D 2.3.2 Specification of Knowledge Acquisition and Modeling of the Process of the Consensus

Coordinator: Anna V. Zhdanova (UIBK)

Matteo Bonifacio (UniTn), Stamatia Dasiopoulou (CERTH), Jérôme Euzenat (INRIA), Rose Dieng-Kuntz (INRIA), Loredana Laera (LivUni), David Manzano-Macho (UPM), Diana Maynard (USFD), Diego Ponte (UniTn), Valentina Tamma (LivUni)

Abstract.
EU-IST Network of Excellence (NoE) IST-2004-507482 KWEB
Deliverable D2.3.2 (WP2.3)
In this deliverable, specification of knowledge acquisition and modeling of the process of consensus is provided.

Document Identifier:	KWEB/2004/D2.3.2/v1.0
Class Deliverable:	KWEB EU-IST-2004-507482
Version:	v1.0
Date:	January 28, 2005
State:	Final
Distribution:	Public

Knowledge Web Consortium

This document is part of a research project funded by the IST Programme 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	**University of Karlsruhe (UKARL)** Institut für Angewandte Informatik und Formale Beschreibungsverfahren – AIFB Universität Karlsruhe D-76128 Karlsruhe

E-mail address: asun@fi.upm.es	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: hamish@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: robert.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:

- National University of Ireland Galway (NUIG)
- University of Innsbruck (UIBK)
- University of Karlsruhe (UKARL)
- Universidad Politécnica de Madrid (UPM)
- University of Liverpool (UniLiv)
- Vrije Universiteit Amsterdam (VUA)
- Institut National de Recherche en Informatique et en Automatique (INRIA)
- Free University of Bozen-Bolzano (FUB)
- Centre for Research and Technology Hellas / Informatics and Telematics Institute (ITI-CERTH)
- University of Sheffield (USFD)

Changes

Version	Date	Author	Changes
0.1	28-05-2004	Anna Zhdanova	First outline
0.2	12-08-2004	Anna Zhdanova	UniTn, USFD inputs added
0.3	12-09-2004	Anna Zhdanova	CERTH, INRIA inputs added
0.4-0.9	05-15-12-2004	Anna Zhdanova	Added inputs of UPM, LivUni, final pre-QA version
0.91	20-12-2004	Anna Zhdanova	Added revised UniTn input
1.0	28-01-2005	Anna Zhdanova	Corrections with regards to QA comments are integrated, final version

Executive Summary

In this deliverable, a specification of knowledge acquisition and modeling of the process of the consensus are described. The deliverable is structured to give an overview of theoretical and practical approaches to knowledge acquisition and consensus making, that have potential to contribute and be combined into improved knowledge acquisition and consensus making processes. Knowledge acquisition is considered both in terms of knowledge acquisition from information sources and in terms of knowledge acquisition of views of experts. Both knowledge acquisition notions are important for consensus process modeling, since consensus making process is tightly connected with analysis and acquisition of different knowledge types, including direct communication flows from communicating parties and extraction and delivery of the relevant knowledge to communicating parties for decision making.

The deliverable consists of three major conceptually important parts:
1) survey of existing techniques/methodologies for knowledge acquisition and consensus modeling (theory),
2) survey of existing tools and prototypes for knowledge acquisition and consensus making (practice),
3) specification for a solution for knowledge acquisition and consensus modeling, and outline of the implementation supporting the consensus making framework.

Contents

1 Introduction

Originally, an ontology should reflect the "truth" of a certain aspect of reality. Earlier in history, to find such truth was a task of a philosopher. Today, ontologies are used as a means of exchanging meaning between different agents. Ontologies can only provide support for exchange of meaning if they reflect an inter-subject consensus. Consensus implies involvement of multiple possessors of heterogeneous knowledge parties (e.g., individuals, communities, user groups, agents) reaching an agreement in a social process. Association with a social process gives ontologies a dual status for the exchange of meaning:

- Ontologies as pre-requisite for consensus: Agents can only exchange meaning when they have already agreed on a joint body of meaning reflecting a consensual point of view on the world.
- Ontologies as a result of consensus: Ontologies as consensual models of meaning can only arise as result of a process where agents agree on a certain model of the world and its interpretation.

Thus, ontologies are as much a pre-requisite for consensus and information sharing, as they are the results of them. For this reason, ontologies cannot be understood as a static model. An ontology is as much required for the exchange of meaning as the exchange of meaning may influence and modify an ontology. Consequently, *evolving* ontologies describe a process rather than a static model. Having protocols for the process of evolving ontologies is the real challenge. Evolution over time is an essential requirement for useful ontologies. As the daily practice constantly changes, ontologies that mediate the information requirements of these processes must have strong support in *versioning* and must be accompanied by *process models* that help to organize consensus.

Evaluation and *meaning negotiation* are other two areas which are to a certain extent reflected in the deliverable, due to their close association with the main topic of the deliverable, i.e., knowledge acquisition and consensus processes modeling.
***Evaluation**:* The ontology evaluation field is just emerging. From the methodological perspective, content evaluation activities should be included in more detail in ontology building methodologies. The purpose of these activities is to raise ontological engineers' awareness of the fact that evaluation should be performed throughout the entire ontology life cycle in order to detect errors at the earliest possible time, and should not be left until the end when the ontology has been implemented.
Meaning Negotiation: Another aspect of dynamic ontologies relates to the fact that agents will inevitably encounter agents with a different ontological history. Successfully interacting with such agents will require the ability *to reach a dynamic consensus* on a shared ontology while maintaining the integrity of the agent's original ontology base, and while extending capabilities to adapt to new concepts, facts, and protocols. The ability of agents to dynamically negotiate will be critical here: both over object level issues ("how much should I pay for this service?") and over meta-level issues ("how should we refer to this concept?").

The deliverable is organized as follows. In Section 2, definitions and contexts of usage of the core terms used in the deliverable are provided. In Section 3, the theory of knowledge acquisition and consensus modeling in described from different perspectives: human interaction in semantic web environments, collaborative ontology construction, approaches arriving from agent negotiation studies, sociology/economy view, etc. In section 4, a deeper insight on the technology and implementation details of the systems supporting theories, which are described in Section 3, is given. In Section 5, an approach to the proposed consensus modeling solution and its implementation are addressed. Section 6 concludes the deliverable.

2 Knowledge Acquisition and Consensus Modeling – Definitions

In this section, we define or describe our understanding of the core terms used in the deliverable and the context in which we use these terms.

2.1 Knowledge Acquisition

Several methods, techniques and protocols were proposed for *knowledge acquisition (KA) from multiple experts*. They aim at: (a) expressing common parts and differences between experts' models; (b) detecting and solving terminology conflicts between the experts; (c) taking into account the different viewpoints of different experts: several experts according to their specialty or their way to tackle the problem solving, may have divergent analyses or divergent understandings of a same object. So, the notions of multi-expertise and multi-viewpoints are closely related.

An expert can have different viewpoints or perspectives on a domain or on a problem; so, for each expert, the knowledge engineer (KE) can identify several viewpoints (see Figure 1).

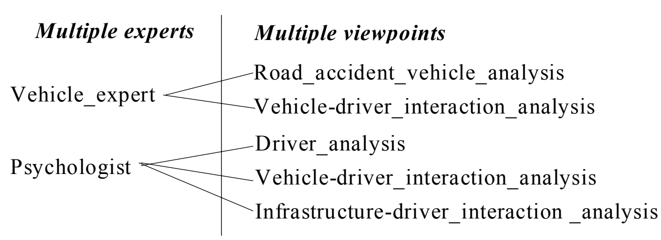

Figure 1: Example of link multi-expertise/multi-viewpoints

The viewpoints of each expert are actually the different points of interest of this expert concerning the application. This relation establishes a partition of the knowledge base (KB) where each viewpoint is a coherent and partial view of the KB. But as in the example of Figure 1, some viewpoints may be shared. This leads us to consider the relationship between multi-expertise and multi-viewpoints otherwise. Indeed, on a same identified viewpoint, there may be different “viewpoints” of different experts from the same domain or from different domains. In that case, the knowledge engineer must either integrate knowledge from the different experts or make those different “viewpoints” on a viewpoint live together.

As mentioned above, the term knowledge acquisition has roots in the field of expert systems research. On the other hand, knowledge acquisition is also a broader knowledge engineering area covering theories and approaches to

- collect knowledge,
- deliver knowledge,
- communicate knowledge,
- share knowledge,
- reuse knowledge

All the listed above actions can be automatic, semi-automatic and performed without involvement of automation. Naturally, when modeling and implementing a process of consensus that presumes operation with knowledge, we are interested in automation and improvement in knowledge annotation, processing and delivery of different types of information, such as natural language texts, multimedia, etc. Therefore, the relevant works in the corresponding knowledge processing areas are considered.

2.2 Ontology

Ontology is a *specification of a shared conceptualization* [Gruber93]. “Shared” requires consensus in community employing ontologies as the means of information exchange. Consensus as common understanding and agreement can only be the result of a social process involving individuals and communities. Thus, ontologies have a dual status in information exchange:

- Ontologies as pre-requisite for consensus: Agents can only exchange information when they have already agreed on a joint body of specification reflecting a consensual point of view on the world.
- Ontologies as a result of consensus: Ontologies as consensual models of meaning can only arise as result of a process where agents agree on a certain model of the world and its interpretation.

2.3 Consensus, Conflicts, Viewpoints

In this subsection, we address the terms related to consensus, as they are seen from the common sense point of view, and in ontology-related research.

2.2.1 Consensus

Generally, consensus can be seen as understanding ontology in the same way and interoperating in a consistent, mutually beneficial way.

In the Webster's dictionary "consensus" is defined as "agreement; accord; consent". In the Semantic Web context, the consensus can be reached at the data level (e.g., "how much should I pay for this service?") and at the metadata level (e.g., "how should we refer to this concept?"). An ability to reach a consensus at both data and metadata levels is a must for successful cross-application interaction.
The necessity to cope with the following issues makes reaching understanding and agreement between two or more parties a difficult task:
Dynamicity (i), i.e., rapid change of outside world, its conceptualization and specification of conceptualization,
heterogeneity (ii), i.e., presence of various description formats and ontological histories.
In addition, the process of reaching a consensus is often followed by the requirement of *maintaining the integrity of the parties' original ontology bases* (iii).
The requirement (iii) meets the common need to have an opportunity of access the data via once used schemata and protocols while extending capabilities to adapt to new concepts, facts, rules and processes.

2.2.2 Conflict Types

We recognize conflicts at three levels:
Conflicts at Syntactic Level (problem of interpretation)
e.g., you speak F-logic and I speak OWL, and we do not understand each other,…
Conflicts at Semantic Level (problem of understanding)
E.g., you say "a" and mean "b", and I understand your "a" as "c", …
Conflicts at Conceptual Level (problem of coming to an agreement)
e.g., you say "you have to restructure your ontology", I understand you, but what you propose does not fit me. How can something be done to come up with something that would fit both of us?

In modeling of the process of consensus, we focus on the conflicts of the Conceptual level and refer to solutions for the conflicts at the Semantic level (in terms of knowledge processing and re-formulation). Conflicts on the Syntactic level are out of scope if our deliverable.

2.2.3 Viewpoints

Ribière and Dieng-Kuntz (Ribière, 1999; Ribière and Dieng-Kuntz, 2002) define a viewpoint as " an interface allowing the indexation and the interpretation of a view composed of knowledge elements. A viewpoint is characterized by a focus (corresponding to a contextual dimension) and a view angle (corresponding to a personal dimension)". More precisely:

- The focus describes the expert's work context (task and objective). Several experts can have the same focus: for example, in Figure 2, the focus is the security of a vehicle. According to this focus, we need to express the different viewpoints of different experts involved. Therefore we characterize the viewpoint by a personal dimension: the view angle.
- The view angle describes the characteristics of an expert or of a group of experts. It can describe the name of this expert (or of this group of experts), his/her application field (domain), his/her expertise level or skill, his/her experience in other domains interesting for the application, his/her role and place in his/her organization.

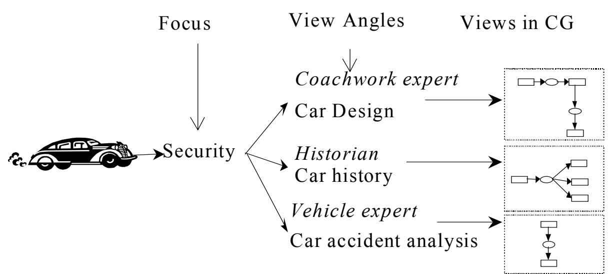

Figure 2: Example of multi-viewpoint description of a car

Viewpoints may index consensual and non-consensual knowledge:

- They may help in knowledge acquisition process by providing a support to represent non-consensual knowledge from several experts (i.e. express the "viewpoint" of each expert on the same object).
- They also enable to keep non-consensual viewpoints on a same object.
- In terms rof KB building, a viewpoint allows to index knowledge. In terms of access to the KB, it plays the role of a filter on the KB and helps the user to avoid to get lost in the whole KB by enabling access only to relevant knowledge according to the user profile. We distinguish two kinds of viewpoints:
- Viewpoints defining perspectives that index consensual descriptions of a same object by different experts. Those views are complementary and give a whole vision of the object. The object is supposed to be unique, but may have some characteristics interesting or visible only for some experts. Therefore a given expert will focus only on some perpectives on the object (the ones relevant for him). The models proposed in ROME (Carré et al, 1990), TROEPS (Marino et al, 1990), VBOOL (Marcaillou et al, 1993) for management of multiple viewpoints or in View Retriever (Acker and Porter, 1994) for extraction of viewpoints from a frame-based KB rely on the hypothesis that viewpoints are partial representations of a unique, coherent set of objects. So, these models handle perspectives.
- Viewpoints defining opinions that index non-consensual descriptions corresponding to the different, specific approaches of the experts. Such views are incomplete descriptions of the studied object and could be

collectivelyinconsistent. For example, such opinions are useful in case of design of an artefact on which several experts will have different (possibly contradictory) propositions.
- The first type of viewpoints "perspective viewpoints" and the second type "opinion viewpoints", in (Ribière, 1999; Ribière and Dieng-Kuntz, 2002).

2.4 Personalization and Community Support

Personalization and community involvement are areas of high potential contribution to the consensus modeling process. Specifically, personalization techniques can enable an individual to access the most relevant for her/him ontology items and instances in the most convenient manner. Thus, for the consensus achievement, personalization is important as it can serve as a mean to avoid consensus forming in certain cases: in particular, when in the course of personalization, substituting/excluding community activities and ontology items of higher potential benefit comparing to the ones an individual had to agree were discovered. Meanwhile, community support and awareness can enable an individual to gain maximum benefit from relevant ongoing community activities, ontology items and instances.

Personalization is traditionally defined as the ability to customize each individual user's experience of electronic content [McCarthy01]. The objective of personalization for the purpose of delivery of personalized information is fairly straightforward. It is to deliver information that is relevant to an individual or a group of individuals in the format and layout specified and in time intervals specified [Won02]. While personalization was applied extensively for individual users (especially in eCommerce area) [Aggarwal et al., 02; Instone04; Schiaffino and Amandi, 04], the problem of supporting communities with personalization-based information exchange on the Semantic Web context is still open.

By a community support, we understand delivery of certain objects by a community member that are reused or shared by the other community members and thus unite the community. The objects delivered by an individual community member are the basis for information exchange in the community and the information exchanged itself, thus these objects may range from portal content to ontology mapping schemas.

3 Methodologies and Models for Knowledge Acquisition and Consensus Making

In this section, we give an overview of theories and methodologies for knowledge acquisition and modeling of process of consensus that serve as a starting point for elaboration of the solution for knowledge acquisition and modeling of process of consensus in KnowledgeWeb. We describe substantial works of the state-of-the-art in different research areas of the Semantic Web, such as knowledge acquisition and consensus in human-Semantic Web interaction, upgrading the current natural language

Web to the future knowledge web, agent negotiation processes for forming a consensus on the Semantic Web, application of socio-economic theories for the process of consensus formation.

3.1 Knowledge Acquisition from Human Contributors and Communities

A very common way to acquire knowledge and information from human contributors and communities is via interaction with the Web applications that enable/require users to input information. Therefore, Semantic Web portals are of high potential importance for knowledge acquisition with respect to being natural platforms for acquisition of complex knowledge from human contributors and communalities. In this subsection, we overview the typical kind of information currently acquired by Semantic Web portals, which is mainly restricted to acquiring ontology instances, according to our observation.

To get an overview about how the state-of-the-art procedure looks like concerning ontology instantiation, a number of typical knowledge environments was reviewed. In the following section, it is shown how knowledge is acquired at the existing knowledge community environments, and how the data are created and maintained.
We review five semantic web portals: the Esperonto project, Knowledge Web portal, Mondeca environment, OntoWeb and K42. We chose all these project portals and Mondeca's platform as they are substantially advanced and typical knowledge-based semantic portals.

The description of each portal was divided in five subsections: *user roles, validators, creating instances, editing instances and deleting instances.* In the section about user roles information about the various user roles in a portal and their rights can be found. The section about validators explains which form of quality control is implemented in the portal. The summary is given in Table 1.

Esperonto Portal
User Roles
In the Esperonto [http://esperonto.semanticweb.org/] portal users have different permissions, which means that various user roles exist, such as Administrators, guest users and members. While Administrators are allowed to create, edit and browse any information on the portal, guest user can only browse the public information. Members have access to various areas. Basically, they are allowed to create and edit information items.
Validators
Esperonto is the only portal that does not have any form of quality control: every user can add anything to the portal. There are no validators. Administrators can delete inappropriate items, but basically anything can be published.
Creating instances
To create a new item the user has to select the type of information (ontology concept) the user is about to create. After having selected an ontology concept, the user is presented with a form to enter the name of the instance and its description. Immediately a new

instance is created and thus assigned to the knowledgebase. The name the user enters when creating the information item also works as an identifier, which may cause problems, because the user does not necessarily know all the names of all the instances in the portal.

Editing instances

Depending on the user right of an instance, existing instances can be edited later, even though the name, which works as an identifier, and description of the instance cannot be changed.

Deleting instances

Furthermore, as a member it is not possible to delete instances, even if they belong to the member who wants to delete the instances. Deletion is restricted to administrators.

Knowledgeweb Portal

The Knowledgeweb Portal [knowledgeweb.semanticweb.org] is based on WebODE as Esperonto portal . Knowledge Web Portal is more recent and more advanced than the Esperonto project.

User Roles

While Administrators are allowed to create, edit and browse any information on the portal, guest user can only browse the public information.

Members have access to various areas. Basically, members are allowed to create and edit information items.

Validators

There are no validators. Administrators can delete inappropriate items, but basically anything can be published.

Creating instances

To create a new item the user has to select the type of information (ontology concept) the user is about to create. After having selected an ontology concept, the user is presented with a form to enter the name of the instance and its description. Immediately a new instance is created and thus implicitly assigned to the ontology. The name the user enters when creating the information item also works as an identifier, which may cause problems, because the user does not necessarily know all the names of all the instances in the portal.

Editing instances

Depending on the user right of an instance, existing instances can be edited later, even though the name, which works as an identifier, and description of the instance cannot be changed.

Deleting instances

Furthermore, as a member it is not possible to delete instances, even if they belong to the member who wants to delete the instances. Deletion is restricted to administrators.

Mondeca

User Roles

Mondeca [http://mondeca.com] implements three kinds of users: administrators, validators and members. Administrators have unlimited rights to write, read and delete in the portal. Validators are responsible for quality control on the portal. They decide whether an item can be published or not. Each member is assigned a certain workspace and can create instances.

Validators

Validators decide whether an item of information can be published or not. Basically, there are two states an information item can have: “proposal” or “validated”, which follows the same line as many other Semantic Web portals except Esperonto do – an instance of a concept has to be validated before it can be published on the portal. The drawback here is, that a validator does not get a notification about a new information item waiting to be validated.

Creating instances

In the Mondeca portal instances of ontology concepts can be added in various ways: it can be added by the end user via forms, automatically using linguistic tools or by extracting information from structured or semi-structured sources. Initial filling of the knowledgebase with data can be done by importing files in OWL and in eXtended Topic Maps format.

The user can create instances by using a form. The type of form depends on the concept that the user wants to create, who provides the necessary information to create the ontology instance. The information item is implicitly assigned to ontology. Additionally, it can be assigned to one or more keywords in a defined thesaurus. Furthermore, it is possible to establish relations among items of information.

