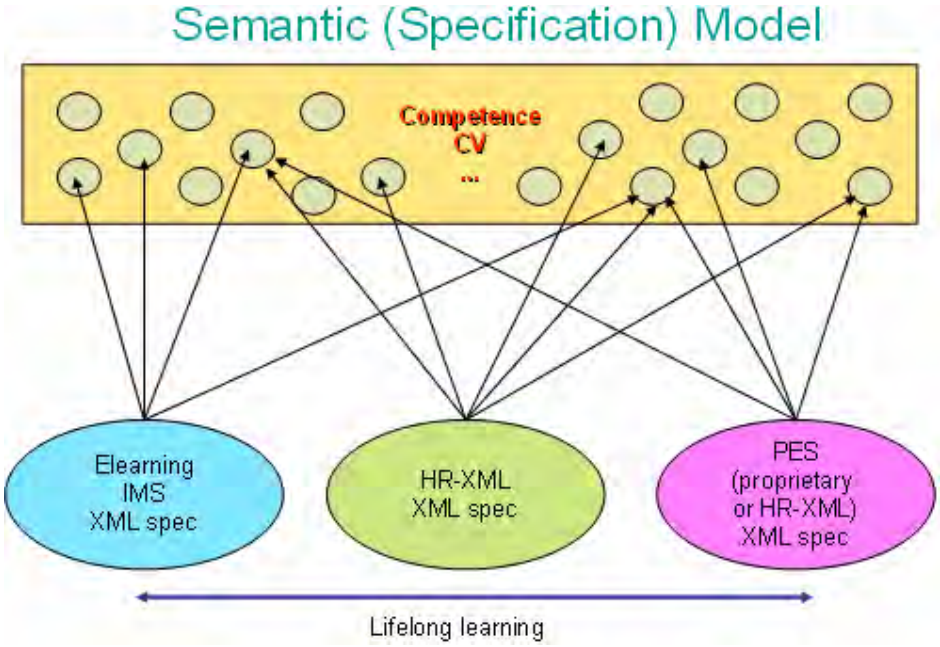
OOA is in a partnership with Synergetics who is chairing the competency workgroup of HR-XML Europe, and is board member of www.eife-l.org and www.elig.org. HR-XML is setting up an CEN/ISSS workshop on a competency framework fir which the OOA will be invited as the "semantics authority".

Furthermore, within the standards arena, some OOA HR-chapter partners have been working internationally towards a model for a symbiosis between the pragmatic data model specifications, usually expressible in XML, and the ontology infrastructures that can bring meaning and context to the application of those specifications.

7.2.2.2 Planned activities

The IEEE RCD draft standard covers the representation and exchange of the portion of competency data that can be reused in different contexts or for different people, maybe with different metrics. It assumes that the context is defined somewhere else. To be operational, a competency definition requires some context, of course, and so the definition of competency context seems a particular good place to start such a semantic exercise. Some promising research has already been done in this area. The

insights from such early research were implemented in the EU Codrive project [111]. This project demonstrated the benefits of adding context as a semantic model, which could then be shared by various learning en HR applications in order to contextualize, competencies, index learning objects, and so on.



Research is expanding in this area. The semantic specification method will be discussed and used in a (two-year) CEN/ISSS workshop which will start Q1 2006 on a global lifelong competency framework. In fact all competency related standards bodies (IMS, IEEE, HR_XML) have agreed to join the HR-XML initiated workshop to set up an aligning guidance framework. The goal of this framework is to provide guidance in aligning the existing and emerging competency specifications and standards within a broad societal perspective to support lifelong learning and competency management. The OOA HR-Chapter will act as the reference contact point for all things semantic within this exercise.

If this succeeds, this may be the start of a next era in standardization where interoperability goes beyond mere data management and reaches into meaningful dimensions.

8 Future steps: Analyses of other domain Sectors

In this section, we provide a brief analysis of the other promising sectors that we aim to target in the future: legal, financial, e-learning and e-government.

8.1 The legal sector

In the last few years, Legal Informatics (the study of methods for automating the treatment of legal information) has been significantly influenced by Artificial Intelligence (AI) approaches. For instance, Machine Learning techniques have been successfully applied to problems of legal document classification, legal information retrieval, legal knowledge discovery and extraction. As the use of these techniques becomes more widespread, it also becomes clearer how to enhance their performance. One way of doing this is to employ structured (domain) knowledge in order to reduce complexity and support correct reasoning.

In 1997, the First International Workshop on Legal Ontologies was held in conjunction with the Sixth International Conference on Artificial Intelligence and Law at the University of Melbourne, Australia. It was a successful workshop in which 8 papers were presented on issues ranging from proposals for legal (core) ontologies, through the comparison of different ontologies, to means of building them automatically from legal sources [112]. Since then, much research has been done, especially in the broader fields of ontological engineering and knowledge management: systems using legal ontologies have been built, and both practical and theoretical problems and opportunities have been encountered. Such legal ontologies play a crucial role in providing such knowledge at various levels of specificity and formality [113].

The potential usefulness of ontologies in the legal world has been demonstrated on various occasions. For example, LOIS is a new international research project on multilingual information retrieval from European legal databases [114]. This aims to enable citizens and professional users to search a multilingual legal database for European legislation and other legal documents (such as court cases) across six European languages (Italian, English, German, Czech, Portuguese and Dutch), by means of a large legal ontology, LOIS-WN (see below).

Other examples where legal ontologies are used include:

- complaint regulation [115];
- the problem of normativity [116];
- representation of legal contracts [117];
- Digital Rights Management [118, 119].

One of the main ontologies in this sector is LOIS-WN, an electronic legal **thesaurus** with multilingual capabilities, based on the semantic structure of WordNet and EuroWordNet.

The main standardization initiatives in this sector are:

- **LeXML**: co-ordination and a workforce for the development of XML standardized structures, vocabularies and data exchange tools (lexml.nl, lexml.de, lexml.at, lexml.se).