Editing instances

When editing an instance, values and semantic associations have to be changed separately. A big drawback is that there is no user interfaces to perform this functionality, so external tools, such as Protégé have to be used.

Deleting instances

Furthermore, the user can determine an expiration date to define for how long the information should be on the portal. This is also how deletion of items works. Moreover, one can also delete an item by using the user interface - depending on the right a user has. When an information item is deleted, all the semantic associations are deleted, too.

OntoWeb

User Roles

In this portals, there are two different user roles: "normal" users and administrators. While administrators can publish, reject, retract, delete and change all information items, normal user can only edit their own information items.

Validators

In the OntoWeb portal there are validators, the administrators, who are responsible for the quality. In the OntoWeb portal there is a distinction between private and public information items. Private items are only visible to the creator and the administrator, while the public ones can be viewed by every user once it has been reviewed by a portal reviewer.

Creating instances

The OntoWeb portal [http://www.ontoweb.org]differentiates between several types of information items defined in the ontology. For each of these types there is a form, which is used to create new instances. This form is divided in three sections of information: a

section called "base data" containing a short description; furthermore another section containing the
values for the attributes of the ontology concept, which includes an upload functionality; finally, a section called "meta data" containing instructions on how the portal should handle the new item. Public items also include an expiration date until which it will be accessible. When a user creates a new information item, it is automatically immediately assigned as an instance of the ontology.

Editing instances

Once created, each information item has its individual editing options. Users who have the appropriate rights can edit instances to a certain extend: as soon as an item has been published or submitted, the user can only rename the item or change its status, but not change its declarative description.

Deleting instances

Users are allowed to delete their instances independent from the status of the information items.

K42

User Roles

In the K42 portal [http://www.research.ibm.com/k42/]users are only allowed to browse through existing information while the administrator is responsible for creation, publication and maintenance, which means that all information items are provided and maintained by the portal administrator.

Validators

The Administrator who is responsible for everything on the portal is also performing quality control.

Creating instances

A tool named WebAuthor allows web-based editing of topic maps. A tool named Ontogen
can be used to create and edit information items based on web forms. Still, it is not possible to upload documents associated with the instance. This has to be done
separately. The information items are published immediately which means that they are published when they are submitted to the system via the creation opportunities. So the border and between the creation and the publication phase is intangible.

Editing instances

The administrator is able to edit all the information items on the portal.

Deleting instances

The administrator is able to delete all the information items on the portal.

	Esperonto	KnowledgeWeb	Mondeca	OntoWeb	K42
User Roles	Administrator, User, Guest	Administrator, user, guest	Administrator, validators, users	Administrator, users	Administrator, users
Validators	no	no	Validators	Administrator	Administrator
Creating instances	forms	forms	forms, wrapping possible	forms	forms, only administrator
Editing instances	possible. Name as an identifier must not be changed	possible	possible, but there is no user interface; tools must be used (e.g. Protege)	possible. Declarative description must not be changed.	only possible for the administrator
Deleting instances	only administrator	only administrator	depending on the expiration date	possible for users and administrator	only possible for the administrator

Table 1: Ontology Instantiation at the Current Knowledge Portals

Investigation of existing typical Semantic Web Portals reveals that there are various approaches to knowledge acquisition.

Interoperation between knowledge portals (e.g., using ontology alignment techniques) currently does not have a due implementation support. Application of these techniques is an opportunity to consider when the ontologies are populated and has been used for a longer period of time, and the techniques themselves become more mature. Although, ontology mapping support is important for gaining and sharing instances with other systems and portals.

Using forms is an appealing possibility. It is a very straightforward approach and is most often chosen solution for maintaining the portal and updating instances. It is an easy to use technology, so that users do not need expertise. The big drawback is that it is a very static approach. The instances have to be maintained manually and also the creation process is quite time-consuming.

The alternative is to consider using wrappers to gather information dynamically from the static web pages. Setting up a procedure using screen scrapers can be complex but in the end can pay off, if the information is well-structured, and the volume of the information is large. How time consuming and complex the implementation of both, wrappers and forms, is depends on the complexity of acquired information.

3.2 Knowledge Acquisition from Natural Language Sources

While conventional KA, where knowledge is derived from humans, usually consists of interviews and/or protocol analysis [Potter, 2001], these methods have theoretical limitations in that they assume that the required knowledge can be expressed in this form (natural language), and practical limitations in the amount of time and resources necessary not only for the interviews but also for transcription, subsequent analysis etc.

Semi-automatic and fully automatic techniques for KA from Natural Language attempt to overcome these problems by speeding up the process and eliminating the subjective nature of the knowledge elicited through human interaction. The current state of Natural Language Engineering technology has meant that the greatest benefit to this approach comes from the time saved rather than the improved quality of the knowledge elicited, especially since human intervention is largely still required at some level.

3.2.1 Top-down vs bottom-up approaches

Top-down (TD) approaches basically take an existing framework and populate the ontology, while bottom-up (BU) approaches, on the other hand, start from the text and cluster instances in order to form an ontology from scratch. BU approaches make integration difficult and mean that the ontology created is very specific to that kind of text. TD approaches, on the other hand, are more generic and may be more useful in practical terms, but have the disadvantage that they may be more difficult to create in the first place. Most research to date that uses NLP techniques has been of the bottom-up variety, though the increasing use of statistical techniques, machine learning and data mining approaches has led to the development of some promising TD approaches.

3.2.2 Text Analysis approaches

Text analysis approaches use techniques from Natural Language Processing (NLP) to identify important domain terms and/or concepts and possibly also relations between them. Approaches are generally quite simple and may take the form of frequency information about nouns, adjectives and verbs in the text, for instance. Such methods require some manual intervention to verify the validity of suggested items, and do not attempt to classify the found items, again relying on human intervention for this.

3.2.3 Pattern matching approaches

A more complex version of text analysis approaches involves the use of pattern matching and/or templates. These typically make use of lexical-to-semantic or syntax-to-semantic rules, based on the ideas of Harris' distributional hypothesis [Harris,1968] and later Hirschman et al. and Sager's information formatting [Sager, 1972], [Hirschman et al, 1975]. The idea here is that syntactic patterns from domain-specific texts can be used to determine sets of sublanguage-specific word classes. By examining sets of lexical items found in specific syntactic environments, semantic word classes can be established for the domain. For example, Hirschman et al. and Sager collected instances of the lexical items found as objects of the verb "develop" together with the subject "patient", and were able to develop from these a class "sign or symptom" consisting of lexical items such as "mild cold", "fever", slight cough", etc.

There has been much debate as to whether semantic information can be obtained more easily and reliably from lexical or syntactic patterns (see e.g. [Montemagni and Vanderwende, 1993; Davidson et al., 1998; Grefenstette, 1994]). In general, syntactic patterns produce more coverage, while lexical patterns provide more precision but only within a limited range, because it is hard to define all the possible lexical patterns.

3.2.4 Contextual approaches

Pattern-matching approaches enable us to create contextual clusters [Maynard and Ananiadou,1999], which are useful for various forms of KA. Principally, contextual clustering helps with the sparse data problem and provides us with more information than we would have by looking at individual words or phrases. For example, if we have a context containing a term about which we have no information, it is not very useful in itself. But if we can show that this context is similar to a set of other contexts about which we do have knowledge, then this information can be extended to the context in question. Let us take an example. Suppose we have the following:

- a term T: proliferative retinopathy
- a set of contexts S: [lower third of cornea, lower half of iris, top half of retina]
- a new context C: lower section of cornea

Assuming that we know nothing about C except that the word section is similar to the words third and half. Given this, however, we can predict that C is similar to the other contexts in S and should form part of the set S. Armed with this knowledge, we can now predict that the relationship between T and C is similar to the relationship between T and any of the members of S.

3.2.5 Information extraction approaches

Traditional information extraction (IE) systems filter documents in order to extract relevant pieces of information, for example, names of people, companies, dates, times etc. They may be tailored to different domains and to extract different kinds of information depending on the requirements. They may also find relations between entities in the text. Typical state-of-the-art IE systems require a specialized lexicon of terms not found in general-purpose dictionaries; domain-specific word or concept classes for semantic generalization; and syntactic-semantic patterns for locating facts or events in text, among other kinds of knowledge

Approaches to traditional IE generally take one of 2 forms: rule-based, knowledge engineering approaches or statistical approaches using machine learning. The former require little training data, making use of human intuition, but may be time consuming to develop and adapt to new domains. The latter require large amounts of training data, which may also be time-consuming to create, and reannotation is required for new domains; however, they have the advantage that human language experts are not required.

Typically the output of an IE system is in the form of annotated data or a database of relevant information. This may then need to be manipulated manually into the required format. IE can clearly be seen as a form of Knowledge Acquisition; its main strength being the speed and efficiency of extraction compared to humans (particularly when small pieces of information are buried in large documents). However, quality of extraction is quite varied, depending on the task and domain, so manual intervention may still be required to verify the results.

3.2.6 Ontology-based approaches

Unlike text-based approaches, these make use of ontological knowledge of the background domain in order to aid the elicitation and organisation of information. The use of an ontology enables the acquisition and representation of semantic information Typically these use an ontology in combination with either pattern-matching approaches, e.g. [Hahn and Schnattinger, 98] or with traditional IE approaches. Unlike the previous methods, ontology-based approaches work on the top-down principle, using an existing ontology to identify relevant new instances in the text and to add these to the ontology. The pattern-matching method uses inheritance information about concepts in the ontology to recognise new instances in the text of existing concepts. The ontology-based IE method tries to find similar instances in the text to those already in the ontology, using mechanisms such as coreference to find variations of existing instances, and to populate the ontology with these once identified.

3.2.7 Machine Learning Approaches

Nowadays, the volume of electronically stored information continues to expand across computer networks. The increasing amount of collected data in organizations and the world wide web is demanding for methodologies to automatically, or semi-automatically, extract hidden, previously unknown, and potentially useful information out of it [Wrobel et al., 2003]. Typical KA approaches need the assistance of a user who supervises the process and evaluates its results. These tasks are sometimes too laborious and time consuming as the amount of data to process increases. This situation involves the need of intelligent access to collections of web documents and information stored in databases. For this purpose, ML offers a set of techniques, tools and systems that can help to develop techniques and principles for automating acquisition of knowledge [Mitchell, 1997] and to solve effectively related problems, such as [Karakoulas and Semeraro, 1998] semantic indexing, content-based search, knowledge acquisition from experts and information sources, document classification, semantic querying, integration of ontologies or knowledge bases into Internet search technologies, etc. Some ML methods can help the KE reduce the time and the cost of developing knowledge-based software by extracting knowledge directly from existing databases and textual repositories [Webb, 2002]. Other ML methods enable software systems to improve their performance over time with minimal user intervention [Palous, 2002].

ML research has bred a number of automated techniques for knowledge capturing and revision [Wrobel et al., 2003]. During the last years, researches on KA have looked for integrative approaches that exploit synergies between traditional KA approaches combined with ML techniques in order to discover, capture, represent, store, retrieve and reuse knowledge. A clear example in the context of the Semantic Web is the research on the ontology learning field that is devoted to reduce the time and effort needed to build up an ontology from different information sources [Maedche , 2002]. The application of several ML algorithms for automating the acquisition of new knowledge to enrich an ontology has proved to be successful.

Data mining (or Knowledge Discovery in Databases) research is perhaps one of the most widely known demonstration of the application of ML techniques in various domains, such as industrial settings, financial prediction, medical diagnosis, administration, commerce, etc. DM aims to discover patterns, information and knowledge in large, complex data sets and it is defined as the nontrivial extraction of implicit, previously unknown, and potentially useful information from data [Frawley et al., 1992]. It uses machine learning, statistical and visualization techniques to discover and present knowledge in a form which is easily comprehensible to humans [Pinto et al., 2001]. DM can help, for example, companies and institutions to focus on the most important information in their data warehouses.

Similar techniques have been also satisfactorily applied to text analysis and to extract previously unknown knowledge from text. This field is known by Text Mining (TM). TM is [Hearst, 1997] an interdisciplinary field involving information retrieval, text understanding, information extraction, clustering, categorization, visualization, database technology, machine learning, and data mining. Information can be extracted to derive summaries for the words contained in the documents or to compute summaries for the documents based on the words contained in them. Hence, words, clusters of words used in documents, etc. can be analyzed, or a set of documents could be analyzed to determine similarities between them, creating cluster of them with similar content.

TM approaches can then be incorporated in other analyses such as unsupervised learning methods (i.e. clustering) [Fayyad et al., 1996; Simoudis, 1996]. TM has a close relationship with NLP systems [Fakotakis and Sgarbas, 1998] , specially in those problems related to syntactical and semantically analysis (section 3.2.2), and information extraction (section 3.2.5). NLP systems provide an analysis of the texts contents, with variations depending on the tool that is used. TM approaches aim to transform these analyses into usable data. A general TM approach [Manning and Schütze, 1999] computes all words found in the input documents and counts them in order to build a matrix of documents and word frequencies (number of times that each word occurs in each document or other similar measures). This basic process can be further refined to exclude certain common words and to fuse different grammatical forms of the same words. Once this information has been totally collected, statistical and ML techniques can be applied to build clusters of words or to identify relevant words from the text.

Part of the most usually applied ML methods [Mitchell, 1997] that can be used for KA are : decision tree learning [Quinlan, 1996] that are mostly used for variety of classification tasks; neural networks methods [Bishop, 1995] that have been successfully applied to problems such as speech recognition, bayesian methods [Berry, 1996] provide a probabilistic approach to inference and it is widely used to classify text documents, being one of the most effective classifiers; inductive logic programming [Flach, 1998] has been used for creating logical programs from training data set, being mostly used for this purpose in DM for finding rules in huge databases; case-based reasoning [Aamodt and Plaza, 1994]; support vector machines [Boser et al., 1992] has also become very useful method for classification, etc. This list of methods is far to be exhaustive, since the applicability of a method rely on the type of problem, the structure of the input data, the possibility of training the system, etc. These methods have been applied in several DM problems with successful results, extracting relevant patterns from large databases [Pinto et al., 2001].

Another important role played by the ML techniques is constituted by the classifiers . The construction of a classification procedure [Michie et al., 1994] from a set of data for which some classes are known has also been variously termed pattern recognition, discrimination, or supervised learning (in order to distinguish it from unsupervised learning or clustering in which the classes are inferred from the data).

3.2.8 Evaluation

Evaluation of knowledge acquisition techniques is an extremely complex issue, partly because it is not clear how much knowledge should be acquired in the first place, and partly because there is not necessarily one correct solution to the organisation of the knowledge obtained. Comparing the performance of different KA systems is also not straightforward since they are typically tuned to specific domains and applications, and use different texts as their starting point.

According to [Hirschmann98], there are three kinds of evaluation for language engineering systems:

- adequacy: determining the suitability of a system for some specific purpose;
- diagnostic: determining the performance of the system with respect to some clearly defined dimension of the possible set of inputs;
- performance: determining the performance of the system for comparative purposes.

These measures are not completely independent, but it may be possible to score well on one type and badly on another. For example, a system's performance may be good, but it may not be suitable for a particular task (for example, because it is difficult to use by a non-expert, or because it is slow), and likewise the system that is the most suitable for the task may not be the one with the highest performance.

The demands of research have generally placed the focus on performance evaluation, and in particular on comparative performance evaluation, prompted largely by organizations such as DARPA who have been running a series of competitions (e.g. MUC [ARPA93], [Chinchor92], and more recently ACE [ACE04]) to promote research in certain areas and focus development of working systems. While this has had a number of benefits, such as the standardisation of test sets, evaluation metrics and formats, it has also led to stagnation of research in areas not evaluated, and has overshadowed the development and use of adequacy evaluation and system portability.

3.3 Knowledge acquisition process for multimedia - learning and ontology-based approaches

Owing to the recent progress of computing and telecommunication technologies, multimedia has become a major source of content. A wide range of applications such as content production and distribution, telemedicine, digital libraries, distance learning, GIS etc. are expected to use general purpose multimedia database systems. Such systems present great challenges in terms of efficient storing, querying, indexing and retrieval. Although there has been a significant progress evident on automatic segmentation, scene-change detection and recognition and extraction of low-level features for multimedia content, the possible different interpretations and intended uses of such content render these efforts insufficient, since they leave out the underlying semantics. Consequently, bridging the gap between human interpretation of multimedia content and the one derivable by means of a computer emerges as the key-enabling factor to support multimedia content target applications and users. This has resulted in a growing demand for efficient methods for extracting and representing knowledge of the semantics associated with such content in order to enable its management and exploitation. This particular need for semantic modelling of multimedia information is also evident in many emerging application fields including among others semantic transcoding, filtering, personalization and summarization.

Generally, the knowledge associated with a multimedia document regards two kinds of information: structural and semantic respectively. The former describes content in the form of signal segments and their properties, comprising thus a low-level and machine oriented description. Example descriptions of the former would be “an image with a large round green textured object” or “a video segment containing spoken content”. On the other hand, semantic information provides a high-level, human oriented kind of knowledge, allowing the description of different aspects of multimedia content at different levels of abstraction, varying from objects identification (e.g. a car, a horse) to events recognition (e.g. scoring a goal). Approaches relying only on low-level features manage to reveal some kind of similarity between multimedia data but lack the potential to convey their perceptual meaning to the user. As a consequence, primary emphasis is given on the development of methods to incorporate higher-level semantics in the process of capturing low-level multimedia features through appropriate intelligent mappings. This challenging problem caused by the gap between the information that can be extracted automatically from visual data and the interpretation that the same data has for a user in a given situation is what the content-based retrieval (CBR) community often refers to as the *semantic gap*.

Low-level features, such as pixel luminance, region's contour, motion activity and other, despite their inability to provide a human understandable representation of the underlying multimedia content, consist the basis of typical content-based retrieval. The corresponding techniques are based on the extraction of visual or audio properties and use distance scores between the query and each of the referenced multimedia documents to determine a set of matches. An automatic way for transition from the low level features to semantic entities or equivalently the automatic extraction of high-level characteristics is an extremely hard task [Chang02]. The latest efforts have been focused on the extraction of medium level features, such as automatic summarization and key frames extraction [Chang00], [Uchihashi99], [Christel02], [Calic02]. Similarly, automatic categorization of images in pre-defined classes, such as indoors/outdoors, city/landscape, faces/non faces can be achieved after a training phase [Luo01].

However, as already mentioned, approaches relying only on low-level multimedia features cannot provide the means for generating high-level, semantic representations of multimedia content. The consequent performance limitations have resulted towards an inclination to integrate semantics with audiovisual features into a unified framework, thus moving from low-level to high-level features allowing them to benefit from each other. Since in many real applications the effectiveness has been shown to greatly improve when a priori knowledge is included, several research methods on exploiting domain-specific knowledge for multimedia analysis have been undertaken. A rough categorization of the developed knowledge-assisted approaches yields two classes, as further described in the following: approaches that use stochastic methods and exploit automatic learning capabilities to derive knowledge and approaches that take advantage of explicitly defined domain-specific knowledge to drive the extraction of high-level semantic concepts based on automatically extracted audiovisual information, i.e. low-level features and spatio-temporal behaviour.

Both classes of knowledge-assisted multimedia analysis approaches consider a priori knowledge in two different contexts:

- the actual analysis process that aims at understanding the conveyed audio-visual information, i.e. the segmentation, localization and identification of visual salient objects, the automatic speech recognition etc.
- the acquisition of higher-level semantic information in terms of meaning, thereby enabling faster and easier information browsing, interpretation and deduction by the end user

As will be illustrated in the following the main focus of the relevant work in literature addresses mostly the generation of annotations regarding a set of recognized objects or simple events. It is important to stress that object detection in terms of extracting visual and/or audio segments that represent meaningful concepts as perceived by humans is the first step towards generating high-level semantic interpretations. Practically this means that exploiting the available knowledge an initial set of identification labels is produced and subsequent processing of this initial set of metadata enables the detection of higher level semantic concepts, such as complex objects and events, e.g. an island, a submarine scene or sailing activity in the domain of beach vacations. In the following subsections, a brief description of the recent state of the art is given for both categories.

3.3.1 Knowledge acquisition in multimedia: learning approaches

Machine learning refers to a broad class of probabilistic and statistical methods for estimating dependencies between data in order to perform tasks associated with artificial intelligence such as recognition, diagnosis, planning, prediction etc. The main characteristic of learning-based approaches is their ability to adjust their internal structure according to input and desired output data pairs in order to approximate the relations implicit in the provided (training) data. Consequently, machine-learning approaches constitute an appropriate solution when the considered a-priori knowledge cannot be explicitly defined because it is ill-defined, incomplete or too large in terms of amount to be efficiently represented. Many techniques have been developed for realizing the learning process and among the most widely used ones are neural networks, maximum likelihood estimation techniques, support vector machines, case-based reasoning approaches etc.

Taking into account the challenges posed by multimedia analysis in terms of extracting semantic descriptions (recognition of concepts corresponding to meaningful objects and events), the class of machine learning approaches appears to be a particularly appealing solution. The reason is that because of their aforementioned characteristics, learning systems are able to tackle elegantly the intrinsic features of visual information, i.e. subjectivity, ambiguity and complexity. This becomes particularly evident in cases where multimedia analysis addresses semantic concepts whose relevant visual (audio) structure might be too complicated to be explicitly expressed in a manually constructed model. It is this particular property of audiovisual information that proves beneficial the use of hybrid approaches where machine-learning and (explicitly defined) rule-based methods provide a unified framework for acquiring semantic descriptions of multimedia data acting complementary.

Some examples of knowledge acquisition applications based on machine-learning techniques address tasks like face detection and human recognition. In [Sigal00], reliable skin segmentation despite wide variation in illumination during tracking is achieved. In [Kouzani03], an intelligent system that locates human faces within images using neuro-fuzzy networks is presented, while in [Jones98] the construction of a skin pixel detector is described. 3D human figures tracking in monocular image sequences is performed in [Sidenbladh00]. Machine learning approaches have also been widely used for the knowledge acquisition process in the sports domain. More specifically, in [Pingali00] tennis broadcasts are enhanced with ball tracking and some impressive virtual replays, while in [Kijak03] HMMs are used for structure analysis of tennis videos using visual and audio cues. Event detection and summarization from snooker broadcasts is presented in [Renman03]. As for soccer, a fully automatic framework for analysis and summarization is presented in [Ekin03] and does not require strictly the use of object-based features, but can be efficient using only cinematic features. HMMs are also used in [Xie04] to analyse the structure of soccer programs, and more specifically to detect play and break segments while in [Assfalg02] HMMs are applied for the detection of soccer highlights.

Support Vector Machines [Burges98], [Vapnik98], [Haykin99], a quite new method in pattern recognition in the Neural Networks area have received lot of attention due to their ability to solve classification problems, that are non-separable by a hyperplane in the input space, using a higher-dimension feature space transformation. It must be noted that in the context of multimedia analysis classification refers to the identification of specific semantic labels in the examined content. Such concepts might represent objects such as street, tree, roof etc and events such as explosion, interview etc. Some interesting SVM applications include [Chapelle99], where they are used for histogram-based image classification, [Wang03], where an SVM framework is used for robust semantic labelling of image regions, while in [Snoek2003] SVMs are employed for the detection of concepts such as goal, yellow card, substitution in the soccer domain. Another learning method considers neuro-fuzzy networks, which encode structured, empirical (heuristic) or linguistic knowledge in a similar numerical framework in contrast to neural networks that encode sampled information in a parallel-distributed framework. Although they can describe the operation of the system in natural language with the aid of human-like if-then rules, they do not provide the highly desired characteristics of learning and adaptation. The use of neural networks in order to realize the key concepts of a fuzzy logic system enriches the system with the ability of learning and improves the sub-symbolic to symbolic mapping as illustrated in [Kosko92], [Lin95], [Stamou01], [Tzouvaras03]. In [Wallace03], fuzzy ontological relations and context-aware fuzzy hierarchical clustering are employed to interpret multimedia content for the purpose of automatic thematic categorization of multimedia documents.

3.3.2 Knowledge acquisition in multimedia: knowledge-based approaches

As previously mentioned, the use of explicitly defined a priori knowledge can facilitate the extraction of higher-level semantics in case of videos of well-structured domains [Yoshitaka et al., 99], [Al-Khatib99]. The approaches examined in this subsection tackle the problem of bridging the gap between low-level descriptions and high-level interpretations by exploiting a priori domain knowledge in the form of explicitly defined object (event) models and manually or semi-automatically constructed rules. These rules drive the reasoning process on the embodied in the multimedia data content aiming both at the detection of valid semantic descriptions as well as the generation of higher-level descriptions based on the ones already acquired. Thus, in contrast with the dynamic learning approaches presented in the previous subsection, the current ones are mostly based on static knowledge structures in the form of predefined models and rules.

A key issue of model-based approaches is knowledge representation. Towards the direction of enabling the automatic generation and understanding of audiovisual descriptions for retrieval and filtering purposes, the Moving Pictures Expert Group (MPEG) has developed the Multimedia Content Description Interface (MPEG-7), which aims to define a rich set of standardized tools advancing among others applications interoperability. On the other hand, the World Wide Web Consortium (W3C) has developed important initiatives within the Semantic Web context such as the RDF (Resource Description Framework) and the OWL (Web Ontology Language. In the majority, the developed approaches have been highly influenced by the knowledge representation community and the recent advances in the context of the Semantic Web, since the use of such semantic descriptions enables more sophisticated semantic querying and ensures that services, agents and applications on the Web have a greater chance of discovering and exploiting the provided multimedia information. In the following, we briefly present a subset of representative efforts undertaken within this framework. As will be illustrated, some of the model-based knowledge-assisted approaches do not follow a formal logic-based data model, but rather develop internal knowledge representation and management techniques.

A priori knowledge models have been used as a knowledge base that assists semantic-based classification and clustering in [Yoshitaka and Ichikawa, 94], [Mezaris04a], [Kompatsiaris04b]. In [Tsechpenakis02], semantic entities, in the context of the MPEG-7 standard, are used for knowledge assisted video analysis and object detection, thus allowing for semantic-level indexing. In [Chen01], [Chan02] hybrid methods extending the query-by example strategy are developed, while in [Benitez00], MediaNet, a knowledge representation framework that uses multimedia content for representing semantic and perceptual information is presented. In [Naphade02] the problem of bridging the gap between low-level representation and high-level semantics is formulated as a probabilistic pattern recognition problem. In [Meghni97], the problem of injecting semantics into multimedia data is addressed by introducing a logic-based (a description logic based) data model for describing both the form and content of multimedia documents, while in [Petkovic03], inferencing semantics automatically from raw video data is addressed by introducing a layered video data model where object and event grammars formalize the descriptions of high-level concepts and facilitate their extraction based on features and spatio-temporal reasoning.

Among the possible knowledge representations ontologies provide a set of particularly appealing properties. On one hand, they provide the formal framework required for consensus and information sharing, thus enforcing interoperability and communication of knowledge. On the other hand, they provide the means to support inference and derivation of implicit knowledge from the already existing one. Ontology modelling and ontology-based metadata creation has addressed mainly textual resources [Schnurr00] for the past decades, while in multimedia, ontologies have been mostly used in the form of thesauri-aided approaches for photo annotation [Schreiber01], [Hyvönen02], [Hollink03], [Luo04], audio structuring and retrieval [Khan00] and image organization, browsing and retrieval [Tansley98], [Yang01] among others. However, acknowledging the importance of coupling domain-specific and low-level description vocabularies for analysis purposes has recently set focus on using ontologies to drive the extraction of semantic descriptions instead of only providing a formal structure for annotations.

In [Kompatsiaris04a], an object ontology, coupled with a relevance feedback mechanism to improve precision, is introduced to facilitate the mapping of low-level to high-level features and allow the definition of relationships between pieces of multimedia information. In [Hunter01] an RDFS ontology for expressing MPEG-7 metadata terms is described, in order to make MPEG-7 accessible, re-usable and interoperable with other domains, while in [Tsinaraki04] a methodology for enabling interoperability of OWL domain-specific ontologies with the complete MPEG-7 MDS is described. In [Jaimes03] a framework for video content understanding that uses rules from knowledge bases and multimedia ontologies is presented, and in [Troncy03], formal descriptions of video content are provided based on languages and technologies underlying the Semantic Web and in particular ontologies. In [Town03] symbolic terms are related to visual information by utilizing syntactic and semantic structure in a manner related to approaches in speech and language, and analysis takes place within the ontological domain defined by the structure of the problem and the corresponding goal set. In the MUMIS project [Reidsma03], ontology based information extraction is applied to improve the results of information retrieval in multimedia archives, making use of a domain specific ontology, multilingual lexicons and reasoning algorithms. In the FUSION project [Hunter04], a user-assisted approach to generate ontology-based semantic descriptions of images from low-level automatically extracted features is presented, where Semantic Web technologies and image analysis techniques are combined to develop a knowledge management system aiming at the optimization of designing fuel cells.

To conclude, independently of the followed approach, machine-learning or model-based, the process of acquiring knowledge from multimedia content has as first step the recognition of a set of simple yet representative objects and/or events of the examined domain. The next step comprises the inference of more complex semantic descriptions that incorporate visual context as well and the extraction of information implicitly contained in the content. The integration of ontologies in the multimedia analysis process, apart from the already mentioned advantages in terms of interoperability and formal foundations, has the additional benefit of enabling intelligent search and retrieval and consequently provides support for applications such as personalization, filtering, etc.

3.4 Modeling of the Process of Consensus Between Individuals and Communities

In this subsection, we discuss the consensus process modeling that involves knowledge acquisition from individuals, user groups and generalization at the community level.

3.4.1 Personalization and Community Support Approaches

We outline two directions in consensus making between individuals and communities: personalization and community support. The application fields of these techniques include recommender systems and efficient information delivery for knowledge acquisition in consensus making process.

Personalization is a field that aims at making applications more useful. Personalization is traditionally defined as the ability to customize each individual user's experience of electronic content [McCarthy01]. The known areas of personalization application are:

- handling different sources of content
- arrangement of content on a screen
- delivery mechanisms ("push" vs. "pull")
- delivery vehicles (web browser, mobile phone, etc.).

The objective of personalization for the purpose of delivery of personalized information is fairly straightforward. It is to deliver information that is relevant to an individual or a group of individuals in the format and layout specified and in time intervals specified [Won02]. While personalization was applied extensively on the ordinary Web portals for the individual users (especially in eCommerce area) [Aggarwal et al., 02; Instone04; Schiaffino and Amandi, 04], the studies for community and consensus aspects of personalization in the Semantic Web context are still lacking to a large extent. Previously existing techniques such as collaborative-filtering for recommender systems are explored to be applied in the Semantic Web context [Agarwal et al., 03; Konstan and Riedl, 03].

Theoretical works supported by implementation were done in the Semantic Web personalization fields. A large attention was paid to the notion of *context*, whether time context, delivery context or other context. Specifically, attention was paid to specification of user profiles: a person is the most often modeled object in the currently available ontologies[1]. Theoretical and practical studies around user profiles include their segmentation into long-term, specific medium term and short term user profiles [Agarwal et al., 03], making personalized semantic bookmarks to produce personal views corresponding to personal preferences and profiles [Maedche et al., 03].

The community issues are currently usually studied with respect to computing communities by means of clustering based techniques and identifying the communities to which pages belong [Greco et al., 04]. For the issues of individual personalization issues, there are no solution frameworks that support these issues extensively. The state of the art is mainly in establishing the theoretical basics for the further work on the application level, e.g., developing languages, such as a view language that picks up the unique situation of data in the Semantic Web and allows easy selection, customization and integration of Semantic Web content [Volz et al., 03]. Semantic log files to track usage patterns and identify answers to questions such as if a single authenticated user has a special interest in a certain part of the ontology, if there are user groups, etc [Maedche et al., 03].

[1] For a comprehensive repository of currently developed schemas, check http://www.schemaweb.info

3.4.2 Model for the Cooperative Building of an Ontology: C-VISTA

Several methods have been proposed for knowledge acquisition from multiple experts and in particular, for building a knowledge base from multiple experts through cooperation among the knowledge engineer and the experts. Inspiration can be taken from such methods for cooperative building of an ontology from multiple experts. Let us cite:

- methods for building cooperatively ontologies (Euzenat, 1996; Garcia, 1996; Tennison and Shadboldt, 1998),
- method for building terminological concept bases (Falquet and Mottaz Jiang, 2000)
- methods for integration of ontologies (Dieng and Hug, 1998a),
- method for comparison of conceptual graphs from several experts (Dieng and Hug, 1998b),
- collective elicitation protocol for knowledge acquisition from multiple experts (Dieng et al, 1998),
- an agent-based method for knowledge acquisition from multiple experts (Dieng, 1994; Dieng et al, 1996, 1998; Labidi, 1996).

In (Mueller and Dieng, 00), several types of conflict among human agents (in particular during knowledge acquisition phases) or among software agents in distributed artificial intelligence applications are analysed and techniques for detecting them and solving them are described.

The following sections will detail the C-VISTA model of viewpoint proposed in (Ribière, 1999; Ribière and Dieng-Kuntz, 2002) or cooperative building of an ontology organized in viewpoints.

C-VISTA models aims at enabling a knowledge engineer to: (1) Identify and index terminological differences between experts and establish a link between different terminologies; (2) Enable multi-representation of an object according to different experts.

3.4.2.1 Problems in Ontology Cooperative Building

Some past experiments of knowledge engineering (Labidi, 1996 ; Dieng et al, 1998) showed that experts working together with the supervision of a KE tend to integrate their vocabularies and to create a common vocabulary in order to understand one another. But the concepts underlying such terms may be actually used differently by the different experts at different levels of granularity and for different situations, so with divergent interpretations. Moreover, the KE does not model the context and the objective of use of each concept or term. This notion of context and objective on a concept must be normally deduced from the concept hierarchy organization (kind-of link). But when in a huge ontology, the whole context of interpretation is not modeled explicitly, it leads to misunderstandings (in particular when a part of the ontology must be reused for another application).

For example, in a hierarchy using only the "kind-of link " to provide the interpretation of a concept, the Accident_Factor concept could be subtyped by the Lack_of_road_exit_specific_signal concept (that is a kind of accident factor) and by the Indicator_ambiguity concept (that is also a kind of accident factor). Such accident factors are proposed by different experts. The first concept is based on the analysis of the infrastructure as accident factor and conforms to the Infrastructure viewpoint while the second one is based on the analysis of the driver as accident factor and conforms to the Driver viewpoint.
By the same way, in a hierarchy on Health-care-network, the Health-care-network concept could be subtyped by the Territorial-network concept (from a viewpoint “Geographical-area”), by the Pathology-centered-network concept (from the viewpoint “Goal”) and by the Purely-social-network (from the viewpoint “Activity”).
So, a viewpoint enables the explicit expression of a particular subtype relation existing between two concepts. Most of the methods described in related work build a consensual ontology, without such particular subtype links that provide documentation and track of the knowledge integration process. As a concept hierarchy can often be built using several different criteria, *our notion of viewpoint enables to make explicit the criteria underlying the subdivision of a concept into its subconcepts*.

3.4.2.2 The C-VISTA Model

The C-VISTA model was defined using the conceptual graph (CG) formalism (Sowa et al, 1984; Chein and Mugnier, 1992). This formalism enables to build a support S and a base of conceptual graphs. A support S is composed of a concept type hierarchy (noted Tc), an ordered set of relation types with their signatures, a set of markers or referents (M) and a conformity relation between concept types and markers. A CG is a bipartite graph built according to the support S and composed of two types of nodes: 1) concept-nodes, each labelled by a concept type and a referent, 2) relation-nodes, each labelled by a relation type. The support S corresponds to terminological knowledge (or ontology) while the base of CG corresponds to assertional knowledge.
But here, we will present the C-VISTA model independently of the CG formalism. We will rather rely on the terminology: concept (instead of concept type) and instance instead of concept).
The C-VISTA model enables to express viewpoints in the concept hierarchy, to describe the multi-representation of an object and to link different terminologies in a same concept hierarchy, thanks to the organization of the ontology in several viewpoints.

Expression of Multiple Viewpoints

Basic and v-oriented concepts
Let tc and tc' be two concepts. If tc' is a subtype of tc, then there may exist a viewpoint p such that tc' is a subtype of tc according to the viewpoint p. In that case, tc is called “basic concept” and tc' “viewpoint-oriented concep” (noted “v-oriented concept”).
For example, the v-oriented concept Highway is a subtype of the basic concept Infrastructure according to the Administrative viewpoint.

A given concept may have several immediate supertypes: a concept t' may be both a subtype of t1 w.r.t. the viewpoint p1 and a subtype of t2 according to the viewpoint p2. So, the concept hierarchy corresponds to a partial order but not to a rooted tree.
A given concept can be both basic (i.e. have v-oriented subtypes) and v-oriented (i.e. be itself a subtype according to a viewpoint). So the presence of basic concepts is not necessarily restricted to the higher levels of the concept hierarchy.

Criterion

A viewpoint is characterized by the explicit criteria according to which a v-oriented type will be considered as a subtype of its basic type.

Viewpoint template

A viewpoint template is composed of two sets of generic criteria, the first set characterizing the focus and the second one the view angle.
We will suppose that a given ontology relies on one viewpoint template: before developing the ontology, the KE and the experts must agree on a viewpoint template on which they will rely for organizing the ontology. The viewpoint template proposed in Figure 3 is composed of:

- The focus, characterized by the generic criteria Context and Objective,
- The view angle, characterized by the generic criteria Person, Field of skill, Expertise level, Other expertise fields and associated levels of expertise.

	Example of viewpoint template
Focus	Context Objective
View angle	Person Field of skill Expertise level Other expertise field/level of expertise

Figure 3: Example of viewpoint template

Viewpoint

A viewpoint p is defined on a viewpoint template Ptemplate by instantiating some (or all) of the generic criteria of this template.
Figure 4 gives an example of a particular viewpoint built according to the viewpoint template shown in Figure 3.

	Example of viewpoint
Focus	Context: Accident analysis Objective: Security of the crossroad
View angle	Person: Manuel Field of skill: Infrastructure Expertise level: 9 (Expert) Other expertise field / level of expertise: Vehicle dynamics / 5

Figure 4: Example of viewpoint

Viewpoint link

A viwpoint link (VPT:p) enables to express that a concept tc' is a subtype of another concept tc according to this viewpoint p.

C-VISTA model, summarized in Figure 5, allows to express for one given concept, a set of more specialized concepts provided by the experts. For the process of terminology integration, we need to express different kinds of viewpoint links in order to distinguish consensual and non-consensual knowledge in the ontology.

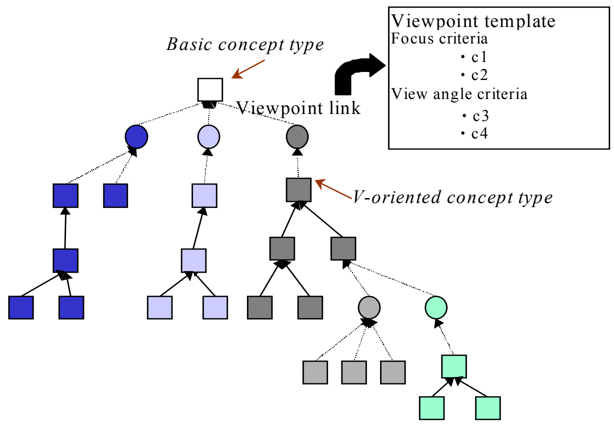

Figure 5: C-VISTA Model

Therefore the following links are introduced in (Ribière, 1999):
The *perspective link* noted (PERSP: p) can index, according to the viewpoint p, a concept having a consensual definition, i.e. shared by all the experts. The subtypes of such a concept are considered as having also consensual definitions.
The *opinion link* noted (OPINION: p) can index, according to the viewpoint p, a concept having a non-consensual definition. This concept stems from the opinion of an expert, not yet shared by the other experts. The subtypes of such a concept are considered as also having non-consensual definitions.

Figure 6 shows an example of application of C-VISTA model. It describes three viewpoint links and shows the different v-oriented subtypes of Accident_factor according to those three viewpoints. Each of those viewpoints is characterized by a particular instantiation of the same viewpoint template. So, the C-VISTA viewpoint model enables to organize the concept hierarchy into explicit viewpoints, making the resulting ontology more accurate and readable.