- **LISA**: open non-commercial network to enhance legally founded trust in the use of information standards. In order to achieve this, LISA works with legal issues related to the introduction and use of IT in society. http://www.lisan.org
- **MetaLex**: open XML standard for the markup of legal sources [120].
- **OASIS:** Organization for the Advancement of Structured Information Standards

8.2 The Financial sector

In the world of finance, knowledge is pervasively present. The artefacts of banking and investment such as banking accounts and financial instruments contain a wealth of knowledge both in their structure and content; furthermore the policies that govern their use are rich in knowledge. Financial service providers need to manage various risks, anti-money laundering and regulatory compliance. The adoption of Semantic Web Technology and ontologies is motivated by, amongst other things, the following:

- Achieving enterprise data: integration and interoperability across multiple databases and information sources
- Effective integration of data and capabilities in mergers and acquisitions
- Operational Integrity: policy-driven information and application access
- Proof and Security: e.g. fraud prevention
- Customer Intimacy: knowledge-driven call centre support
- Smarter Content Management of external and internal knowledge resources
- Knowledge Asset Management: knowledge classification and knowledge networks
- Regulatory Compliance: governance and accountability.

The main ontologies in this sector are:

- SUMO Financial Ontology, a domain ontology that forms part of SUMO [121]

http://ontology.teknowledge.com/

- DIP Financial Ontology [122].

The main standardization initiatives in this sector are:

- **fpML.org**, XML standards for financial products.
- **OASIS**

8.3 The e-Learning sector

E-learning environments are claimed to be a significant means to improve communication and understanding between users in the education area such as teachers, students, information systems users and more generally for training. Many users have already described their learning materials, respecting more or less the standards and norms, using Web semantic languages such as XML, RDF, OWL.

These users had to develop ontologies in order to index and annotate their materials. Ontologies are a pivot for the indexing, annotation and research of learning materials. The use of these materials, the integration of new materials and the annotation process performed by teachers or students all need some specific approaches. A lot of Web-based courses and other educational applications have been made available on the Web. Nevertheless, most of them are not well integrated in the personal information system of the general public. Many methodological, technical and social problems need to be solved in order to give means for the life long learning. Ontologies for e-learning applications may contribute in several ways [123]:

- by providing semantic interoperability of educational contents on the web;
- by assisting in the development of personalised and adaptive learning environments;
- by helping teachers to produce web materials with web resources;
- by making maintenance of web-based learning environments much simpler.

8.4 The e-Government sector

Essentially, every Public Administration makes use of knowledge in order to increase productivity of its activities [124] and therefore efficient, scalable and flexible knowledge management systems are required to support these activities [125]. Government administration tends to involve information that is highly distributed and heterogeneous, and therefore requires information management applications that allow data to be streamed seamlessly across different environments. In particular, the e-government sector poses additional demands on such applications in view of the fact that the data is constantly changing on a large scale, for example when new laws or statutes are introduced this can have huge impacts on the nature of the information required and produced.

In 2001, a report was published by the OECD (Organisation for Economic Co-operation and Development) stating that all OECD countries[3] regard new information and communication technologies (ICTs) as a powerful tool and are making significant efforts to bring their administrations and their citizens on-line [126]. While many believe ICTs have great potential for engaging citizens in policy-making, they are still currently seen as an addition to rather than as a replacement for traditional tools. Most OECD governments are working to bridge the digital divide, and recognise the need to ensure that all citizens, whether on-line or not, continue to enjoy equal rights of participation in the public sphere. All OECD countries recognise the need to develop tools and to improve their capacity for evaluation. However, at the time of the report, no OECD had conducted a systematic evaluation of government performance in providing information, conducting consultation and engaging citizens in policy-making.

The main standardization initiatives in this sector are the UK e-Government Interoperability standards and specifications, and **OASIS.**

[3] The OECD encompasses 30 countries including most major EU contries, the US, Japan, Korea and Australia, amongst others.

9 Appendices

9.1 The OOA-HC Invitation Letter

Invitation to join
The Ontology Outreach Advisory
Healthcare and Life Sciences Chapter

Dear Colleague,

We hereby invite you to join us as member of the **"European Ontology Outreach Advisory Board"** (OOA) and to play an active role in the human resources chapter of this advisory board.

The OOA board is an initiative of KnowledgeWeb, the world's largest research network on Semantic Web Technology in general and on ontology's in particular, funded by the European Commission. KnowledgeWeb is also the first research network with a specific focus on bringing its research results to industry.

The main instrument for achieving this "outreach to industry" is the "Ontology Outreach Advisory" (OOA), an industry-sector-led community board that will help promote greater awareness and faster take-up of ontology's and at the same time act as an "*recommending body*" of industry domain specific ontologies.

Semantic Interoperability

Now that XML-based standards are abundant, the world is starting to realize that syntax alone is not enough. Fortunately the EC through its 5th and 6th Framework Programme has directly and indirectly spent tens of millions of euros in semantic research funding. This leaves us with the task of scaling up the transfer of technology and research results to industrial and governmental best practice.

The **Healthcare and Life Sciences domain** is one of these areas where the need for and benefits of semantic technology start to be well understood. Semantic web technologies have to potential to shortening the drug discovery and development cycle by linking genotype to phenotype data and providing accurate reasoning algorithms for it. Medical mistakes may be prevented by a better semantic representation of clinical data in electronic patient records and the decision support facilities that may come true by linking patient data to ontologies and epidemiologic databases. Researchers and scientists will have better and more efficient access to literature databases using advanced text mining tools. These are not just far fetched claims made by developers or early adopters of Semantic Web technologies, but true beliefs that are reflected already in political decisions as witnessed by the Action Plan for the European e-Health area. The OOA HC board intends to investigate & promote the use of ontology's in the HC and LSc domain.