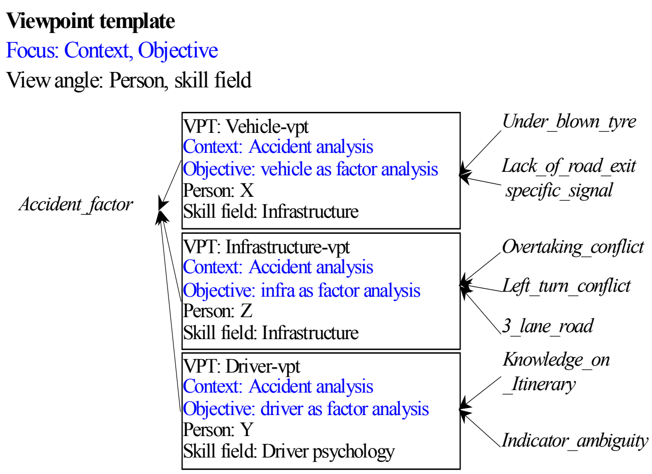

Figure 6: Example using C-VISTA model

Multi-representation of a Concept

Having multiple representations of an object allows, for a given instance, to obtain different perspectives describing this instance according to different viewpoints. For example, an infrastructure could be seen as a straight_road or a curved_road according to the curve viewpoint; it could also be seen as a highway, a national_road or a departemental_road according to the administrative viewpoint; and last, according to the nb_lanes viewpoint, it could be a 3_lane_road or a 2_lane_road or a One_way. The different possible viewpoints for an instance are represented in the concept hierarchy with the C-VISTA model.

By the same way, a specific healthcare network, DIABETO, can be considered as a territorial network according to the viewpoint “Geographical-area”, a town-hospital network according to the viewpoint “Administrative”, a pathology-centered network (cf. dedicated to diabetes) according to the viewpoint “Goals” and a mixed-network (both medical and social), according to the viewpoint “Activity”.

In CG formalism, the creation of an instance establishes a link between the original concept of the hierarchy and the name of the instance. This link is called instantiation link (it corresponds to the is_a link of object-oriented representations). The instantiation of a basic concept is called a basic instance and the instantiation of a v-oriented concept is called a v-oriented instance. We also introduce another link, called representation link (see model in Figure 7 and example in Figure 8), and inspired by ROME (Carré and Dekker, 1990).

Representation link

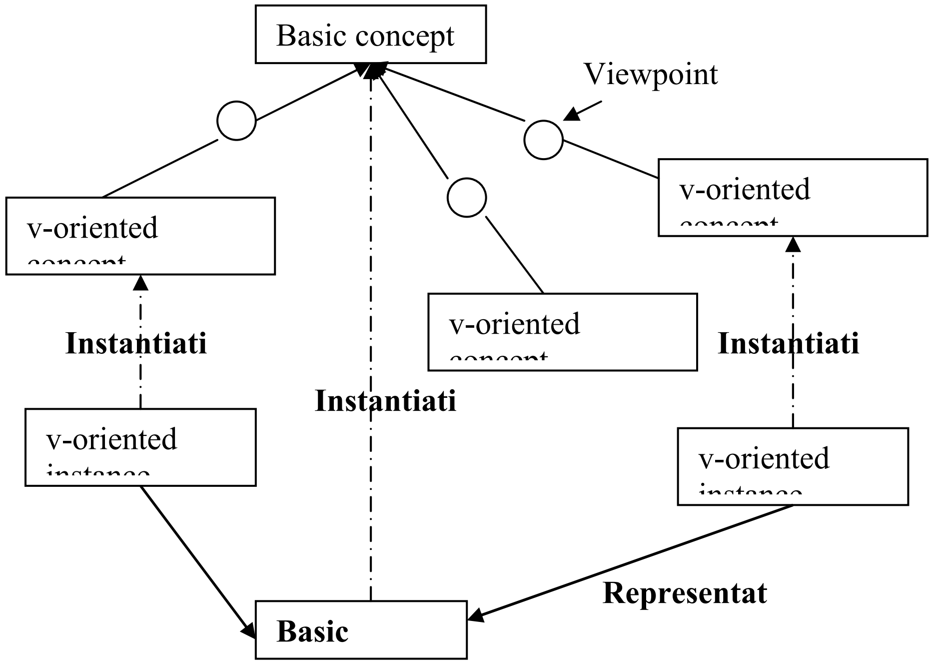

Figure 7: Model of multi-viewpoint representation

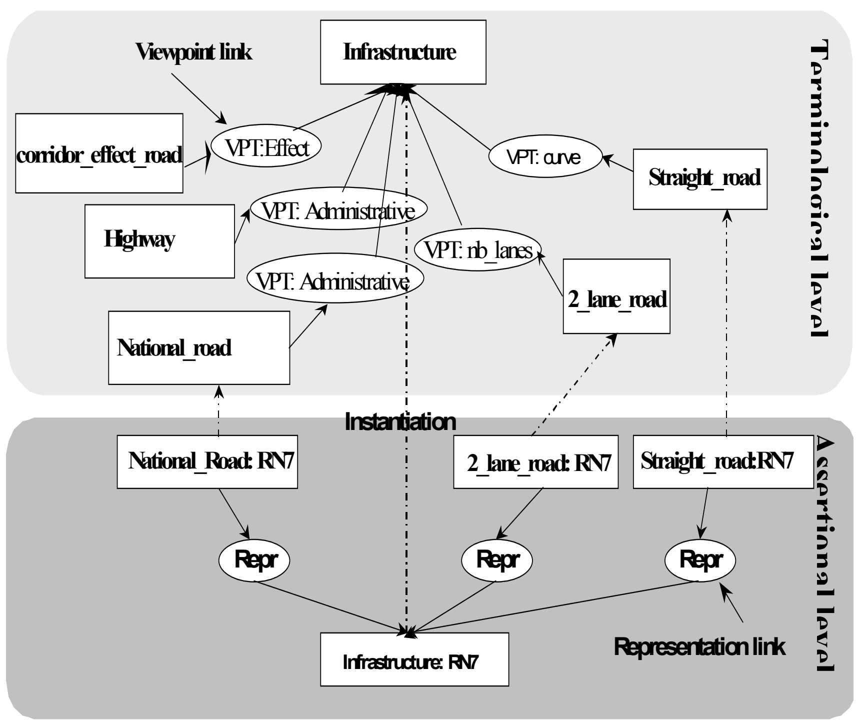

Figure 8: Example of multi-representation with C-VISTA

As such v-oriented concepts are different perspectives on a same object, they are defined from v-oriented concepts, which are indexed by perspective viewpoints.

This model has two advantages:

- It allows to gather information about an object in a same model. So it takes advantage of object-oriented formalism but keeps the advantage of a relational-based formalism.
- It is a dynamic structure allowing to modify an object by adding or removing a perspective on it without having to destroy and recreate the object.

Links between Terminologies

If we try to integrate terminologies in the same concept hierarchy, it is important to interconnect v-oriented concepts. In the example of Figure 6, we notice that some subtypes of Accident_factor are equivalent such as Lack_of_road_exit_specific_signal and Indicator_ambiguity.They are defined in different viewpoints (having the same focus but different view angles) but if they are used in an assertion about the accident, they could be used to mean the same thing.
Therefore, Ribière and Dieng-Kuntz (Ribière, 1999; Ribière and Dieng-Kuntz, 2002) defined three types of links to handle the different possible relations existing between terminologies:

- An equivalence link between two v-oriented concepts stemming from two different viewpoints enables to identify two concepts having the same meaning but used in different contexts (and perhaps named differently) by two experts. It corresponds to the bridge among classes of different perspectives, offered by TROEPS (Marino et al, 1990).
- An inclusion link enables to express that the meaning of the first concept implies that of the second one. It can be useful if two experts express their concepts with different grain levels. For example, in Figure 9, *Overtaking_conflict* and *Left_turn_conflict* are types of accident factors identified by the expert in infrastructure. But in fact, a discussion among the experts reveals that their definitions are included in the concept defined by the psychologist and called *Indicator_ambiguity*. So, the KE can add inclusion links between *Overtaking_conflict* and *Indicator_ambiguity* on the one hand and between *Left_turn_conflict* and *Indicator_ambiguity* on the other one.
- An exclusion link enables to identify the concept s that cannot be at the same time representations of the same instance. For example, the concepts *Straight_road* and *Curved_road* stemming from the *Curve* viewpoint cannot be both used for representations of a same infrastructure.

In C-VISTA model, those three links are proposed as they were the most useful for the intended applications (Ribière, 1999), but other links between v-oriented concept s could be defined: e.g. the composition of viewpoints proposed in (Acker and Porter, 1994).

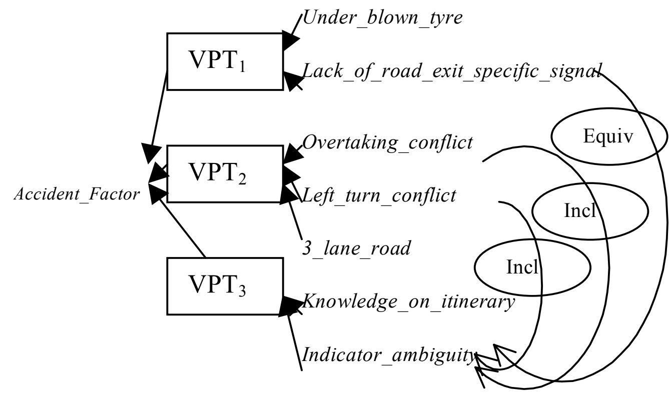

Figure 9: Example of links

3.4.2.3 User's Viewpoint for Accessing the Ontology

A user's viewpoint can be defined using the same viewpoint template as the one used for the ontology building. Thanks to the viewpoint representation, the set of the ontology viewpoints which are included in this user's viewpoint can be determined automatically, so as to restrict to them when visualizing the ontology for this user.

3.4.3 Method for Ontology Cooperative Building with C-VISTA

The KE can exploit C-VISTA model as follows:
1. Elicit knowledge from the experts or analyse their documents (using a knowledge engineering method or possibly a knowledge acquisition method from texts)
2. For each expert, determine the concepts used by this expert.
3. Agree on the viewpoint template to be used for the ontology.
4. Establish the "common top level" of the ontology, composed of the *common concepts* (i.e. denoted by the same terms by all the experts and having the same definition for all of the experts).
5. For each expert, index by *opinion viewpoints* the concepts that are both specific to this expert and subtypes of a common concept. Such concepts are *v-oriented*. Make explicit the focus and view angle of these opinion viewpoints, as well as the initial set of definitions, without yet working on their interpretation: some definitions may be redundant (resp. conflictual).
6. Work with the experts upon this first representation (by analysing the so far obtained concept hierarchy), in order to confront the concept definitions so as to detect the consensual ones.
7. Index the set of consensual definitions by *perspective viewpoints* according to the following rules:
In the *view angle* of the perspective viewpoint, indicate the experts from which the considered concept stems.

Name differently the concepts denoted by the same term by the different experts, if there is no agreement on a common definition: e.g. concatenate the initial name and a feature of the expert.
If two concepts from two different viewpoints have equivalent definitions, they can be linked by an *equivalence link*.
If two concepts from two different viewpoints have compatible definitions, one including the other, an *inclusion link* can be set between them.
If two concepts describe different properties of a basic concept and if, according to the experts, these properties cannot occur simultaneously on a same instance, an *exclusion link* can be set between both concepts.
8. Compare the definitions of the concepts indexed according to *opinion viewpoints* having the same focus. Index the definitions that can be integrated, by *perspective viewpoints* as described in step 6. Track of this integration relies on opinion viewpoints keeping the intermediate definitions elaborated during the construction of the ontology.

This method, obtained by abstraction of our own experiments, offers a methodological guide for helping the humans involved (i.e. the knowledge engineer and the experts) to create the ontology collaboratively.

3.4.4 Co_4: Collaborative construction of consensual knowledge bases

CO4 (for Collaborative construction of consensual knowledge bases) is an infrastructure enabling the collaborative construction of a knowledge base through the web. The consensual knowledge base is meant to represent the consensus among a community about a domain to model (Euzenat, 1995, 1996). The knowledge base is accessible from a HTTP client and can be consulted or edited by authorised people.

A key idea in the approach taken here is that formally expressed knowledge serves a variety of purposes including knowledge and data search, but above all knowledge elaboration (i.e. the organisation and formalisation of knowledge). Knowledge elaboration can be though of as a social process, involving the cooperation of a variety of agents. The CO4 system aims at supporting the elaboration with the help of knowledgeable people, i.e. by enforcing a kind of peer review process on the modifications attempted.

Formality and consistency require more strictness in the protocol than pure peer-reviewing because it is not possible to deal with an inconsistent knowledge base contrary to a paper journal in which the articles do not have to be consistent. This justifies the consensus requirement in which a modification, for being accepted, must have been agreed by all members.

The task of the editor-in-chief is automatically carried out according to a formal protocol, which handles the communication between knowledge bases. The protocol has been fully described and proved consistent, consensual, live and fair under reasonable assumptions (Euzenat, 1997). CO4 can be thought of as a formalised scientific journal (both in its content and in its functioning).

In order for the proposal to be feasible, the cooperators do not directly modify the knowledge base but their own personal workspace. In CO4, anyone is viewed by the system as a knowledge base. In order to build a consensual knowledge base, the individual knowledge bases must be linked together. Knowledge bases are organised into a tree whose leaves are *user knowledge bases* and whose intermediate nodes are called *group knowledge bases*. Each group knowledge base represents the knowledge consensual among its sons (called *subscriber knowledge bases*).

As soon as the knowledge base is part of a group knowledge base, it receives its complete contents, it is entitled to give its opinion on all submissions currently under examination and it is allowed to submit knowledge. A group knowledge base sends to its subscribers messages in order to broadcast a change accepted by everyone and calls for comments in order to establish whether a change must be committed or not. A (group or individual) knowledge base sends to its group knowledge base changes that it wants the group knowledge base to integrate.

When subscribers are sufficiently confident about some pieces of knowledge, they can submit them to their group knowledge base. This proposal is then submitted to the other subscribers as a call for comments. In response, users must answer by one of the following: *accept* when they consider that the knowledge must be integrated in the consensual knowledge base, *reject* when they do not, and *challenge* when they propose another change. When the group knowledge base has gathered enough comments, three cases may happen:

- All of them agree to accept the modification, then the modification is committed into the group knowledge base and broadcast to every subscriber knowledge base;
- One of them rejects the proposal, then the changes are not committed and the comments provided by the rejecter are sent to the submitter;
- One submitter sends a counter-proposal, then the call for comments is replaced by a call for comments about all the available proposals.

The CO4 protocol applies to several levels: the group knowledge bases can be grouped together into a more important group knowledge base and so on.

3.5 Argumentation techniques among agents complying to different ontologies

In this subsection, argumentation techniques and their existing formalizations are overviewed. Argumentation techniques and their formalizations can be implemented on top of the basic ontology structures and knowledge acquisition processes to accomplish the consensus making process.

3.5.1 Argumentation-based inference

Argumentation systems define the notion of an **argument** in terms of an **underlying logical language** and an associated notion of logical consequence. By argument we mean a reason supporting a given conclusion. The underlying logical language and the notion of an argument can be partly or fully unspecified. There exist three main kinds of arguments*: explanation-conclusion pair* [Simari and Loui, 92] [Kohlas et al., 00] [Amgoud and Cayrol, 02a], *inference tree* [Lin and Shoham, 89] [Vreesswijk, 97] and *sequence of inferences* (on line of reasoning) [Pollock, 92], [Prakken and Sartor, 97].

Explanation-conclusion pair approaches use one kind of inference rules but two kinds of premises (certain and uncertain), and leave the structure of the connection between promises and conclusion implicit.

Inference trees and *sequence of inferences* are also known as *derivation-based*. They used two kinds of inference rules (strict and defeasible), and make the structure of the connection between premises and conclusions explicit. An inference tree records the logical dependencies between the various propositions while a sequence of inferences records a particular order in which such a tree can be constructed.

The process leading to the construction of an argument is a monotonic process: new knowledge cannot rule out an argument but only gives rise to new arguments which may interact with the former ones. We distinguish between two kinds of interactions: conflicting arguments or arguments supporting arguments.
The notion of conflict-type interaction can be defined on a set of arguments (collective conflict) or between two arguments (binary conflict). This definition might depend on the argument structure and on the use of a preference relation between arguments.

In order to decide which argument to use, of argumentation systems associate **valuations** with arguments. The aim of valuation is to assign a weight to an argument in order to make comparisons on the set of arguments. It is possible to assign a priori a weight to each argument, but this weight can also be computed.
Three examples of such computation processes:

- *Use of preferences*: frameworks for preference-based argumentation have been proposed by Amgoud and Cyrol in [Amgoud and Cayrol, 02a] and Bench-Capon in [Bench-Capon, 03]. Preferences relations are usually defined from priorities over the belief and take the structure of arguments into account. The priorities can be implicit (e.g. specificity), explicit (e.g. take the form of a partial pre-ordering on the knowledge base) or expressed in the knowledge base itself.

- *Intrinsic value of an argument*: In this case, we do not have priorities over the belief, but we have a weight for each belief. This weight is used to deduce the weight of the arguments, independently of the other arguments. Probability calculus [Kohlas et al., 00], [Krause et. all, 95] and Weakest Link Principle [Pollock, 92] are two examples of approaches.

- *Interaction-based evaluation*: This valuation aims at reflecting the way an argument is defeated and/or supported by other arguments. Among the existing approaches, there are local approaches (which take into account only the arguments directly related to the argument to be valued) and global approaches (which consider all the path leading to the argument to be valued).

Given all the elements above, the argumentation process aims at **selecting** the most acceptable arguments. This step consists in choosing some arguments in the set of arguments using different criteria. The selected arguments are said to be acceptable and the different criteria correspond to different semantics, each semantics being defined by a set of constraints. There exist two kinds of acceptability:

- *Individual acceptability*, which takes into account only one level of interaction, directly selects an argument, and assign to this argument a unique status; see for example [Elvang-Gøransson, et al., 93], [Krause *et. all*, 95];

- *Collective acceptability*, which selects sets of arguments using all the levels of interaction between arguments; the main work is the one proposed by dung [Dung, 95], but the work of Bochman [Bochman, 03] and the labeling approaches by [Jakobovits and Vermeir, 99] are also to be noticed. Arguments can be assigned a unique or a multiple status. Some approaches combine valuation and acceptability in order to define new processes of selection; see for instance [Amgoud and Cayrol, 02a] [Jakobovits and Vermeir, 99] [Cayrol and Lagasquie-Schiex, 04].

The status of the arguments determines the status of the conclusions supported. To this end, argumentation-based inference relations are defined. A proposition is inferred if there exists an acceptable argument in favour of it.The argumentation-based inference process is summed up on Figure 10.

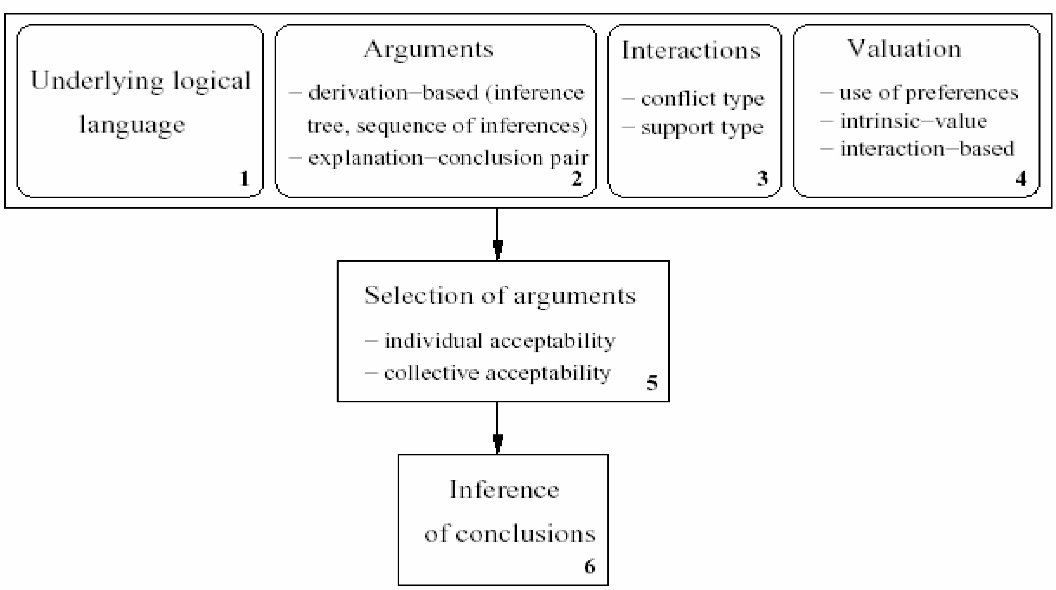

Figure 10: Argumentation-based inference process

The most abstract approach to argumentation is the one proposed by **Dung** [Dung, 95]: it is based on no more than a set of arguments that have no internal structure and a binary defeat relation between arguments. Argumentation is just one domain to which the framework can be applied. Dung defines several semantics for the collective acceptability of arguments (stable, preferred, grounded and complete semantics).
Instantiations of Dung's framework are essentially made in terms of the structure of arguments. An example of derivation-based system is the one of **Pollock** [Pollock, 87], [Pollock, 95] [Pollock, 92] [Pollock, 01].
The approach of **Bondarenko**, **Dung**, **Kowalski** and **Toni** (BDKT system) [Poole, 88] is an assumption-based counterpart o Dung's framework, in which the structure of arguments takes the form of explanation-conclusion pair.