The OOA will be organized into domain specific sector boards. The first two will cover (1) Human resources & Employment and (2) Healthcare & Life Sciences. Additional strategic sectors being planned are: Regulatory Compliance, Telecom, Education, etc… Each of these domains is the focus of currently ongoing and representative ontology and Semantic Web research and application efforts. **More details about the goals and the organization of the OOA are provided in the attached presentation material.**

For this reason, KnowledgeWeb is looking for thought leaders and innovators active in these sectors both within industry and government, who are willing to assist their sector in making the next step towards "meaningful computing", by signing on to the European Ontology Outreach Advisory.

"What is in it for me?"

Being on the forefront of semantic computing, you will have first-hand knowledge, deep insight and high level academic contacts to lead your industry sector or governmental organization into the age of "meaningful computing".
The **Ontology Outreach Advisory** board will be a non-profit member organization that consists of industry, government, and research leaders and innovators in their respective fields.

What is needed?

At this stage, we need representatives from companies and organizations active in these sectors who are ready to show interest to the goals and activities of the

OOA. Little or no technical knowledge is required. Instead we ask to bring a thorough understanding of your sector and its interoperability needs (different languages, meaning, and processes). We also appreciate the willingness to promote and disseminate the potential (awareness, services, tools) made available through the OOA.

As a member (no membership costs involved) and after being trained, you will be expected to help promote and introduce the use of semantics and ontology's for use in your professional environment.

How can I join?

Just sign this document and return it to the address below. Once accepted you will get an invitation to attend our next meeting (2 June, Crete), receive the OOA email newsletter and you will be contacted to help start the "meaningful computing" era in your company, sector or region.

Kind regards,

Werner Ceusters

Chair OOA HC Chapter

European Centre for Ontological Research
Universität des Saarlandes
Postfach 151150
D-66041 Saarbrücken
Germany

www.ecor.uni-saarland.de
werner.ceusters@ecor.uni-saarland.de

direct: +32 475 486 587

The Ontology Outreach Advisory
Healthcare and Life Sciences Chapter
Letter of interest

From :

Company name:

Company Representative Name:

Email:

Tel: Fax:

To: **Professor Dr. Robert Meersman**, OOA Chair, VUB/Starlab
OOA Board
VUB - STAR Lab
Pleinlaan 2 - Gebouw G - 10
1050 Brussels
Tel: +32 (0)2 629 12.37

Dear Professor Meersman

I hereby declare that (company name) _____ is interested in joining the "Ontology Outreach Advisory" initiative as member.

_____ (company name) accepts to help promote the use of semantics and ontology's in the HC and LSc domain within my industry sector or governmental organization. The OOA activities will be attended by:

_____ (company representative name).

Name:

Date:

Signature:

9.2 The OOA-HR Invitation Letter

Invitation to join
The Ontology Outreach Advisory
Human Resources Chapter

Dear Colleague,

We hereby invite you to join us as member of the **"European Ontology Outreach Advisory Board"** (OOA) and to play an active role in the human resources chapter of this advisory board.

The OOA board is an initiative of KnowledgeWeb, the world's largest research network on Semantic Web Technology in general and on ontology's in particular, funded by the European Commission. KnowledgeWeb is also the first research network with a specific focus on bringing its research results to industry.

The main instrument for achieving this "outreach to industry" is the "Ontology Outreach Advisory" (OOA), an industry-sector-led community board that will help promote greater awareness and faster take-up of ontology's and at the same time act as an "*recommending body*" of industry domain specific ontologies.

Syntax and semantics

Now that XML-based standards are abundant, the world is starting to realize that syntax alone is not enough. Fortunately the EC through its 5th and 6th Framework Programme has directly and indirectly spent tens of millions of euros in semantic research funding. This leaves us with the task of scaling up the transfer of technology and research results to industrial and governmental best practice.

The HR domain consists of many processes in which – particular in Europe – competencies are play an increasingly important, if not pivotal role. Programs such as 'Matching on Competencies' (MoC, replacing the traditional Job-CV matching), Mobility@Work, competency-driven qualifications and even competency-based exit strategies, are laying the foundation for competencies to readily become **the common currency' of the European labor market**. However, a 'conditio sine qua non' for this to happen is a "meaningful" way to exchange competency data between industry, education and public/private employment services. The OOA HR board intends to investigate & promote the use of ontology's in the HR domain.

The OOA will be organized into domain specific sector boards. The first two will cover (1) Human resources & Employment and (2) Healthcare & Life Sciences. Additional strategic sectors being planned are: Regulatory Compliance, Telecom, Education, etc… Each of these domains is the focus of currently ongoing and representative ontology and Semantic Web research and application efforts. **More details about the goals and the organization of the OOA are provided in the attached presentation material.**

For this reason, KnowledgeWeb is looking for thought leaders and innovators active in these sectors both within industry and government, who are willing to assist their sector in making the next step towards "meaningful computing", by signing on to the European Ontology Outreach Advisory.

"What is in it for me?"

Being on the forefront of semantic computing, you will have first-hand knowledge, deep insight and high level academic contacts to lead your industry sector or governmental organization into the age of "meaningful computing".

The **Ontology Outreach Advisory** board will be a non-profit member organization that consists of industry, government, and research leaders and innovators in their respective fields.

What is needed?

At this stage, we need representatives from companies and organizations active in these sectors who are ready to show interest to the goals and activities of the OOA. Little or no technical knowledge is required. Instead we ask to bring a thorough understanding of your sector and its interoperability needs (different languages, meaning, and processes). We also appreciate the willingness to

promote and disseminate the potential (awareness, services, tools) made available through the OOA.

As a member (no membership costs involved) and after being trained, you will be expected to help promote and introduce the use of semantics and ontology's for use in your professional environment.

How can I join?

Just sign this document and return it to the address below. Once accepted you will get an invitation to attend our next meeting (2 June, Crete), receive the OOA email newsletter and you will be contacted to help start the "meaningful computing" era in your company, sector or region.

Kind regards,

Luk Vervenne

Chair OOA HR Chapter

VUB/Starlab
Pleinlaan 2
1050 Brussels
luk.vervenne@vub.ac.be
http://starlab.vub.ac.be
direct: +32 476530021

The Ontology Outreach Advisory
Human Resources Chapter
Letter of interest

From :

Company name:

Company Representative Name:

Email:

Tel: Fax:

To: **Professor Dr. Robert Meersman**, OOA Chair, VUB/Starlab
OOA Board
VUB - STAR Lab
Pleinlaan 2 - Gebouw G - 10
1050 Brussels
Tel: +32 (0)2 629 12.37

Dear Professor Meersman

I hereby declare that (company name) _____ is interested in joining the "Ontology Outreach Advisory" initiative as member. _____ (company name) accepts to help promote the use of semantics and ontology's in the HR domain within my industry sector or governmental organization. The OOA activities will be attended by: _____ (company representative name).

Name:

Date:

Signature:

9.3 Example of Evaluation Form

Evaluation Category	Evaluation Criterion	Value	Explanation
Ontological Content	**Ontology domain**:	A list of keywords describing the domain of the ontology. The keywords might belong to a pre-defined vocabulary.	This information is required in order to assign the submitted ontology appropriately to the corresponding OOA chapter.
	Content type	One of the following categories: knowledge representation ontology, upper-level ontology, domain ontology, task ontology, application ontology [127]. The classification scheme is subject of further discussions.	The classification of the ontology w.r.t. the level of generality of its domain. This information is required in order to assign the submitted ontology appropriately to the corresponding OOA chapter and to specify a particular evaluation strategy of the OOA.
	Formality level	One of the following categories: controlled vocabulary, glossary, thesaurus, informal taxonomy, formal taxonomy, formal taxonomies containing formal instances, frames, restrictions etc. See [128] for a more detailed description of the formality levels.	The classification of the ontology w.r.t. the richness of their structure, which corresponds to particular formality level. This information is important for the selection of an appropriate evaluation strategy by the OOA.
Information Sources	**Sources used for knowledge acquisition**	A list of documents which have been used for knowledge acquisition purposes during the building of the ontology: dictionaries, thesauri, user interviews and glossaries from which the ontology was derived.	These information sources provide additional input for the description of the ontology domain and might trigger a specific evaluation procedure.
	Directives	A list of directives, official documents and standards, laws or regulations in a particular domain with which an ontology might be compliant with.	

Application scenario	**Application system**	A document describing the application system the submitted ontology was originally designed for. Application systems might reference to pre-defined vocabularies for industrial sectors or software classifications.	Knowledge about the application scenario is important for the evaluators as this specifies the requirements the submitted ontology claims to fulfil.
	Ontology use cases	A document describing the use cases in which the ontology is involved.	
	Ontology task	A list of keywords describing the system/technical tasks the ontology is originally developed for (e.g. information extraction from documents).	
	Ontology requirements	The ontology requirements specification document	
Evaluation	**Evaluation method**	A list of methods used to evaluate the ontology in each stage of the ontology life cycle (pre-development, post-development etc.). Documentation of the proprietary methods should be enclosed.	The authoring organization should provide information about the syntactic validity, the logical consistency, application-specific evaluation activities and/or general-purpose evaluation activities (such as expert reviewing, ontological evaluation etc.).
	Evaluation tool	A reference to the tools used as evaluation support	
	Evaluation results	A document reporting on the evaluation results.	
Documentation		The complete documentation of the engineering process, including comments and definitions of the ontology primitives, information about the engineering methodology, design guidelines and patterns eventually applied to build the ontology	This information should be accessible to the OOA chapter in case a more detailed investigation of the submitted ontology is required.