Most existing derivation-based system are for defeasible reasoning while most existing explanation-conclusion pair systems are for plausible reasoning (this is mainly due to the kind of promises and inference rules involved).

Issue concerning self-defeating arguments and odd-defeat loops, hang-yourself arguments, floating conclusions, inductive definitions of argument acceptability, types of defeasible inference rules, reasoning about the strength of arguments and the status-assignment approach to define acceptability are also addressed.

3.5.2 Dialectical proof theories

Several argumentation systems have defined dialectical proof theories aiming to establish the status of an individual argument. These approaches can be explained in terms of an *argument game* between *two players*, a proponent and a opponent. The players *move* alternatingly, moving in general one argument at each turn.

The proponent starts with the argument to be tested. The game evaluates *legality* of moves through some functions that define which arguments can be moved at each point in the game. A *dispute* is an alternating sequence of moves by the two players. A *winning criterion* is a partial function that determines the winner of a dispute, if any. If one player wins, the other one lose, so the argument game is a so-called zero-sum game.
Dialectical proof theories were proposed:

- for Dung's grounded semantics, in order to determine if an argument belongs to the acceptable set under this semantics [Amgoud and Cayrol, 02b] [Prakken and Sartor, 96], [Dung, 94].

- for Dung's preferred semantics, in order to determine if an argument belongs to at least one acceptable set under this semantics ([Vreeswijk and Prakken, 00], [Cayrol, Doutre and Mengin, 03]) or to every acceptable set ([Vreeswijk and Prakken, 00] [Doutre and Mengin, 04]).

- For some generalization of Dung's abstract argument system proposed by Jakobovits in [Jakobovits, 00].
- For the assumption-based framework (BDKT) of [Bondarenko *et. Al,* 97] ([Kakas and F. Toni, 99]).
- For Pollock reasoning architecture [Pollock, 95].

3.5.3 Argumentation for decision making

Argumentation-based decision systems are defined in terms of: an **underlying logical language** and an associated notion of logical consequence, a **Knowledge base**, a **base of goals**, a definition of the notion of an **argument** in favour (or against) a decision, a definition of the **force** of an argument, which depend on the certainty of the beliefs and the priority of the goals which compose the argument, and an **aggregation function** (in the case where the same decision is supported by several arguments, the different arguments have to be aggregated into a single one with a single force).

The level of certainty of beliefs affects the process of decision-making. There are three main approaches for argumentation-based decision making under uncertainty.

The first one is an original approach to qualitative decision, proposed by **Bonet and Geffner** [Bonet and Geffner, 96]. It is based on action rules that link a situation and an action with the satisfaction of a positive or a negative goal. Two measures are associated to each rule: a priority degree, which is the priority degree of the goal of the rule, and a plausibility degree, which depends on the plausibility degree of a input situation. Positive goals provide reasons/argument for actions, whereas negative goals provide reasons/argument against actions. Negative goals should be discarded, and hence any action, which may lead to the satisfaction of such goal, should be avoided. Decisions, which satisfy the most important goals, are privileged. Note that this approach does not refer to any model in argumentative inference.

The second approach, by **Fox and Parsons** [Fox and Persons, 97], is the first one that proposes an argumentation-base framework for reasoning about actions. This framework extends their framework (LA) for reasoning about beliefs.

The last approach, by **Amgoud and Prade** [Amgoud and Prade, 04c], is a logical framework for optimistic and pessimistic decision. This framework uses possibilistic logic to define a knowledge base whose beliefs are assigned a level of certainty, and a base of goals whose goals are assigned a level of priority. Two kind of arguments are defined according to the following two activity towards decision:

- *Pessimistic attitude*: a decision is justified if it leads to the satisfaction of the most important goals, taking into account the most certain part of the knowledge. Arguments for a decision are hence defined: the decision, together with the subset of beliefs, entails the satisfaction of the given goals.

- *Optimistic attitude*: goals may be attained as soon as their negation cannot be proved. Arguments against a decision are defined: the decision, together with the set of beliefs, does not entail the given goals. A decision is the entire better as there does not exist any strong arguments against it.

A level of certainty (depending on the level of certainty of the beliefs) and a degree of satisfaction (depending on the priority of the goals) are associated to the arguments.
This framework is achieved in the case where the knowledge base is consistent, but it is only preliminary in the case where the knowledge base in inconsistent. In this last case, arguments about beliefs are defined, and they are combining with arguments about decisions. Amgoud and Prade intend to extend this framework to multi-criteria decision and to deliberative negotiations.

3.5.4 Dialogue systems

Dialogue systems define the principles of coherent dialogues, the conditions under which an utterance is appropriate, that is, if it furthers the goal of the dialogue in which it is made. Several types of dialogues with different goals can be distinguished. Walton and Krabbe in [Walton and Krabbe, 95] have classified human dialogues as:

Information-seeking dialogues. One particular seeks the answer to some question(s) from another participant, who is believed by the first to know the answer.

Inquiry dialogues. The participants collaborate to search for a truthful answer to some question(s) whose answer(s) is (are) not known to one participant.

Persuasion dialogues. One participant seeks to persuade another participant to endorse a statement (containing beliefs or actions) she does not currently satisfiable.

Negation dialogues .The participants bargain over the division of some scare resource, for which participants have claims, which are not mutually satisfiable.

Deliberation dialogues. Participants collaborate to decide what course of action should be adopted in some situations. This class contains the class of negation dialogues.

According to the typology above, dialogues can focus on *beliefs*, on *actions*, or on both. When a dialogue type focuses on beliefs (resp. actions), then the dialogue system may use *inference* features (resp. *decision-making* features), possibly argumentation-based. These links are summarised in Table 2.

Dialogue type	Focus on beliefs / Use of inference features	Focus on actions / Use of decision-making features
Information-seeking	•	
Inquiry	•	
Persuasion	•	•
Negotiation		•
Deliberation		•

Table 2: Relation of Dialogue Types to Beliefs and Actions

In this document, the term dialogue system only covers the rules of the game, not the principles for playing the game well (i.e. strategies and heuristic).

Elements of dialogue systems

The common elements of dialogue systems are:

- A *dialogue purpose* (or *dialogue goal*)

- A *set of participants* (containing at least two participants), and a *set of roles*. Participants can be humans or computer systems. A participant can have various roles. A participant may or may not have a, possibly inconsistent, *belief base*, or *mental states* (beliefs, goals, etc) which may or may not change during the dialogue. Each participant has a, set of commitments (possibly empty) which usually charges during the dialogues. No relation is assumed between a participant's commitments and belief base or mental states. Participants might also have some *reasoning capabilities*.

- A *communication language* (or ACL) defining the set of *dialogues* and finite dialogues. The syntax is composed of the different *illocutionary acts/speech acts/performatives* that participants can perform during a dialogue. The semantics of the language must be well defined.

- A topic *language* based on a logic, which may or may not be monotonic and may or may not be arguments-based.

- *A context*, built on the topic language. It contains the knowledge that is presupposed and must be respected during a dialogue. The context is assumed consistent and it is fixed and undisputable.

- A *protocol*, which is a set of rules specifying the dialogue states, the permissible moves (speech act) and the allowed replies to each move at each point in a dialogue. There are different types of protocol rules: some regulate dialogical or participant's internal consistency (rationality rules), some are about dialogical coherence, others about the dialogical structure. One can also distinguish the rules that regulate *turntaking* and dialogue *termination*. A protocol P can have several properties:

    - P is *context-independent* if the set of legal moves and the outcome is always independent of the context.

    - P is *fully deterministic* if P always return a singleton or the empty set.

    - P is *deterministic* if the set of moves returned by P at most differ in their prepositional content.

    - P is *single-move* if the turn shifts after each move; otherwise, P, is *multiple-move*.

    - P is *single-reply* if at most one reply to a move is allowed throughout a dialogue; otherwise, P, is *multiple-reply*.

    - P is *immediate-response* if the turn shifts just in case the speaker is the current winner and if the shifts to a current loser.

A participant has some dialectical obligations: for instance, making an allowed move when it is one's turn.

- A set of *effect rules* specifying the effects of utterances on the participants' commitments.
- A set of *outcome rules* defining the outcome of a dialogue.

Participants in a dialogue can also have *strategies* and *heuristic* for playing the dialogue, given their individual dialogue goals.

Persuasion dialogues

Persuasion dialogue systems are dialogue systems instantiated in the following way:

- *Dialogue purpose*: resolve a conflict of opinion about one or more propositions, called the topics. The conflict is resolved if all the parties share the same opinion on the topics.

- *Participants* are at least two. The participant's individual goal is to persuade the other participant(s) to take over her opinion. Participant's roles can be:

- proponent, if the participant has a positive opinion towards a given topic;
- opponent, if the participant has a doubtful opinion towards a given topic;
- third party, if the participant is neutral towards a given topic.

- *Commitments* have several roles: to enforce dialogical consistency in participants; to enlarge the hearer's means to construct arguments; determine termination and outcome of a dialogue; and to determine certain dialectical obligations. Three kinds of commitment can be distinguished; assertions (created by claims) and concessions (created by conceded commitments), and in some games, tacit commitments (entered into by not making particular responses). Assertions have a dialectical obligation attached to them.

- *Logic*: can be used to determine consistency of a participant's commitments, to determine whether the reasons given by a participant for a challenged indeed imply the proposition, and to determine whether a participant respects his assertion or acceptance attitude. Most of the logics in persuasion-dialogue systems are argument-based.

- *Communication language*: it contains some speech arts:
  - claim x: the speaker asserts that x is the case.
  - why x: the speaker challenges that x is the case and asks for reasons why it would be the case.
  - concede x: the speaker admits that x is the case.
  - retract x: the speaker declares that she is no longer committed to x.
  - x since S: the speaker provides reasons why x is the case.
  - question x: the speaker asks another participant's opinion on whether x is the case.

- *Protocol*: A move is a speech act. The typical allowed moves (or replies) after a given move are the following:

In some protocols, like in [Amgoud et al., 00], participants have assertion and acceptance attitudes and have to comply with these attitudes:

- *credulous attitude*: adopted by a participant who can assert any proposition for which she can construct an argument. The participant is said to be confident.
- *cautious attitude*: adopted by a participant who can assert a proposition only if she can construct an argument for it and cannot construct a stronger counter-argument. The participant is said to be careful.
- *skeptical attitude*: adopted by a participant who can assert a proposition only if she can construct an acceptable (in the sense of skeptical inference) argument for the proposition. The participant is said to be thoughtful.

- *Effect rules*: the effect of a speech act on a speaker's set of commitments is:

- *Outcome rules*: they define for each dialogue and topic the winners and losers. When outcomes are fully determined by the participants' opinions and commitments, the dialogue is a pure persuasion dialogue, otherwise, it is a *conflict-resolution dialogue*.

Negotiation dialogues

Negotiation dialogue systems based on argumentation have the following structure:

- *Dialogue purpose*: reaching agreement on the division of some scarce resource, for which participants potentially have claims, which are not mutually satisfiable.

- *Participants* are at least two. A participant has some mental states (beliefs, goals, etc.) and some *reasoning capabilities* that allow it to:

    - generate arguments from its mental states and evaluate those arguments (argumentation rules)
    - make decisions: the participant should select the content of a move if necessary, decide when a given move may be played and choose the following move to play among all the possible ones (decision rules)
    - revise its beliefs or goals thanks to revision rules.

A participant's goal is to get what they most wants.

- *Communication language*: offers, arguments, promises, challenge, accept, refuse, withdraw are examples of speech acts used for negotiation.

Argumentation is used to support offers by arguments. This is one of the main advantages of argument-based negotiation over the other approaches to negotiation (game theoretic and heuristic-based approaches).

Among the argumentative negotiation systems, one can find:

Parson and Jennings' system [Parsons and Jennings, 96] whose basic idea is to construct arguments and counter-arguments to evaluate proposals made during a negotiation; no dialogue protocol is given in this system.

Kraus, Sycara and Evenchik's system [Kraus et al., 98] describes a logic-based approach to agent specification and is implemented using logic programs. An important contribution of this work is the introduction of different types of arguments in the negotiation protocol: appeal to prevailing practice, counter-example, appeal to past promise, appeal to self-interest, promise of future reward, threat. These arguments are treated as speech acts. The recent framework of [Amgoud and H. Prade, 04b] also handles these types of arguments.

Amgoud, Maudet, Parsons' system [Amgoud et al., 00] have proposed a protocol said to be based on James MacKensie's philosophical dialogue game DC. This game allows two participants to argue about the truth of a proposition. The communication language allows four distinct locutions (assert, accept, question, challenge) which be instantiated with a single proposition or an argument for a proposition. The syntax of this protocol has been extended to more than two participants in [Amgoud and Prade, 02]. Note that when a participant asserts a proposition or an argument, then this proposition or argument is inserted in the participant's commitment store. This protocol has a private axiomatic semantics: pre-conditions and post-conditions are defined for each locution, imposing requirements on the participant's mental states; each participant is assumed to be vested with a private reasoning mechanism using argumentation, permitting a preference ordering over the arguments. This system enables modelling of inquiry, persuasion, information-seeking, deliberation and negotiation dialogues (in this last case, as indicated in [Amgoud et al., 00], three additional locutions are suggested: request, promise, refuse).

Amgoud and Prade [Amgoud and Prade, 03] [Amgoud and Prade, 04a] are particularly interested in deliberative negotiation, in which the participants try to find an agreement on a given subject. Possibilistic logic is used as a unifying setting. Their negotiation protocol is similar to the one proposed by Amgoud, Maudet and Parsons in [Amgoud et al., 00].

Information seeking dialogues

An information seeking dialogue system has the following structure:

- *Dialogue purpose*: One participant seeks some information about some statement(s) from another participant who is thought by the first to have some more information in regard to the statement(s).

- *Participants* are two. A participant's individual goal is to acquire or give information.

- *Protocol:* in [Amgoud et al., 00], the dialogue is initiated with a 'question' move, asking if it is the case that the statement(s) hold(s). If the other participant has an argument of or against the statement(s), it asserts this. Then the participants argue about the acceptability of this argument using the argumentation system proposed by [Amgoud and Cayrol, 02b]. When this acceptability is decided, the aim of the information seeking dialogue is reached.

3.6 Socio-Economic Aspects Affecting the Process of Consensus

As said in the introduction, ontologies can be considered as software artifacts that represent a consensual agreement between agents stating which knowledge in a particular setting is expressed and represented [Mentzas, 2002]. More in particular, ontologies provide syntactic and semantic terms and relations between objects to describe knowledge into a domain. It is noteworthy to revise one of the main important definitions of ontologies proposed by Gruber [Gruber, 1993]:

> "An ontology is a formal, explicit specification of a shared conceptualization that holds in a particular context."

The most important goal of ontologies is to provide a formal means to manage and facilitate communication and data/services exchange [Farquhar et al. 1997]. Said differently, ontologies are used for practical reasons i.e. they serves as means to communicate between different entities facilitating interoperability. Furthermore, an ontology has as requisite a well defined formal semantic because it should be used from computer machines that are able to interpret data and exchange information according to the ontology. The ontology is also a description of concepts and relations describing a domain within a community of agents [Russel & Norvig, 1995].
Since the agreement on the terms defined by an ontology is a social process and is influenced by the agreement process this section wants to investigate the consensus from an economic-organizational point of view. In particular, we will use the sensemaking approach, a growing approach in organization studies used to understand how organizations change meanings of things, and how meaning and actions influences and are influenced by the organization structure. As we will see, this approach emphasizes important considerations about ontologies [Garigue, 2003].

3.6.1 Starting considerations

To understand the approach here presented three considerations about organizations and the process of consensus should be introduced.
First of all, the sensemaking approach states that information coming from the environment is often unstructured and ambiguous and need an interpretation ("making sense of it") in order to be transformed into a more structured information asset. The structured information asset is properly an ontology. As we shall see in the following, ambiguity means that information coming from the environment can be interpreted in different even conflicting ways [Weick, 2000]. Such situation has been called ontological ambiguity[2] and allows the enactment of different ontologies. As a result, this condition

[2] "The central proposition is that ontological ambiguity occurs when one concept is represented via several ontologies but in each case there evidence of structural variation such as incomplete list of properties or relationships among the various object or discrepancy in the properties, values or constraints. In particular these applications may use the same concepts in different contexts. (…) This leads to multiple interpretations of unique concepts across different users." [Garigue, 2003].

originates a "confusing similarity" between terms within a community and/or among different communities. Notice that this kind of ambiguity can be viewed as a negative concern from the perspective of information exchanges between different systems but is not considered as such from the sensemaking perspective since it emphasizes the fact that information is intrinsic unstructured and full of meanings. In others words, different ontologies within an organization are the results of the wealth of meanings that can be given to events and objects and not a maladaptive result of irrational persons.

Second, it is also clear that ontologies are both the premise and the product of a social process directed at creating consensus over a particular meaning structure. This structure is a necessary condition to share information between different agents (single, communities or organizations). Even in ambiguous environments, the communication process between agents would be a trivial problem if meaning of things where stable i.e. it does not change during time. In fact, a possible agreement between different ontologies terms could be found. But since meaning inevitably change during time, different agents can have developed different ontologies according to their past experiences i.e. they could have enacted different plausible views of the world. In such a situation, communication and information sharing between agents became difficult because of the discrepancies that could emerge among the different ontologies (among different agent and among different communities/organizations).

Third, when an ontology is developed it became the reference point for the overall community that share it. In this sense the ontology is a structure through which the information system of an organization or a community is developed. This brings to an important consideration: when investments (in this case knowledge-based systems) are developed by organizations, they became both strategic assets - since they represent the particular perspective of the community - and sunk cost (irreversible investment) - since they are developed upon a specific ontology and are not easily changeable.

The interrelation between these three considerations, deeply analyzed in the following, affects the development of ontologies from an organizational point of view. Briefly, the perspective of this section is that meaning negotiation and process of consensus among different actors (either single or communities) is heavily influenced by the irreversible investments developed in the past. This is because these investments became useless if a new ontology, that does not permit the use of the developed investments, is developed.

3.6.2 Ambiguity in organizational settings

It's to be underlined that, according to a sensemaking approach, ambiguity can be referred to subjective computational limitations, but also to an objective configuration of the environment that can be shaped, through manipulative actions, according to alternative interpretations.

On the one hand, the above described form of a subjective lack of information of the decision maker. It is caused by the decision maker's cognitive limitations even when all the possible events are predetermined or *ex-ante* foreseeable. On the other hand, the lack of meaning can be referred as the possibility of several interpretations of environmental signals and objects. Ambiguous environments are those in which there is a lack of meaning since there is no sufficient information in order to formulate a unique frame according to which events can be interpreted and probabilities formulated [Cohen et al. 1972; Hatch, 1997]. As a consequence, agents can formulate alternative and even conflicting frames and, according to these, information can be interpreted in different and plausible ways [Gioia and Chittipetti, 1991; Weick, 1995]. As a consequence, in ambiguous environments people could be not able to communicate adequately since language is not consciously articulated to support them [Mintzberg, 1978]. From this perspective, ambiguity qualifies situations in which there are multiple plausible readings of the same "raw data". This means that the feedback received from the environment needs to be transformed into a finished artifact: it is ambiguous and as such needs an interpretation. In ontological terms, this constructivist approach underlines that there is not an *a-priori* correct interpretation of the world but rather a continuous accomplishment of a possible coupling between a subject and the context. Such a vision of the reality is emphasized by the fact that continuously changing environments, for instance due to technological changes, causes continuous changes of the meanings of things [Daft et al, 1984; Stone et al., 1996; Gomez et al. 2000].