9.4 Example of Evaluation Metadata

```
<?xml version="1.0"?>

<rdf:RDF
        xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns #
        xmlns:xsd=http://www.w3.org/2001/XMLSchema#
        xmlns:rdfs=http://www.w3.org/2000/01/rdf-schema#
        xmlns=http://http://nbi.inf.fu-berlin.de/research/swpatho/context/omv-
        evaluation.owl#
        xmlns:owl=http://www.w3.org/2002/07/owl#
        xmlns:omv=" http://omv.ontoware.org/2005/05/ontology# "
        xml:base=http://http://nbi.inf.fu-berlin.de/research/swpatho/context/omv-
        evaluation.owl>

<owl:Ontology rdf:about=””>
        <owl:imports rdf:resource=http://omv.ontoware.org/2005/05/ontology#/>

</owl:Ontology>

<OntologyImplementation rdf:about="#swpatho1">
        <omv:implementationAcronym
        rdf:datatype="...XMLSchema#string">SWPATH=1</omv:implementationAcronym>
        <omv:ontologyLanguage
        rdf:resource=http://omv.ontoware.org/2005/05/ontology#owl-dl/>
        <omv:naturalLanguage
        rdf:datatype="...XMLSchema#string">de</omv:naturalLanguage>
        <omv:implementationDescription rdf:datatype="...XMLSchema#string">The
        implementation of the swpatho ontology in OWL DL.
        </omv:implementationDescription>
        <omv:versionInfo rdf:datatype="...XMLSchema#string">1.0</omv:versionInfo>
        <omv:implements>
                <omv:OntologyConceptualisation rdf:about="#swpatho">
                        <omv:conceptualisationDocumentation
                        rdf:datatype="...XMLSchema#anyURI"> The ontology describes
                        concepts of lung anatomy and lung diseases and is used as basis for
                        retrieval and semantic annotation tasks on medical data. It is based on
                        common medical libraries like SNOMED, DigitalAnatomist and the
                        UMLS Semantic Network.
                        </omv:conceptualisationDocumentation>
                        <omv:conceptualisationDescription
                        rdf:datatype="...XMLSchema#string">An ontology of lung
                        pathology</omv:conceptualisationDescription>
                        <omv:conceptualisationAcronym
                        rdf:datatype="...XMLSchema#string">SWPATHO
                        </omv:conceptualisationAcronym>
                        <omv:conceptualisationName
                        rdf:datatype="...XMLSchema#string">Semantic Web for Pathology
                        </omv:conceptualisationName>

                        <omv:conceptualisationDomain>
                                <omv:OntologyDomain rdf:about="#Pathology"/>
```

```
                    </omv:conceptualisationDomain>
                    <omv:conceptualisationType rdf:resource=omv#domain_ontology/>
                </omv:OntologyConceptualisation>
        </omv:implements>
        <omv:numAxioms rdf:datatype="...XMLSchema#int">34</omv:numAxioms>
        <omv:creationDate
        rdf:datatype="...XMLSchema#date">15.07.2004</omv:creationDate>
        <omv:numIndividuals
        rdf:datatype="...XMLSchema#int">400</omv:numIndividuals>
        <omv:status rdf:datatype="...XMLSchema#string">stable</omv:status>
        <omv:hasSyntax rdf:resource="http://omv.ontoware.org/2005/05/ontology#rdf-
        xml"/>
        <omv:designedForOntologyTask>
                <omv:AnnotationTask rdf:about="#annotationtask">
                <omv:taskName rdf:datatype="...XMLSchema#string">Semantic Annotation
                Task
                </omv:taskName>
                <omv:taskDescription rdf:datatype="...XMLSchema#string">Annotation of
                pathology reports
                </omv:taskDescription>
                </omv:AnnotationTask>
        </omv:designedForOntologyTask>
        <omv:numClasses rdf:datatype="...XMLSchema#int">1200</omv:numClasses>
        <omv:numProperties rdf:datatype="...XMLSchema#int">50</omv:numProperties>
        <omv:implementationType rdf:resource=omv:application_ontology/>
        <omv:implementationName rdf:datatype="...XMLSchema#string">Semantic Web for
        Pathology
        </omv:implementationName>
        <omv:ontologyURL rdf:datatype="...XMLSchema#anyURI">http://nbi.inf.fu-
        berlin.de/research/swpatho/owldata/swpatho1.owl</omv:ontologyURL>
        <omv:naturalLanguage rdf:datatype="...XMLSchema#string">
                en
        </omv:naturalLanguage>
</omv:OntologyImplementation>

<OntologyEvaluationProcess rdf:about="#swpatho1evaluation">
        <usesMethod rdf:about=”#consistencyOWLDL”/>
        <usesDataSet rdf:about=”#...”/>
</OntologyEvaluationProcess>
<LogicalEvaluation rdf:about=”consistencyOWLDL”>
        <usesTool rdf:about=”#racer1.8.2”/>
</LogicalEvaluation>
<EvaluationDataSet rdf:about="#http://…"/>
<OntologyDirective rdf:about="#pathologyreportcorpus">
        <directiveAcronym rdf:datatype="...XMLSchema#string">
                SWPATHOCORPUS
        </directiveAcronym>
        <directiveDescription rdf:datatype="...XMLSchema#string">
                A corpus of 700 medical reports in XML form
        </directiveDescription>
</OntologyDirective>
```

10 References

[1] KWEB/2004/D1.1.1/v2

[2] http://www.loa-cnr.it/core_onto.html (Last visited Sept. 2005)

[3] Guarino, N., Welty, C., and Christopher, P. 2000. Towards Ontology-based harmonization of Web content standards. In S. Liddle, H. Mayr and B. Thalheim (eds.), Conceptual Modeling for E-Business and the Web: Proceedings of the ER-2000 Workshops. Springer Verlag: 1-6.

[4] Jarrar, M.: Towards Methodological Principles for Ontology Engineering PhD thesis, Vrije Universiteit Brussel, May 2005.

[5] López-Pérez, A.: "Integración de la aplicación OntoMetric en la plataforma WebODE". Proyecto fin de carrera. Facultad de Informática (Universidad Politécnica de Madrid). 2004.

[6] http://lirics.loria.fr/index.html. Last visited: Nov 12th, 2005.

[7] http://syntax.inist.fr/. Last visited: Nov 12th, 2005.

[8] http://syntax.inist.fr/helpManual/index.php. Last visited: Nov 12th, 2005.

[9] http://ncor.us/inaugural. Last visited: Dec 21, 2005.

[10] http://www.oasis-open.org/home/index.php. Last visited: Dec 21, 2005.

[11] Obrst, L., R. Wray, H. Liu. 2001. Ontological Engineering for B2B E-Commerce. In: Formal Ontology in Information Systems: Collected Papers from the Second International Conference, October 17-19, 2001, Ogunquit, ME. Chris Welty, Barry Smith, eds. ACM Publishing, http://www.fois.org/fois-2001/index.html.

[12] www.moniteur.be. Last visited: November 12th 2005.

[13] Suárez-Figueroa, M., García-Castro, R., Gómez-Pérez, A., Palma R., Nixon, L., Paslaru, L., Hartmann J., Jarrar, J.: Identification of standards on metadata for ontologies. Deliverable D1.3.2. EU-IST Network of Excellence (NoE) IST-2004-507482 (KWEB), Luxemburg (2005).

[14] http://ontoware.org/projects/omv/. Last visited: November 21, 2005.

[15] Jouanna J. The Birth of Western Medical Art. In: Grmek MD, editor. Western Medical Thought from Antiquity to the Middle Ages. Shugaar A, translator. Cambridge (MA): Harvard University Press; 1998: 22–71.

[16] Groninger H. What do we mean by 'the art of medicine'. The Faros, Automn 2004, 27-30.

[17] W.J. Mackillop. The Importance Of Prognosis In Modern Medicine. 16th International MASCC/ISOO Symposium, Miami Florida - June 24-27 2004.

[18] Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312:71-72

[19] Wiley V, Carpenter K, Bennetts B, Wilcken B. Information overload--new technologies, can we store the data? Southeast Asian J Trop Med Public Health. 2003;34 Suppl 3:59-62.

[20] Boyle P, Ferlay J. Cancer incidence and mortality in Europe, 2004. Annals of Oncology 2005 16(3):481-488.

[21] House of Commons, Public Accounts Committee. Tackling cancer in England: saving more lives. Second Report. 25 January 2005 .

[22] Newman Dorland. The American illustrated medical dictionary (1923).

[23] Keating P, Cambrosio A. Does biomedicine entail the successful reduction of pathology to biology? Perspect Biol Med. 2004 Summer;47(3):357-71.

[24] Hersh WR. Medical Informatics: Improving Health Care Through Information. JAMA. 2002;288:1955-1958.

[25] Eich HP, de la Calle G, Diaz C, Boyer S, Pena AS, Loos BG, Ghazal P, Bernstein I. Practical Approaches to the Development of Biomedical Informatics: the INFOBIOMED Network of Excellence. Stud Health Technol Inform. 2005;116:39-44.

[26] Sanz F, Diaz C, Martin-Sanchez F, Maojo V. Structuring European biomedical informatics to support individualized healthcare: current issues and future trends. Medinfo. 2004;11(Pt 2):803-7.

[27] Dennis JL and Oien KA. Hunting the primary: novel strategies for defining the origin of tumours. Journal of Pathology. Vol.205:236-247. 2005.

[28] Carr KM, Rosenblatt K, Petricoin EF and Liotta LA. Genomic and proteomic approaches for studying human cancer: prospects for true patient-tailored therapy. Human Genomics. Vol 1(2): 134-140. 2004.

[29] Strausberg RL, Simpson AJG, Old LJ and Riggins GJ. Oncogenomics and the development of new cancer therapies. Nature. Vol 429: 469-474. 2004.

[30] Anand Kumar1; Barry Smith; Daniel D. Novotny. Biomedical informatics and granularity. Comparative and Functional Genomics, Volume 5, Numbers 6-7, August 2004, pp. 501-508(8).

[31] Outreach in KnowledgeWeb: Use cases, stakeholders and solution providers in Healthcare and The Life Sciences

[32] Guarino N., (1998), Formal Ontologies and Information Systems, in Guarino N. (ed), Proc. of FOIS98, IOS Press, pp.3 – 15

[33] The Description Logic Handbook. Franz Baader, Deborah McGuinness, Daniele Nardi (eds), Cambridge University Press, 2003.

[34] Farrugia J., (2003), Model-Theoretic Semantics for the Web, in Proceedings of the 12th International Conference on the WWW, ACM, pp. 29 – 38.

[35] Ceusters W, Smith B. Strategies for Referent Tracking in Electronic Health Records. Journal of Biomedical Informatics. In press.

[36] Kumar A, Smith B. The Unified Medical Language System and the Gene Ontology, KI2003: Advances in Artificial Intelligence (Lecture Notes in Artificial Intelligence 2821), 2003; 135–48.

[37] Smith B, Williams J, Schulze-Kremer S. The ontology of the gene ontology. In: Musen MA, editor. AMIA 2003. Proceedings of AMIA 2003 Annual Symposium; 2003 Nov 8-12, Washington D.C., USA. AMIA; 2003. p. 609-13.

[38] Grenon P, Smith B, Goldberg L. Biodynamic ontology: Applying BFO in the Biomedical Domain, in Pisanelli DM (ed). Ontologies in Medicine. Proceedings of the Workshop on Medical Ontologies, Rome October 2003. IOS Press, Studies in Health Technology and Informatics, vol 102, 2004. p. 20-38.

[39] Ceusters W, Smith B, Kumar A, Dhaen C. Ontology-Based Error Detection in SNOMED-CT®. In: M. Fieschi, E. Coiera and Y-C.J. Li, editors. MEDINFO 2004. Proceedings of the 11th World Congress on Medical Informatics; 2004 Sep 7-11, San Francisco, CA, USA. Amsterdam: IOS Press; 2004. p. 482-6.

[40] Smith B, Rosse C. The role of foundational relations in the alignment of biomedical ontologies. In: M. Fieschi, E. Coiera and Y-C.J. Li, editors. MEDINFO 2004. Proceedings of the 11th World Congress on Medical Informatics; 2004 Sep 7-11, San Francisco, CA, USA. Amsterdam: IOS Press; 2004. p. 444-8.

[41] Kumar A, Schulze-Kremer S, Smith B. Revising the UMLS Semantic Network. In: M. Fieschi, E. Coiera and Y-C.J. Li, editors. MEDINFO 2004. Proceedings of the 11th World Congress on Medical Informatics; 2004 Sep 7-11, San Francisco, CA, USA. Amsterdam: IOS Press; 2004. p. 1700-4.

[42] Ceusters W, Smith B. A Terminological and Ontological Analysis of the NCI Thesaurus. Methods of Information in Medicine 2005; 44: 498-507.

[43] http://www.snomed.org/snomedct/index.html. Last visited: Nov 6th, 2005.

[44] http://www.cap.org/apps/cap.portal. Last visited: Nov 6th, 2005.

[45] Spackman KA, Campbell KE, Cote RA. SNOMED RT: a reference terminology for health care. Proc AMIA Annu Fall Symp. 1997;:640-4.

[46] http://www.nhsia.nhs.uk/terms/pages/faq.asp?om=m1. Last visited: Nov 6th, 2005.

[47] Chiang MF, Casper DS, Cimino JJ, Starren J. Representation of ophthalmology concepts by electronic systems: adequacy of controlled medical terminologies. Ophthalmology. 2005 Feb;112(2):175-83.

[48] Penz JF, Brown SH, Carter JS, Elkin PL, Nguyen VN, Sims SA, Lincoln MJ. Evaluation of SNOMED coverage of Veterans Health Administration terms. Medinfo. 2004;11(Pt 1):540-4.

[49] Wasserman H, Wang J. An applied evaluation of SNOMED CT as a clinical vocabulary for the computerized diagnosis and problem list. AMIA Annu Symp Proc. 2003;:699-703.

[50] Humphreys BL, McCray AT, Cheh ML. Evaluating the coverage of controlled health data terminologies: report on the results of the NLM/AHCPR large scale vocabulary test. J Am Med Inform Assoc. 1997 Nov-Dec;4(6):484-500.