3.6.3 The sunk cost effect

As clearly stated by Arkes and Blumer, a sunk cost occurs when an investment in "money, effort or time has been made" [Arkes et al., 1985]. In particular, sunk costs can be considered as irreversible investments that cannot be changed in some other resource. The main characteristics of sunk costs are two. Sunk costs are reusable meaning that they are subject to economies of reuse. Second, sunk costs are irreversible i.e. these resources cannot be transformed into an alternative resource having equal value.

Two different research fields judge in opposite ways the reasoning process influenced by sunk costs. In the psychological decision making research field, a lot of empirical studies show that people, when deciding, are heavily influenced by their past investments. Typically such influence is considered irrational since decision theory states that rational decisions should be based on future expectations and costs. Typically several psychological argumentations are used to explain human behaviors in such a way. For instance, the most known arguments are the "don't waste" rule [Arkes, 1996], the self-justification explanation [Brockner, 1992] the prospect theory effect [Whyte, 1986], the project completion [Garland et al. 1998], and the mental accounting explanation [Thaler, 1994].

Another school of thought in this area that falls under the notion of decision dilemma theory, proposes a critic from a quite different point of view [Bowen, 1987; Bowen and Power, 1993; Nortcraft and Wolf 1984; Hantula et al. 1999]. Their contribution opens sunk costs studies to the wider research inquiry on sensemaking and organizational epistemology, throwing a constructivist light onto these supposedly irrational behaviors.

They state that reasoning influenced by sunk costs is not necessarily irrational when people have to manage unstructured information in which terms and events can have different plausible meanings (such as ambiguous environments).

3.6.4 Ambiguity, sunk cost effect and meaning

Ambiguity and sunk costs, if turned in the ontology field, allow to understand that, for a certain community, an ontology can:

1. Be developed according to the particular perspectives used by each community each of which uses similar concepts for different meanings. For instance, glossaries, repositories, dictionaries, databases schemas or knowledge bases can be developed for each specific domain. In this sense ontologies can be viewed as theories that use a specific language to describe entities, properties and relations within a certain perspective (Fonseca et al. 2002).
2. Be a structure through which irreversible investments are developed. In fact, knowledge bases, information retrieval applications, information systems and the data contained are applications that depend heavily and works well according to a particular formalization of the ontology. The strong formal definition of the semantic in the ontology is properly necessary for the right and coherent functionality of such systems.

An example is useful to understand this underlining concern. Consider as a starting point the ontology of a user[3]. According to this "representation" of the world, the user starts to develop investments like, for instance, acquisition and development of particular information systems, development of databases and of infrastructures that allow for the exchange of data with other users that share the same ontology. The ontology drives the user in doing a certain type of investments. In other words, the contexts can be considered as the "view of the world" that legitimates the irreversible investments made in the past by the users. In fact, a particular definition of the context gives meaning to the investments: it is possible to say that the meanings described by a particular ontology are "looked-in" by past investments and, at the same time, the values of irreversible investments are given by the particular ontology used. Clearly a change in the meaning of the concepts (i.e. of the ontology) is a "costs" for the user if the new ontology does not allow for the use of past irreversible investments.

Two economic principles can be introduced according to what has been presented until now. In ambiguous environments, these principles can influence the negotiation of meaning. First, from the perspective of ontology holders in ambiguous situations, an interpretation is "true" the more they bounded to it irreversible resources (sunk costs). In economic terms, such situation shows the so called "lock-in effect": the sunk cost is a barrier (a wasted cost) to exit (abandon an ontology) (Arthur, 1988; 1989).
Second, the value of an ontology increases the more it is shared since more people and systems become interoperable. In economic terms, an ontology is subject to "network externalities": every agent that joins in increases the value for each network participant (i.e. the more people use an ontology the more the interoperability opportunities increase). Figure 11 summarizes these two principles.

[3] Here user or community are used interchangeably since the considerations are valuable both for single users and communities that share the same ontology.

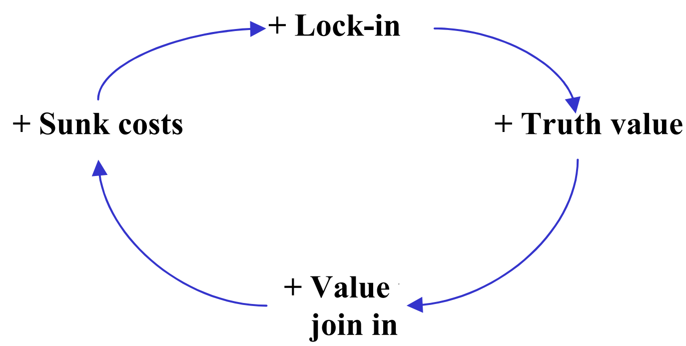

Figure 11: Lock -in effect and network externalities effects

In this sense, the investments stabilize the meaning of concepts while ontologies influence the development of certain investments type favoring the information exchange. As we will see in the description below, if we transpose these conceptualizations in the achievement of consensus within an ontology-based community, users will hardly change idea about a certain concept if, in so doing, they loose a high amount of past irreversible investments.

3.6.5 The process of consensus in ambiguous environments

The process of meaning negotiation and modification of an ontology [O1] can be conceptualized as following the next steps:

1. The process starts when different users exchange data according to a particular ontology. We can say that these users map the different context using as a reference point the ontology (Figure 12).

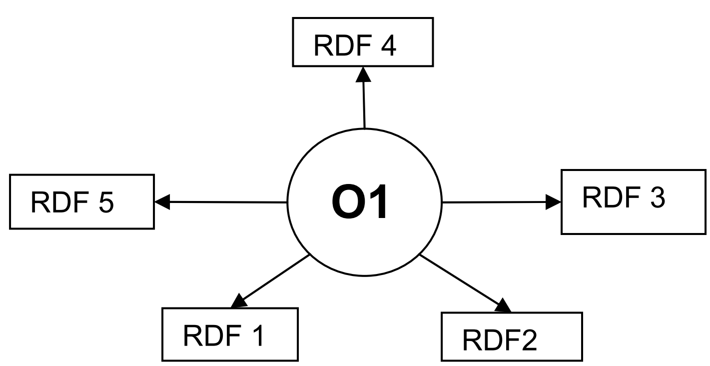

Figure 12: Ontology for information sharing

2. During time the two users exchange data with other entities that have different ontologies. The first two users need also to modify their ontologies in order to have a communication with other entities. When, after a certain period of time, the two users restart to communicate they make sense that the results of the data-exchanges do not fit the requests. In other words, the users are seeking for contents that cannot be found trough the current ontologies. The interaction with other entities has in some way modified the concepts they are seeking for, and the current mappings do not satisfy their communication needs. Users need to change the meanings of things (Figure 13).

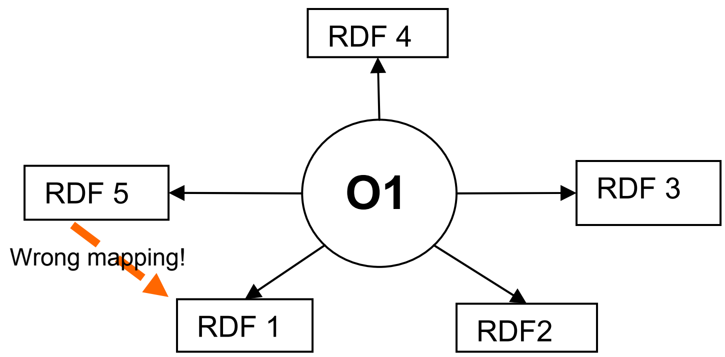

Figure 13: Examples of a wrong mapping

3. If the current mappings do not satisfy the interaction, the users can follow three different ways to re-establish the communication (see Figure 13):
    a. ***Coordination***: they can try to remap all the contexts. If this is possible with zero costs we say that the users are coordinating themselves.
    b. ***Negotiation***: If the simple coordination is not possible (for instance because the mappings are in some way incoherent) the users can try to negotiate the meanings of the concepts in order to find an agreement. With negotiation we mean that users have to change their context performing the economic calculation of the sunk costs as before presented. More precisely, the negotiation can happen in two ways:
        i. In the firs case, the two users can modify their current contexts to find an agreement and re-establish communications disregarding the [O1] i.e. the try to redefine concepts ignoring other users. In this situation the two users can utilize a specific ontology [SO] (the result of the negotiation process) for their particular purposes while use [O1] with the other users (see Figure 14). Negotiation implies a certain costs of change because some concepts needs to be modified, and, as we have proposed above, this implies that users will loose some of the “sunk costs” (the irreversible investments) developed for the original context. Summarizing, this phase corresponds to the fact that versioning is done in a social environment and the change of concepts is done trough a

economic calculus. Point 4 will explain why we consider this phase not enough to generate a stable interaction between users.

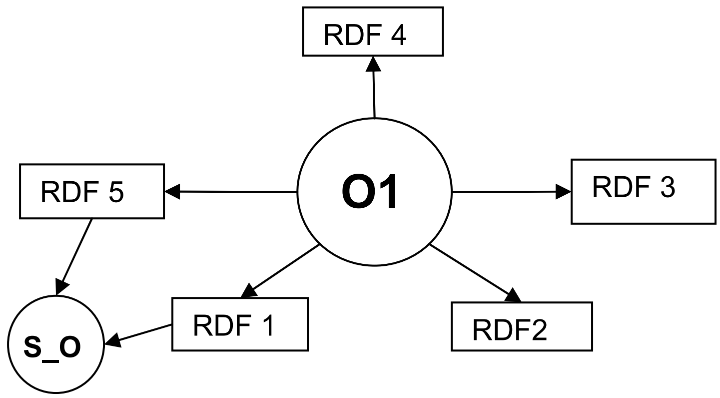

Figure 14: Special ontology

ii. In the second case, the users can try to involve other users in the negotiation process as already described. This case is oriented in changing [O1]. Obviously it implies a more complicated negotiation process but is still similar to the one presented above (see point **i**).

4. Coordination and negotiation are not sufficient to make stable versioning of an ontology. In fact, users have lost irreversible investments in the negotiation process for the developing of the new Ontology but it does not implies for the fact that users will not change ideas about the new concepts in the future. We say that the achieved agreement is not stable. In this case, our statement is that to make an ontology "real" and stable i.e. to force users to use it, each user has to develop irreversible investments in this new categorization. As before said, in so doing users will hardly change idea in the future. We call this phase **escalation** or **reification** (see the entire process in Figure 15). In other words, users should develop sunk costs in the new Ontology (even if we are in the point [**i**] or [**ii**]). Only if subjects invest in the new categorization the new concepts will be sufficiently stable for communications.

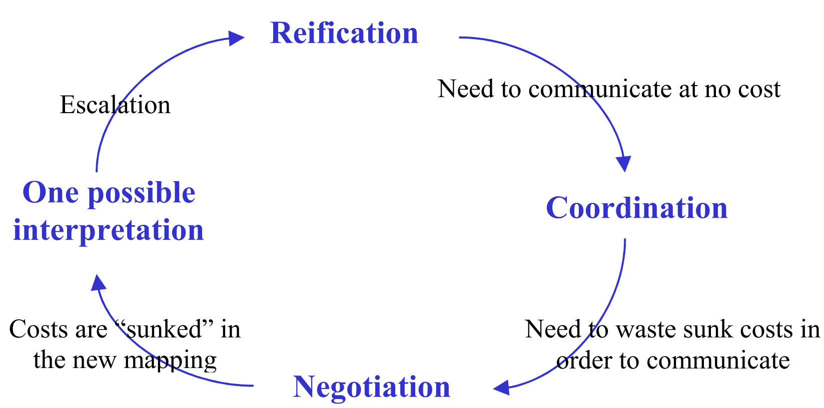

Figure 15: Converging trends of the consensus process

In conclusion, if ambiguity is considered as one of the main characteristics that influences organizational environments (allowing for different ontology conceptualizations), the perspective of doing sunk costs (irreversible investments on a specific ontology), stress the ontology holders at continuously confirm the already developed ontology. That is, the truth of a particular meaning configuration is increasingly assumed according to what has been developed.
In this sense, the development of a new ontology is influenced by the "single" sunk cost of each user that wants to share a new ontology. From this perspective, as we have seen, the coordination as a means to find an agreement about the general structure of an ontology has no cost. It is flexible but not really reliable; in fact, the mapping is not a sunk cost and thus is reversible (one or more participants can freely leave the agreement). On the other hand, the negotiation as a means to find an agreement on an ontology has a cost for the users. It is less flexible but more reliable than negotiation. In fact, a least one part has to invest in the change of its investments and will retrospectively tend to confirm the new ontology. The mapping becomes a sunk cost. In conclusion, a stable agreement (consensus) on a new developed ontology can be achieved forcing users to make investments on this ontology.

4 Tools for Knowledge Acquisition and Consensus Making

In this section, we review the existing tools/prototypes for knowledge acquisition and consensus. The goal of the section is identification of reusable practices, approaches, components for the knowledge acquisition and consensus process modeling and implementation.

4.1 Knowledge Acquisition from Individuals and Communities on the Web Environments

In this section, we describe the tools supporting consensus making processes between individuals and communities.

Ontology development and editing policies are quite simple on most of the current Semantic Web portals [Stollberg et al., 04]: ordinary portal users do not participate in construction of ontologies, though they often can introduce their ontology instances (e.g., as in KnowledgeWeb[4] and Esperonto[5] Semantic Web portals based on ODESeW [Corcho et al., 03]). Exceptionally, the users can propose changes to ontology structure, but these changes need to be approved by the main ontology editor [Pinto et al., 04]. Obviously, this approach to ontology development and editing is not dynamic, does not consider heterogeneity, personalization and community aspects, is not scalable, and thus can not

[4] KnowledgeWeb portal: http://knowledgeweb.semanticweb.org
[5] Esperonto portal: http://esperonto.net

serve as a basis for organization of an effective communication process. Though the People's portal environment [Zhdanova04] supports functions that are typical for Semantic Web portals in general, it is different, because of allowing the portal members to specify knowledge representation issues of their Semantic Web portal, and thus, develop their own portal themselves.

In analogy with the FOAF project[6], the People's portal environment provides means (similar to foaf-a-matic) to create semantic annotations on people's personal details or other portal content the portal members might want to bring in. The specifics of the People's portal environment is that its users actually produce machine readable pages to make use of the portal, whereas FOAF project approach focuses on the promotion and improvement of a specific ontology, but not on the FOAF ontology application, usage and dynamic user-driven evolution. Meanwhile, recent research has shown effectiveness of knowledge acquisition from web users, and the same research also brought understanding that in order to be a success knowledge acquisition applications need to move out from the game and toy area and be tightly integrated with applications that are of actual use to the community [Chklovski03].

In comparison to Wiki and Open Directory Project[7] approaches, where "netizens" are encouraged to bring structured knowledge on the web, the People's portal environment aims at reaching more semantic granularity in specifying the portal content. The People's portal environment provides the means for collaborative development of ontologies. However, it is different from environments for explicit web-based collaborative ontology development [Domingue98] [Farquhar et al., 97], which resulted to be of limited practical usage. The People's portal environment makes the users involved in creation, extension and reuse of ontologies implicitly in order to increase the value of the portal.

4.2 Knowledge Acquisition from Natural Language Sources

In this subsection, tools supporting knowledge acquisition from natural language sources are overviewed.

4.2.1 Text analysis approaches

As discussed in Section 3.2.2, text analysis approaches only really deal with the knowledge elicitation process, relying on human intervention for the trickier aspect of organising the found knowledge.

TOPKAT [Kingston94] is a KA system designed to extract relevant nouns from transcriptions of interviews with domain experts, though presumably the approach would also work with texts such as textbooks. Frequency information is used to determine

[6] FOAF project: http://www.foaf-project.org

[7] Open Directory project: http://dmoz.org

which nouns should be identified as domain concepts. Adjectives modifying such nouns are presumed to suggest some attribute values for the concepts in question. The extracted material is then presented to a human for verification and classification. Clearly the approach is a very simple one, but it acts as a first step in the automation of the KA process and reduces the time required to carry out the task manually. The approach could be combined with further automatic processes to organise the knowledge acquired.

KRITON [Diederich et al., 88] and KITTEN [Shaw and Gaines, 88] use similar techniques to acquire background domain knowledge from textbooks. Frequency information is again used to extract relevant concepts, which are then used as the basis for further KA episodes from the domain expert (at least, in the case of KITTEN).

4.2.2 Pattern matching approaches

PETRARCA [Velardi89] identifies “surface semantic patterns” (SSPs) from NL texts in an attempt to acquire knowledge about word definitions. The corpus used consists of press agency releases in the domain of finance and economics, and is first processed using morphological and syntactic analysis. Then the system attempts to derive interpretations of unknown nouns from the text, using syntax-to-semantics rules as discussed in Section 3.2.3. For example, the preposition “of” suggest a possession relation between the two nouns it links. There are many other similar approaches, mostly linked with a specific domain, such as [Bowden et al., 96] and [Oakes and Paice, 01], both of whom rely largely on “trigger words” to identify relevant patterns. As discussed earlier in 3.2.3, this is both an advantage and a drawback to using lexical information – on the one hand, patterns are easier to define for a restricted domain, but on the other hand, portability remains an issue.

4.2.3 Contextual approaches

The COGNITERM project [Meyer01] aims to find knowledge-rich contexts in the domain of composting, as a means of semi-automatic knowledge extraction. The idea is to build on the concept of a KWIC (KeyWord In Context) concordancer, which shows all the given occurrences of a term together with its context. KWIC concordancers have the drawback (for KA purposes) that a human domain expert is still required to sift through the concordances found in order to identify those that are most knowledge-rich, i.e. those which exhibit the most important contextual characteristics for the search term. The approach uses a pattern-matching methodology similar to those described in Section 4.2.x above, the difference being that the aim is to identify whole contextual patterns themselves given the search term, rather than using the contextual patterns to find relevant terms.

The TRUCKS system [Maynard and Ananiadou, 99] presents similar ideas. In this approach, however, the aim is to combine syntactic and semantic information about terms and their contexts in order to produce clusters of related contexts, grouped by similarity. This is then used as a bootstrapping mechanism for knowledge acquisition and ontology creation and tuning.

4.2.4 Information Extraction approaches

As described in Section 3.2.5, one of the uses of information extraction (IE) is for the purposes of knowledge acquisition from text. GATE contains a generic IE system which uses a knowledge-engineering approach [Maynard et al., 02, Maynard and Cunningham, 03]. This consists of a pipeline of processing resources run sequentially over a set of texts. Modules include tokenization, sentence splitting, part-of-speech tagging, semantic annotation, and coreference, while other optional modules such as morphological analysis can be plugged in. Adapting the system to different domains, applications and languages requires minimal alteration to the system since each module is independent from the others and is designed to be as generic as possible. On news texts the system achieves in the 90th percentile for Precision and Recall.

Most recent approaches to IE are turning towards machine learning rather than rule-based systems such as GATE, mainly because large amounts of training data are becoming available, especially with the development of mechanisms to bootstrap this process. One such system is BBN’s Identifinder [Bikel et al., 99], which uses a cascade of 4 statistically trained models: parsing and name finding, name classification and linking, description classification and linking, and pronoun resolution. Heuristic rules are used for nominal coreference, metonymy resolution and generic detection. IBM's approach [Florian et al., 04] uses Maximum Entropy Models. The model can use arbitrary feature types, being able to integrate a wide variety of lexical, syntactic and semantic features. Crucially, it also uses feature streams derived from different named entity classifiers.

Other systems use a combination of the two approaches, such as NYU’s Proteus Project [Yangarber et al., 00], which uses hand-coded rules augmented with machine learning. The machine learning component has over 33 features, including the hand-coded rules, features from the parse, features from gazetteers, dictionaries and training data statistics, and genre information generated by heuristics.

4.2.5 Ontology-based approaches

As described in Section 2.3.6, ontology-based approaches operate on a top-down principle and rely on a pre-existing ontology and typically either a pattern-matching or IE based approach. The current wave of new tools for the semantic web has led to the development of methods for automatic metadata creation and ontology population, such as the following.