[51] Strang N, Cucherat M, Boissel JP. Comput Methods Programs Biomed. 2002 Apr;68(1):73-85. Which coding system for therapeutic information in evidence-based medicine.

[52] Brown SH, Bauer BA, Wahner-Roedler DL, Elkin PL. Coverage of oncology drug indication concepts and compositional semantics by SNOMED-CT. AMIA Annu Symp Proc. 2003;:115-9.

[53] Smith B, Ceusters W. An Ontology-Based Methodology for the Migration of Biomedical Terminologies to Electronic Health Records. AMIA 2005, October 22-26, Washington DC

[54] S. de Coronado, M. W. Haber, N. Sioutos, M. S. Tuttle, and L. W. Wright. NCI Thesaurus: Using Science-based Terminology to Integrate Cancer Research Results. In: M. Fieschi, E. Coiera and Y-C.J. Li, editors. MEDINFO 2004. Proceedings of the 11th

World Congress on Medical Informatics; 2004 Sep 7-11, San Francisco, CA, USA. Amsterdam: IOS Press; 2004. p. 33-7.

[55] Open Biological Ontologies. http://obo.sourceforge.net/. Last visited 2005, Oct 24.

[56] National Cancer Institute, Office of Communications, Center for Bioinformatics. NCI Terminology browser, ftp://ftp1.nci.nih.gov/pub/cacore/EVS/. Last visited 2005, Jan 18.

[57] National Cancer Institute, Office of Communications, Center for Bioinformatics. NCI Terminology browser, http://nciterms.nci.nih.gov/NCIBrowser/Startup.do. Last visited 2005, Oct 18.

[58] Hahn U, Schulz S. Towards a broad-coverage biomedical ontology based on description logics. Pac Symp Biocomput. 2003:577-88.

[59] Cimino J. Auditing the Unified Medical Language System with Semantic Methods. J Am Med Inform Assoc. 1998 January; 5 (1): 41–51.

[60] Rector AL, Nowlan WA, Glowinski AJ. Goals for Concept Representation in the GALEN project. 17th annual Symposium on Computer Applications in Medical Care, Washington, USA, SCAMC 93.pp, 414-418

[61] http://www.opengalen.org/. Last visited: Nov 12th, 2005.

[62] Rector AL, Rogers JE, Zanstra PE, Van Der Haring E. OpenGALEN: open source medical terminology and tools. AMIA Annu Symp Proc. 2003;:982.

[63] Heja G, Surjan G, Lukacsy G, Pallinger P, Gergely M. GALEN Based Formal Representation of ICD10. Stud Health Technol Inform. 2005;116:707-12

[64] Rodrigues JM, Trombert Paviot B, Martin C, Vercherin P. Integrating the Modelling of EN 1828 and Galen CCAM Ontologies with Protege: towards a Knowledge Acquisition Tool for Surgical Procedures. Stud Health Technol Inform. 2005;116:767-72.

[65] Alan Rector and Jeremy Rogers. Ontological Issues in using a Description Logic to Represent Medical Concepts: Experience from GALEN. IMIA WG6 Workshop: Terminology and Natural Language in Medicine. Phoenix Arixona, Nov 1999. http://ontology.buffalo.edu/Rector_Ontological_Issues.pdf.

[66] Donnelly M, Bittner T, Rosse C. A formal theory for spatial representation and reasoning in biomedical ontologies. Artif Intell Med. 2005 Oct 21; [Epub ahead of print]

[67] Barry Smith, Anand Kumar, Thomas Bittner. Basic Formal Ontology for Bioinformatics. http://www.uni-leipzig.de/~akumar/JAIS.pdf

[68] Rosse C, Kumar A, Mejino JLV, Cook DL, Detwiler LT, Smith B. A strategy for improving and integrating biomedical ontologies. Proc AMIA Symp 2005, Washington DC, in press.

[69] Barry Smith, Anand Kumar, Werner Ceusters and Cornelius Rosse. On carcinomas and other pathological entities (Submitted to SMBI).

[70] http://sig.biostr.washington.edu/projects/fm/

[71] Rosse C, Mejino JLV Jr: A reference ontology for bioinformatics: The Foundational Model of Anatomy. *J Biomed Informatics*, 2003; 36:478-500.

[72] http://www.geneontology.org/. Last visited 2005, Nov 9.

[73] http://www.mged.org/. Last visited 2005, Nov 9.

[74] http://obo.sourceforge.net/. Last visited 2005, Nov 9.

[75] Smith B, Ceusters W, Klagges B, Kohler J, Kumar A, Lomax J, Mungall C, Neuhaus F, Rector AL, Rosse C. Relations in biomedical ontologies. Genome Biol. 2005;6(5):R46. Epub 2005 Apr 28.

[76] http://ncicb.nci.nih.gov/core/caDSR. Last visited: October 28, 2005.

[77] ISO/IEC 11179: Information technology — Metadata registries (MDR) —. http://isotc.iso.org/livelink/livelink/fetch/2000/2489/Ittf_Home/PubliclyAvailableStandards.htm.

[78] http://www.eurorec.org/projects/qrec.cfm?actief=Q-Rec. Last visited: November 12th, 2005.

[79] http://www.eurorec.org/. Last visited: November 12th, 2005.

[80] http://www.srdc.metu.edu.tr/webpage/projects/ride/index.php. Last visited: November 12th, 2005.

[81] http://philosophy.buffalo.edu/ARCHIVE/features/PressRelease9-29-05.pdf.

[82] http://bioontology.org/dissemination.html. Last visited: November 12th, 2005.

[83] http://bphc.hrsa.gov/. Last visited: Nov 9, 2005.

[84] BPHC. Electronic Medical Record (EMR) Functional Requirements VERSION 5.2 11.9.2005. http://www.bphc.hrsa.gov/chc/CHCInitiatives/emrspecs.doc.