Magpie [Domingue et al., 04] is a suite of tools which supports the interpretation of webpages and "collaborative sense-making", by automatically populating an ontology from relevant web sources. It can be used with different ontologies. The principle behind it is that it uses an ontology to provide a very specific and personalised viewpoint of the webpages the user wishes to browse.

KIM [Popov04] is an architecture for automatic semantic annotation developed within a platform for semantic-based indexing and retrieval from large document collections. KIM contains an instance base which has been pre-populated with 200,000 entities (mostly locations), and performs information extraction based on GATE [Cunningham et al., 02]. Essentially, KIM recognises entities in the text with respect to the KIM ontology, and adds new instances where they do not already exist.

The SemTag system [Dill et al., 03] performs large-scale semantic annotation with respect to the TAP ontology. It first performs a lookup phase annotating all possible mentions of instances from the TAP ontology, and then performs disambiguation, using a vector-space model to assign the correct ontological class or determine that this mention does not correspond to a class in the ontology.

[Hahn and Schnattinger, 98] uses a system based on pattern-matching similar to PETRARCA, but its goal is the extension of existing ontologies using information exploited from parsing the text. For example, the pattern "operating system OS2" suggests that "OS2" is a new (unknown) instance of the (known) concept "operating system", and can thus be added to the ontology.

4.2.6 Machine learning approaches

The Proteus project mentioned in Section 4.2.4 also includes an approach to unsupervised, or minimally supervised, knowledge acquisition [Yangarber02]. This is based on bootstrapping a comprehensive knowledge base from a small set of seed elements. The approach is embodied in algorithms for discovery of lexicon, concept classes, and patterns, from raw, un-annotated text.

DBMiner [Han et al., 1996] is a database mining system which makes use of an induction method with attribute oriented induction for learning characteristic rules and discriminate rules in relational databases. It performs dynamic adjustment of concept hierarchies and automatic generation of numeric hierarchies. The system allows to discovers different kinds of knowledge rules and generates different forms of outputs including generalized relations and multiple forms of generalized rules. The system offers a graphical user interfaces for interactive knowledge mining. DBMiner combines machine learning algorithms with database technologies

MOBAL [Sommer, 1994] is a system for developing operational models of application domains in a first order logic representation. It integrates a manual knowledge acquisition and inspection environment, an inference engine, machine learning methods for automated knowledge acquisition, and a knowledge revision tool. The knowledge acquisition environment offered by Mobal allows to develop a model of the domain in terms of logical rules. Mobal also integrates several machine learning methods to automatically discover additional rules based on the facts entered by the user, or to form new concepts. If there are contradictions in the knowledge base due to incorrect rules or facts, there is a knowledge revision tool to help the user locate the problem and fix it.

SOAT [Wu et al., 2002] allows a semi-automatic domain knowledge acquisition from a domain corpus, extracting relationships among existing concepts from parsed sentences based on applying phrase-rules to identify keywords. The system combines NLP tools with different ML algorithms. To perform its goals, a set of rules has been defined to extract keywords. The tool receives as input a domain corpus with the part of speech tags. A keyword, usually the name of the domain, is selected as root in the corpus. Then, with this keyword, the process aims to find a new related keyword to the previous one applying extraction rules and adding the new keyword into the ontology. This new keyword is now taken as root to repeat the process during a determined number of times or until it is impossible to find a new related keyword. The user intervention is necessary to verify the results of the acquisition and to refine and update the extraction rules.

Weka [Witten and Frank, 1999] is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own Java code, allowing an easy use of these techniques by other platforms or systems. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.

4.3 Knowledge acquisition process for multimedia - learning and ontology-based approaches

In the following, an approach for knowledge-assisted semantic multimedia content analysis and annotation, based on an ontology infrastructure is presented [Dasiopoulou et al.., 04]. This work is based on [Mezaris et al., 04b] where a framework for compressed-domain video analysis exploiting domain-specific knowledge is described. In the proposed approach, semantic and low-level attributes of the objects to be detected in combination with appropriately defined rules determine the set of algorithms and parameters required for the objects detection. Semantic concepts within the context of the examined domain are defined in an ontology, enriched with qualitative attributes of the semantic objects (e.g. color homogeneity), multimedia processing methods (e.g. color clustering), and numerical data or low-level features generated via training (e.g. color models, also defined in the ontology). Semantic Web technologies are used for knowledge representation in the RDF(S) metadata standard. Rules in F-logic are defined to describe how tools for multimedia analysis should be applied, depending on object attributes and low-level features, for the detection of video objects corresponding to the semantic concepts defined in the ontology. This supports flexible and managed execution of various application and domain independent multimedia analysis tasks.

The general system architecture, depicted in Figure 16, consists of a knowledge base (including both the developed ontology and rules), an inference engine, the algorithm repository containing the necessary multimedia analysis tools and the system main processing module, which performs the analysis task, using the appropriate sets of tools and multimedia features, for the semantic multimedia description extraction.

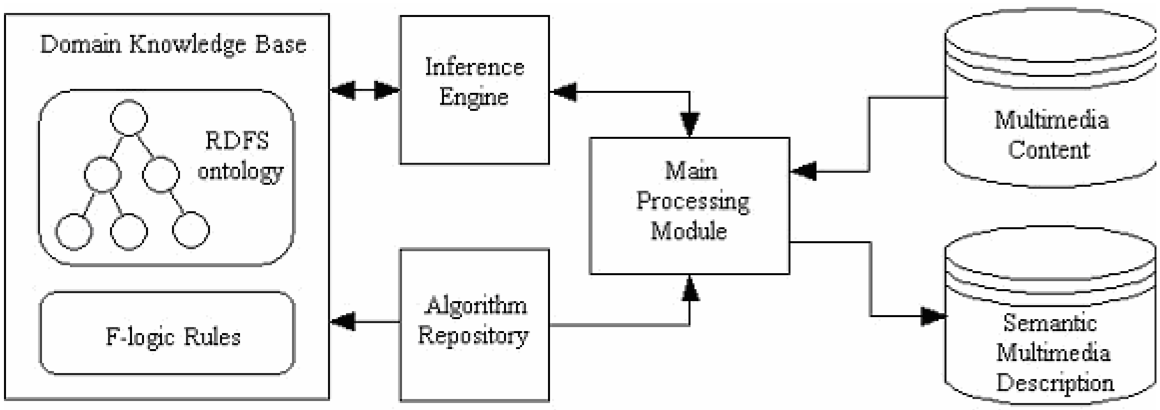

Figure 16: Multimedia Knowledge Acquisition Architecture

Following this approach, the multimedia semantic analysis and annotation process largely depend on the knowledge base of the system and as a result the method can easily be applied to different domains provided that the knowledge base is enriched with the respective domain ontology. Extending the knowledge base with spatial and temporal objects interrelations would be an important step towards the detection of semantically important events for the particular domain (e.g. a car getting out of the road, a player scoring a goal), achieving thus a finer, high-level semantic annotation. In addition, the ontology-based approach also ensures that semantic web services and applications have a greater chance of discovering and exploiting the information and knowledge in multimedia data.

In order to implement the knowledge-assisted multimedia content analysis technique, an analysis and a domain ontology are constructed. The multimedia analysis ontology is used to support the detection of domain specific objects. Knowledge about the domain under discourse is also represented in the form of ontology, namely the domain specific ontology. The domain-independent, primitive classes comprising the analysis ontology serve as attachment points allowing the integration of the two ontologies.

The choice of algorithms employed for the detection of each object is directly dependent on its available characteristic features. This association is determined by a set of properly defined rules represented, as mentioned earlier, in F-logic. The rules required for the presented approach include: rules to define the mapping between algorithms and features (which implicitly define the object detection steps), rules to determine algorithm input parameters and rules to deal with object interdependencies as explained above. However, during the analysis process, a priority is given to certain algorithms, which affects the actual order of execution.

The proposed analysis procedure, as defined by the system ontology and rules, has been applied to Formula One racing and Football videos and produces satisfactory results. The same methodology could be easily applied to different domains by using the appropriate domain ontology. The followed multimedia analysis approach provides a framework for ontology-based annotation of multimedia content enabling semantic transcoding and key Semantic Web functionalities such as querying, retrieval and reasoning.

All ontologies are specified in RDFS format using the OntoEdit engineering environment [Sure et al., 02], while for the inference and querying service the OntoBroker inference engine [Decker et al., 99] was used.

4.4 Modeling of Consensus Between Individuals and Communities [8]

In a distributed environment, concurrent access mechanisms are insufficient for users who are working together editing an ontology or several different ontologies. In the opinion of consensus tool developers, the concurrent access mechanisms to ontology sharing and evolving should be enriched by collaborative facilities that contribute to generating a virtual environment where people can share ideas, discuss approaches or warn other people about certain operations that are currently executed or will be executed. Reaching consensus collaboratively in ontology editing environments, building consensual knowledge in a distributed manner and applying ontology views are the major directions in development of the current consensus tools described below. A larger overview of potentially relevant to the knowledge acquisition and consensus modeling ontology management methodologies and tools was done by Martin-Recuerda et al. [Martin-Recuerda et al., 04].

4.4.1 Collaboration Tool Support

A first step in the development of ontology editing tools with collaborative features is recently introduced by the OWL Plug-in for Protégé [Knublauch et al., 04]. With included set of ontology tests to check against best ontology design practices, ontology engineers can ensure that the developed ontologies have interoperable common features, e.g., OWL DL compliance. The ontology test mechanism has also been exploited to implement a simple but powerful "*to-do-list*" feature. A to-do-list is a proposal of tasks which a user suggests to the other participants in the process of creating and maintaining a particular ontology. This mechanism helps to coordinate shared ontology design efforts [Knublauch et al., 04].

OntoEdit [Sure et al., 02] provides two tools designed expressly for specification of requirements in the design of an ontology. OntoKick includes specific features for collaborative generation of requirements specifications for ontologies, and Mind2Onto is a plug-in for supporting brainstorming and discussion about ontology structures. Also, during the design phase, the participants can store comments (for example, design

[8] The current text of the section is a starting point from Francisco Martin-Recuerda (UIBK) – further the section needs to be rewritten with a bias towards consensus, and selection of reviewed tools needs to be modified accordingly

decisions) in the documentation field of each term of an ontology, and clients are immediately informed of changes that other participants are producing.

KAON ([Gabel et al., 04; Volz et al., 03; Motik et al., 02]), in our opinion does not provide full-fledged facilities for collaboration as opposed to multi-user (concurrent) support. Though, KAON allows users to work together in editing the same ontology and can warn the ontology engineers about conflicts in the proposed by them modifications, but there are no workgroup facilities. The same point arises in relation to Ontolingua [Farquhar et al., 97], which has a similar notification mechanism. In addition, the authors of the Ontolingua system stress its ability to support group sessions where the users working on an ontology are organized in groups.

Tadzebao [Domingue98] is an ontology discussion tool which supports asynchronous and synchronous communication facilities. Tadzebao is an application on top of WebOnto [Domingue98], an ontology library system, designed to support collaborative creation, browsing and editing of ontologies. Tadzebao is integrated in the architecture of WebOnto and includes two main components: a Tadzebao client which is the front-end of the tool and manages the interaction with the users; and a Tadzebao server that represents the back-end and maintains all the annotations that the clients include in the client-side. Tadzebao client uses the idea of virtual notepad that integrates editing tools, such as *text editors* and *drawing tools*, for expressing general ideas about the definition or modification of ontologies. Users can include hand-drawn sketches, GIF images, text comments and ontology components represented in OCML [Motta98]. These inputs are automatically copied to the Tadzebao clients that are involved in the interaction, so the users can follow and participate in the discussion "*in real time*" (synchronous) or see the result of the discussion when they start the client (asynchronous).

APEKS (Adaptative Presentation Environment for Collaborative Knowledge Structuring) [Tennison and Schadbolt, 98], is a tool with collaborative facilities for creating personal ontologies. The approach of APEKS is to allow users to define their own versions of an ontology, and then apply comparison mechanisms to detect differences between versions and prompt these differences to the users in order to start a discussion and reach a consensus between the different proposals. Users can interact with the system and with each other using a program based in a previous development called MOO (Multi-user text-based virtual environment) [Curtis92]. MOO supports synchronous and asynchronous textual communication (no multimedia facilities like in Tadzebao) where the comments of the users are displayed in chronological order.

4.4.2 Distributed Construction of Consensual Knowledge

Several methodologies for building ontologies were proposed [Davies et al., 02], and most of them do not take into account that defining ontology as a shared conceptualization requires consensus between the authors. However, methodologies and tool support targeted at consensual ontology development in distributed environments exist.

CO4 (Collaborative construction of consensual knowledge bases) [Euzenat96] is an infrastructure enabling the collaborative construction of a knowledge base through the web. The main contribution of this approach is a proposal for organizing KBs in a tree structure. The leaves are called user KBs, and the intermediate nodes, group KBs. Each group knowledge base represents the knowledge consensual among its sons (called *subscriber knowledge bases*). When a subscriber wants to extend their group knowledge base, they submit a proposal with the modifications to the other subscribers. In response, users must answer by one of the following: *accept* when they consider that the knowledge must be integrated in the consensual knowledge base, *reject* when they do not, and *challenge* when they propose another change.

DILIGENT (Distributed Loosely-Controlled and evolving Engineering of oNTologies) [Pinto et al., 04] is an approach in the area of decentralized and individualized knowledge management. The content of the work was provision of IT infrastructure for individualized knowledge work and harmonization of vocabularies/ontologies developed in this infrastructure. The approach adheres to applying roles of ontology engineers, ontology users and control board editors and assigning responsibilities to the role owners for the actions in the harmonization process, namely: ontology build, analysis, revise, and local update.

4.4.3 Ontology Views: Reaching Consensus through Personalization

Similar to database views that provide a specific visualization of part of the database instances, the ontology view approaches exploit the idea of views in ontologies. However, ontology views implementation is a more complex task due to presence of terminological specification in addition to instances. Further, extraction of self-contained portions of ontologies and adaptation to further restrictions brings more complexity. The use of views for visualizing the content of ontologies provides a useful mechanism for reaching consensus via personalization. Users can reach easily an agreement about the structure and instances of an ontology, because views will help to adapt the visualization of the data to each particular need.

In KAON implementation of ontology views, an extension of the query language RQL ([Alexaki et al., 00a], [Alexaki et al., 00]) is used to generate views from RDFS ontologies [Volz et al., 03a]. Volz and colleagues distinguish between two types of views: views on classes applied to concepts (classes) returning only unary predicates, and views on properties can be defined using arbitrary queries which return binary predicates.

Another ontology view related work [Noy and Musen, 04] is motivated by the fact that use of extended query languages to generate ontology views as in the KAON approach does not allow users to specify a portion of an ontology that results from a particular traversal of ontology links. Noy and Musen define a Traversal View as "*a subset of an ontology that consists of classes and instances on the path of the traversal specified in the view*". To generate a Traversal view, they propose a method that starts with the selection of a concept that will belong to the Traversal view, a list of relationships (property names) that should be traversed and the maximum distance to traverse along each of the relationships.

The approach of Noy and Musen is oriented on ontologies specified in RDF Schema, though it is also applicable for OWL ontologies with limitations. The advantage of this approach over the previous approaches is that it the output view is not restricted as returning only unary and binary predicates [Volz et al., 03a]. Noy and Musen do not include an example of a query language, but their approach is implemented as a Protégé plug-in with a graphical interface that facilitates the task of formulating the queries on the one hand, and on the other hand, makes the approach implementation monolithically built in the editor, restricting the implementation reuse.

4.5 Implementation of C-VISTA model

C-VISTA model was implemented in C++ above the conceptual graph platform COGITO (Haemmerlé, 1995) that was extended by second-order concept and relation types, as required by C-VISTA: an environment for multiple viewpoint management, and in particular, methods enabling to create a viewpoint template, create the corresponding generic viewpoint, create a viewpoint from the list of specific criteria, add a viewpoint to a list of viewpoints managed in the environment, establish a viewpoint between two concepts, establish a bridge (i.e. a link) between two concepts, establish a representation link between two concepts, identify the coreference set of a referent, extract a subset of the concept hierarchy according ot a user's viewpoint.

C-VISTA was tested in road accident analysis (Ribière, 1999): the author presented an ontology on traffic accident analysis, based on the different viewpoints of seven experts (two specialists in psychology three infrastructure engineers and two vehicle engineers). All the examples in the paper were based on this application.

C-VISTA was also tested in the framework of a memory of a concurrent engineering project in aeronautics (Ribière, 1998; Ribière, 1999), with the objective of representing the artefact to be designed by several participants. For this application, a different viewpoint template, shown in Figure 17, was introduced.

Focus	DesignView: Material view Task: Building of the Brake component Step: 4
View angle	Participant: Mr X Skill field / Level: Mechanics / Expert Objective: Description Satisfied requirements: Cost reduction

Figure 17: Example of viewpoint for a concurrent engineering

C-VISTA method enables to build a multi-viewpoints ontology, with cohabitation of several possibly contradictory viewpoints, but each viewpoint itself being coherent. We must notice that, though it was presented in the framework of conceptual graph formalism, the C-VISTA model can be exploited in the framework of other formalisms. As a conclusion, using viewpoints enables a more accurate knowledge modelling from several experts and a user-oriented access to the ontology thus organized through viewpoints.

5 Specification for Next Generation Knowledge Acquisition and Modeling of Process of Consensus

In this section, a solution for knowledge acquisition and process of consensus modeling is presented. In subsection 5.1, the grounding ideas of consensus making framework and environment are provided. Further, in subsection 5.2, we provide basic details of the prototype that was implemented to support the methodologies and enable their deployment and testing in real scenarios in the future.

5.1 Abstract Specification for the Process of Consensus Support and Knowledge Acquisition

In this subsection, we present the basis of consensus framework elaborated on top of ontology classification and layering approach in the community Semantic Web portal environments, principles of consensual ontology development and editing, and, finally, personalization and community support.

5.1.1 Ontologies and Ontology Layering

We distinguish three main levels and six ontology types in the consensus framework that is suitable to be deployed on the community Semantic Web portals. The proposed classification constitutes the framework and allows introducing similar editing and storage policies for the ontologies and data that are assigned to the same level.

Levels of the community Semantic Web portal environment:

1) *User level* – user profile and personalization data specified according to ontologies of the community level.
2) *Community level* – ontologies and rules associated with a community, used and evolved by the community.
3) *Portal level* – ontologies and rules for cross-community information exchange, that also support inter-portal integration and communication.

Ontology types:

1. *User profile ontologies* – ontologies that specify the content of the portral. For example, if the main content of the community Semantic Web portal is data about people, the user profile ontology data are person's first name, last name, phone numbers, hobbies, etc.
2. *User personalization ontologies*-- ontologies that specify how user profile ontologies and user profile data are delivered to the individual user. These personalization ontologies can be Semantic Web portal specific. The personalization ontology data can also specify which user profile ontology concepts are instantiated by the user and which content and content links the user wants to share and which not.
3. *Community profile ontologies* – ontologies that specify community data, such as lists of the members of this community, their general anonymous interests and preferences.
4. *Community personalization ontologies* – ontologies that specify how and which Web portal content is delivered to a community. These personalization ontologies can be Semantic Web portal specific. The personalization ontology data can also specify which content and content links the community wants to share and which not.
5. *Portal profile ontologies* – ontologies that specify mappings and data transfer protocols across community and user ontologies. These ontologies define both mapping within ontologies (helping to reach consensus at the data level: example of a problem taken from Instone [Instone04]: "if users can specify they are interested in "PlayStation 2" but the information about the product is tagged "PS2" there will be gaps in the personalization") and also specific inter-community ontology mappings (helping to reach consensus at the metadata level: problem of the type "she uses FOAF, he uses VCard").
6. *Portal personalization ontologies* – ontologies that specify inter-portal mappings (helping to reach consensus at the physical level: problem of the type "she is on Friendster, he is on Orkut").

The six ontology types and assignment of the ontology types and instance data to the three levels of the Semantic Web portal are shown at Figure 18.