[85] http://bphc.hrsa.gov/chc/emrspecs.htm. Last visited: Nov 9, 2005.

[86] Karen Strandgaard. Innovative Medicines for Europe Abbreviation (InnoMed): Creating biomedical R&D leadership for Europe to benefit patients and society. December 2004. http://europa.eu.int/comm/research/fp6/p1/innovative-medicines/pdf/vision_en.pdf

[87] http://www.ifomis.uni-saarland.de/. Last visited: Nov 9, 2005.

[88] http://www.ecor.uni-saarland.de/announce/eurorecontologyworkshop.html. Last visited: Dec 21, 2005.

[89] http://www.prorec.be/events/eurorec2004/eurorec/17-WernerCeusters.pps. Last visited: Dec 21, 2005.

[90] http://www.ecor.uni-saarland.de/papers/Diepenbeek20041202.pps. Last visited: Dec 21, 2005.

[91] http://www.who.int/classifications/terminology/en/. Last visited: Dec 21, 2005.

[92] http://www.who.int/entity/classifications/terminology/ceusters.pdf. Last visited: Dec 21, 2005.

[93] http://www.ecor.uni-saarland.de/papers/Pres20050408.pps. Last visited: Dec 21, 2005.

[94] http://www.ecor.uni-saarland.de/papers/SMBM2005.pps. Last visited: Dec 21, 2005.

[95] http://www.ecor.uni-saarland.de/papers/OOAHCoverview.pps. Last visited: Dec 21, 2005.

[96] http://www.mie2005.net/tutorials/T035.PDF. Last visited: Dec 21, 2005.

[97] http://www.ecor.uni-saarland.de/papers/NCOR20051027.pps. Last visited: Dec 21, 2005.

[98] http://www.ecor.uni-saarland.de/papers/Paris2005.ppt. Last visited: Dec 21, 2005.

[99] www.jobpilot.com. Last visited: Nov 9, 2005.

[100] www.arbeitsagentur.de. Last visited: Nov 9, 2005.

[101] www.ams.se. Last visited: Nov 9, 2005.

[102] http://www.aspanet.org/solutionstemp/jobport.html. Last visited: Nov 9, 2005.

[103] D. Maynard, M. Yankova, A. Kourakis and A. Kokossis. Ontology-based information extraction for market monitoring and technology watch, ESWC Workshop "End User Apects of the Semantic Web", Heraklion, Crete, 2005.

[104] http://www.job-search.com. Last visited: Nov 9, 2005.

[105] http://www.innovantage.co.uk/. Last visited: Nov 9, 2005.

[106] http://ieeeltsc.org/wg20comp. Last visited: Dec 21, 2005.

[107] http://i-know.knowcenter.tugraz.at/content/download/382/1523/file/Gualtieri_paper.pdf

[108] http://www.eil.utoronto.ca/enterprise-modelling/tove/index.html. Last visited: Nov 9, 2005.

[109] http://proton.semanticweb.org/. Last visited: Nov 9, 2005.

[110] http://www.hr-xml.org. Last visited: Nov 9, 2005.

[111] http://www.vl-leonardo.be/CODRIVE.html. Last visited: Dec 21, 2005.

[112] http://www.lri.jur.uva.nl/jurix2001/legont2001.htm#_ftn2. Last visited: Nov 9, 2005.

[113] LOAIT Workshop: Legal Ontologies and Artificial Intelligence Techniques. June 6, 2005, Bologna, Italy, held in conjunction with ICAIL-05.

[114] http://www.loisproject.org/. Last visited: Nov 9, 2005.

[115] On The Move to Meaningful Internet Systems 2003: OTM 2003Workshops. Workshop on Regulatory Ontologies and the Modelling of Complaint Regulations (WORM CoRe). Lecture Notes in Computer Science, Volume 2889 / 2003 Springer-Verlag GmbH , October 2003.

[116] Zeibert, LA - Smith, B: Legal Ontology and the Problem of Normativity, http://www.ontology.buffalo.edu/smith/articles/Normativity.pdf.

[117] A Daskalopulu & MJ Sergot (1997). The Representation of Legal Contracts, AI and Society 11 (Nos. 1/2), pp. 6-17.

[118] Jaime Delgado, Isabel Gallego, Silvia Llorente, and Roberto Garc´ıa, 'IPROnto: An Ontology for Digital Rights Management' in D. Bourcier (ed.), Legal Knowledge and Information Systems. Jurix 2003: The Sixteenth Annual Conference. Amsterdam: IOS Press, 2003, pp. 111-120.

[119] Ceusters W, Smith B. Referent Tracking for Digital Rights Management. Submitted to SWCASE05.

[120] http://www.metalex.nl/. Last visited: Nov 12, 2005.

[121] Niles, I., and Pease, A. 2001. Towards a Standard Upper Ontology. In Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), Chris Welty and Barry Smith, eds, Ogunquit, Maine, October 17-19, 2001.

[122] http://dip.semanticweb.org/documents/D10.3.pdf. Last visited: Nov 12, 2005.

[123] Second International Workshop on Semantic Web for web-based learning. Implications in the area of Information Systems in education. 13-17 June 2005, Portugal.

[124] Wimmer, M.A. 2001. European Development Towards Online One-stop Government: the E-GOV project". Proceedings of ICEC2001, Vienna, Austria.

[125] A. Gugliotta. 2005. Knowledge Modelling for Integrating Semantic Web Services in E-Government Applications. Proceedings of ISWC 2005 Doctoral Symposium, Galway, Ireland.

[126] OECD. Engaging citizens in policy making: information, consultation and public participation. PUMA Policy Brief No 10, July 2001.

[127] Gómez-Pérez, A., Fernández-López, M. and Corcho, O. Ontological Engineering. Springer Verlag, 2004.

[128] Lassila, O. and McGuinness, D. The role of frame-based representation on the Semantic Web. Technical report KSL-01-02. Knowledge Systems Laboratory. Stanford University. 2001