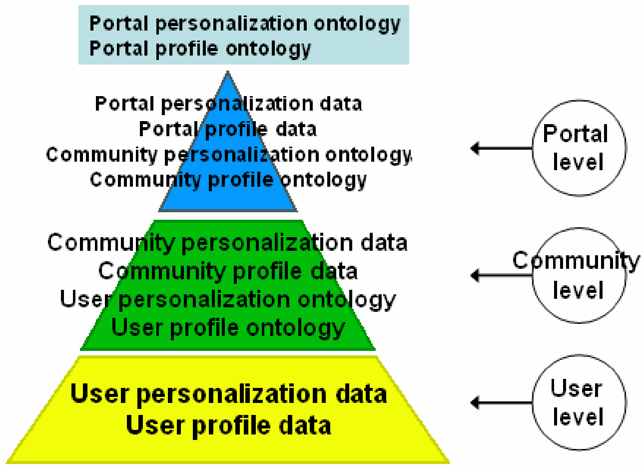

Figure 18: Layering ontologies and instance data

5.1.2 Policies for Ontology Extension and Ontology Data Editing

Here, we outline extension policies for ontologies and editing policies for instance data in the community environment. We identify operations with ontologies at three levels (at the level of an individual user, a community as a whole, and at the portal/community environment level) and two ontology types (such as ontologies specifying content or profile data, and ontologies specifying personalization data) for consensus modeling in community environments/portals. These levels and dimensions serve as a basis for the consensus framework and allow introduction of similar editing and storage policies for the ontologies and data that are assigned to the same level and type. The ontology extension and data editing policies that are enacted at the levels of individual users, communities and portals in a consensus framework are as follows:

1. *User profile ontologies*: All portal users extend profile ontologies in a by-the-way, routine manner with no interaction of central controllers and external experts. Bringing in external ontologies and bringing out ontologies constructed within the portal environment are possible.
User profile data: provided and edited by community members, individually.

2. *User personalization ontologies*: These ontologies are extended by any community member who has expertise and capability to support ontology constructs with personalization rules or services. Here and further, the user/community that has expertise and capability to support ontology constructs is the user/community that can provide functionalities for using the introduced to the other users/communities (such user with expertise and capability can be a portal creator or an external service provider). Bringing in external ontologies and bringing out ontologies constructed within the portal environment are possible.
User personalization data: provided and edited by community members, individually.

3. *Community profile ontologies*: These ontologies are extended by any community member who has expertise and capability to support ontology constructs with rules or services. Extension of these ontologies is done on the basis of user profile ontologies.
Community profile data: generated automatically by analyzing user profile data (e.g., a per cent of community members that have their own cars can be obtained as a community profile data item). Also direct introduction of the community data is possible for the cases when this data can not be received as a result of analysis of other ontologies and ontology data (e.g., the name of the community).

4. *Community personalization ontologies*: These ontologies are mainly generated automatically by adapting user personalization ontologies with specific focus on the requested content and delivery times. Also direct introduction of the community personalization ontology items is possible by anybody who has expertise and capability to support ontology constructs with personalization rules.
Community personalization data: generated automatically by analyzing user personalization data. Also direct introduction of the community data is possible for the cases when this data can not be received as a result of analysis of other ontologies and ontology data (e.g., the name of the community).

5. *Portal profile ontologies*: These ontologies are extended by anybody who has expertise and capability to support ontology constructs with application integration rules or services.
Portal profile data: provided by a community of users. For example, a member of the community can specify/confirm mappings for certain items of ontologies from the community level. After the mappings are specified, these mappings are stored as a portal profile data and can be reused by other communities. Also direct introduction of the portal data is possible for the cases when this data can not be received as a result of analysis of other ontologies and ontology data (e.g., the name of the portal).

6. *Portal personalization ontologies*: These ontologies are extended by anybody who has expertise and capability to support ontology constructs with personalization rules or services for application integration.
Portal personalization data: specified by the communities of users, specifically including portal owners.

5.1.3 Principles of Consensus Making Process

The core principles of the consensus making framework are as follows:
- New items for both content and personalization appear *only because of the efforts of individual community members* who initiate the new items.
- Ontology items *can not be deleted or modified*, they *can be supported or not* by communities. Only introduction of new ontology items is supported in the proposed consensus framework, but not deletion and modification of existing ontology items. After a community member introduces a new item, the item will exist in the system, and the

other community members do not have a possibility to delete and modify the item. Further, the members can support the initiated item by *putting an effort to comply with the initiative* (e.g., by reusing the item and including the item into the personal ontology view) or decline the new item by *ignoring* the item (i.e., not reuse and not put any efforts into the initiative). Thus, we adhere to the principle of backwards consistency in ontology development. The rationale to support the common software development principle of backward consistency in ontology development is to allow the application developers refer to the ontology items which are most appropriate for their tasks (disregarding the fact that these items might not be supported by the majority of the community).
- If two similar ontology items are found in a community environment, *a community member can map the similar items* (for example, for making use of instance data from both items). He/she can indicate the similarity of these items via an ontology mapping pointing that certain ontology items are considered to be related by a certain user. The mapping can be introduced by any user employing a standard functionality that assists to introduce the mapping in compliancy with the community or portal profile ontology. In case the mapping pattern is introduced in a community or portal profile ontologies, the whole community or all portal users can benefit from reusing the mapping.
- *Each ontology item has a measure of importance*, e.g., popularity in the community and relatedness to the community. The value of such measure can be an indicator of how many times the item was instantiated in the community (in association with community personalization ontology). According to the value of the item's importance measure of the community, a decision on how to generally treat the item is executed by personalization rules. For example, a decision on the item's placement on the screen for a default community member can be made via community personalization ontologies and rules as proposed. Apart from the community, the measure of importance of an ontology item can be adapted and applied to an individual user also (e.g., if the user has initiated the item, the item is marked as being important to him/her in association with the user personalization ontology).

5.1.4 Consensus Process Features

In addition to complying with the consensus process modeling principles, the consensus framework supports certain features, which have personalization and community support principles as cornerstones.

Features, based on by personalization:

- Personalization schemata and rules comprise separate ontology-based components and can be applied easily and interchangeably to multiple environments.
- The ontology management is offered to the broadest possible specter of community members, thus the visual ontology representations (web-forms, graphics and natural language descriptions) are the ones viewed in the portal's user interfaces and commonly shared in human-portal interaction. For the member with the basic (weak) expertise and capability of community support, ontology extension and population are downsized to provision of natural language descriptions, filling out forms, and triggering implicit personalization and

ontology instantiation (e.g., resulting from observing actual use of the environment such as calculation of item popularity measure). Meanwhile, the ontology structures and mappings introduced at the natural language and user-form level have potential to be reused at the level of machine-machine interoperation.

Features, based on community support:

- Modeling community decisions takes place at the community level. Specifically, the community members can be timely notified about community trends, e.g., on appearing new concepts or growing or decreasing popularity (i.e., importance) and support of existing concepts. According to these notifications, the members can make decisions on whether to adhere to community trends.
- Creation and evolvement of a depersonalized community member profile encapsulating personal data takes place at the community level. A depersonalized profile of a community member is necessary for acquiring data in community profile ontologies and complies with privacy support. (Privacy guarantee is necessary for obtaining accurate statistical data on sensitive issues, as applying the data on people's preferences and interests diminishes the concerns in providing the data [McCarthy01]. In addition, as indicated by Won [Won02], "…there is in general no cause for concern if information about an individual, even sensitive information, is used merely as a part of broad statistical information (e.g., the number of people in Dallas who purchased a BMW 528i in 2000…)".)
- Identification of web communities can be executed by analyzing user, community and portal data and the changes and dynamics these data undergo in the portal environment. On the basis of this analysis recommendations to join community can be delivered to portal members, and ontology items can be differently presented to different communities. An example of a rationale behind the community recommendation rule is as follows: a member, who already uses a large part of a ontology of a certain community, is likely to be interested in other ontology parts of this community, though he/she might not be a community member at that moment.
- Enhancement of implicit personalization is done at a community and portal level. Implicit personalization is an opposite of explicit personalization. Traditionally, implicit personalization is based on user behavior analysis (e.g., products purchased, pages browsed). Normally, users are turned away by explicit personalization such as need to fill in forms, subscribe to mailing lists, etc. [Instone04]. Within the proposed framework, implicit personalization can be done on the basis of community ontology-based analysis that is a step towards efficient solutions for the users with under-specified profiles.

5.1.5 Consensus Process Stepwise

In the light of different ontologies (describing users, communities, cross-platform interoperation), we specify the consensus process basing on actions of individual users and interactions across communities and platforms.

As for ontologies and policies to edit them, we subdivide actions constituting the process of reaching a consensus into the following categories:

- *Individual actions* – actions taken by individual users and having an effect on individual users only
- *Community actions* - actions taken by individual users and having an effect on more than one individual users
- *Cross-community actions* - actions taken by individual users and having an effect on more than one individual users belonging to different communities
- *Cross-platform actions* - actions taken by individual users and having an effect on more than one individual users of different environments (such as portals, platforms, communication media)

Individual actions:

- *Create* – create a new ontology or information item from scratch.
- *Create with reuse* – create a new ontology or information item employing existing ontology or information items. Discovery of and access to the reused ontology item are the necessary conditions of *create with reuse* action. *Create with reuse* also includes a simple reuse, not necessarily accompanied by creation of a considerable added value.

We say that a user belongs to a community if he/she creates with reuse an ontology or information item basing on an item reused by other (more than one) individual user(s).

Community actions:

- *Join/leave community* – joining or leaving community takes place on the basis of reuse of items created by the community. The strength of connection with a community may be represented in a range from 0 (not reusing any items assigned to the community) to 1 (reusing all items assigned to the community).

Therefore, all individual actions are directed towards weakening or intensifying connections and relations with communities.

Cross-community actions:

- establishing links between communities for gaining benefit for one community from another community and enabling interoperation of these communities

Cross-platform actions:

- establishing links between portals/platforms to benefit from interoperation of environments and enabling interoperation of these environments

We define ***consensus*** as a result of a reiterating *process* consisting of the following three steps:

1) *Creation* or *creation with reuse* an ontology or information item(s) that are estimated as highly relevant by an individual.
2) Discovery of relevance of created or created with reuse items to other individuals The discovery process consists of the following steps:
    a. Ranging communities and individuals as more and less relevant to an individual, e.g., depending on presentation of external ontology items in the individual and community profiles, dynamics and tendency in the evolution of individual and community profiles.
    b. Reception of information on individual and community actions, e.g., as a summary starting from more relevant communities and individuals to less relevant communities and individuals. Reception of information on similar actions (e.g., efforts that can bring benefit via making alignment) and complementing actions (which can influence or be influenced by actions of an individual) is of special importance for estimating relevance.
3) Returning to step (1) with estimation of relevance renewed by a discovery process.

Therefore, consensus is a result of a sequence of individual actions grounded on individual estimates of relevance based on information flows received from the community.

5.2 Implementation Support for the Consensus Framework

As mentioned in the introduction, the process of consensus making is natural to be modeled on community Semantic Web portals. The proposed here consensus principles and features are supported in the People's portal environment [Zhdanova04]. The People's portal environment is an implementation of a community Semantic Web portal infrastructure that provides ontology management facilities to the community members. Technically, the environment is built as Java servlets and Java Server Pages, employing Jena 2 [Carroll et al., 04] for manipulation with ontologies and instance data. The architecture of the People's portal environment is shown at Figure 19. A detailed description of the implementation is out of the scope of this deliverable, and more implementation details can be found in elsewhere [Zhdanova04].

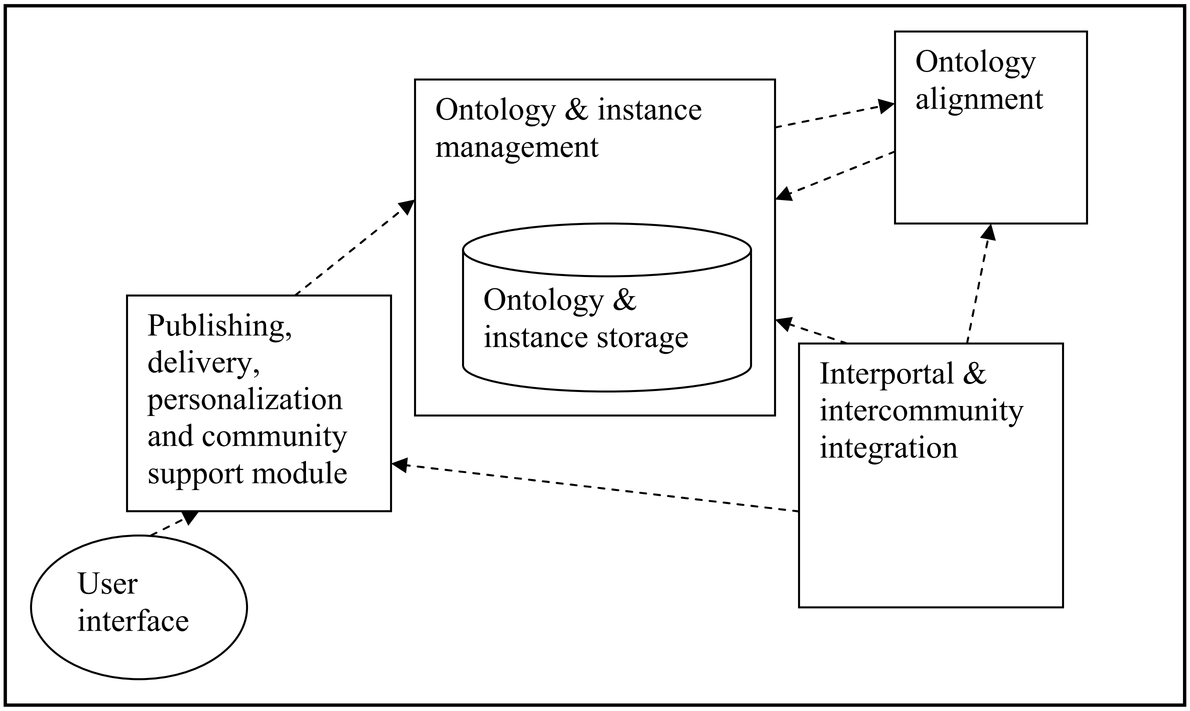

Figure 19: The People's Portal Architecture

The People's portal environment supports acquisition and exploitation of ontological structures by a community, and consensus process maintenance is an indispensable component enabling interoperation between the portal members on the basis of evolving ontologies. The application required involvement and population of domain-dependent and domain-independent ontologies, and service support for the portal's data and metadata (mostly, with publishing services for making the Semantic data human-readable). Typical types of data acquired by the People portal at different levels and dimensions of the consensus framework are listed in Table 3. Currently, the individual user personalization ontologies and data as well as portal profile ontologies and data are not exploited in the People's Portal implementation, due to absence of practical scenarios. At the moment, the information is delivered and rendered to all the community members employing the People's portal community in the same way, thus no data at individual user level and personalization dimension are involved. Further, the People's portal has not yet been involved in interoperation with other portals, therefore, no data at the portal level and profile dimension was introduced.

The ontology instantiation part of the prototype is delivered together with a simple web-based ontology editor that allows every portal member to extend the existing ontology. The importance of ontology extension functionalities on the SW portals is in allowing the community to specify what kind of content they draw to their portal and in bottom-up growth of the quantity Semantic Web pages without which the Semantic Web is impossible as was stated by initiatives such as SWAD-Europe project[9].

Level/dimension	Profile Dimension	Personalization Dimension
Individual User Level	- personal data - links to community items	n/a
Community Level	- ontological concepts and instances - ontological relations	- publishing markup for ontology items (e.g., “stable”/”unstable” markup [10]) - importance indicators for ontology items (number of views, modifications)
Portal Level	n/a	- versioning support (time markup)

Table 3: Data at Different Levels and Dimensions on the People's Portal

The idea of having certain real-life actions (e.g., publishing new instances at the portal) taken place immediately after community members make changes cannot be applied to any case studies due to restrictions, e.g., on what can be published on the portal (censorship, organizational policies). However, in any use cases, the approach can be used and be helpful nevertheless - by letting the users to extend the existing ontology, we learn more about user's interests and receive additional instance and ontology data that can indeed be included (probably after some conversions) in the next "stable"/publishable ontology and data versions.

A view on how ontology extension editing functionality can be incorporated in user forms is presented at Figure 20. These views are generated directly from OWL and RDFS ontologies and their instance data.

[9] SWAD-Europe project, URL: http://www.w3.org/2001/sw/Europe
[10] Term Status schema, URL: http://www.w3.org/2003/06/sw-vocab-status/ns

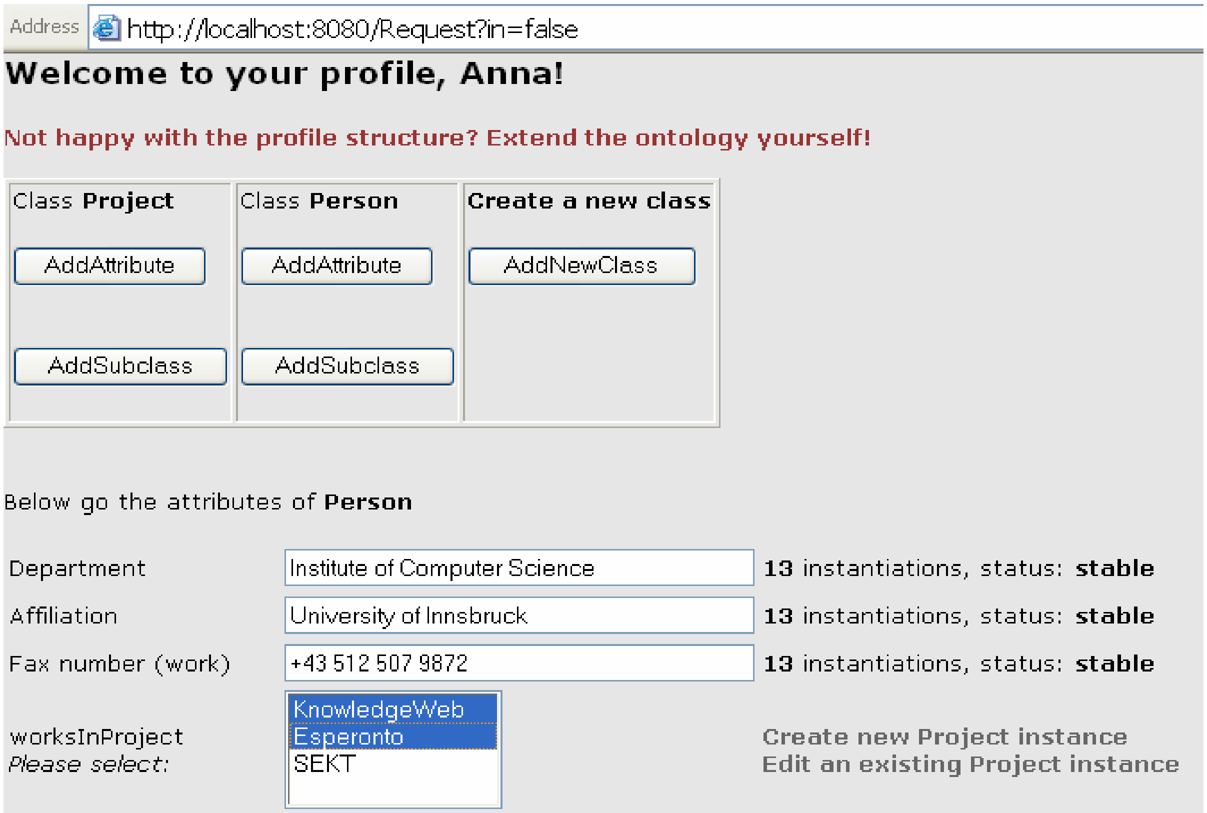

Figure 20: Outlet to Knowledge Acquisition in a Community Environment

At Figure 20, possibility to relate "class to class" and "class to literal" and an outlet to introduction of community instance data is shown.
Here, it is possible to collect "complex" information. An example of such information at the instance level is data of who works in which project, in which working group, on which topic, etc. Community-related instances introduced with this environment are versioned and can be reused in a consensual way. In addition, at the schema level, environment users are enabled to create and reuse wider range of schemata, i.e., perform ontology acquisition further.

For community interoperation support in a consensus making process, INRIA API [Euzenat, 04] was chosen to be applied in the ontology alignment solution for implementation. The choice stems up from the survey of the ontology alignment methods and tools that revealed that most considered methodologies did not gain a reusable implementation and often even vanished without any trace of implementation and INRIA API has clear advantages comparing to other existing tools [Zhdanova et al., 04].

The resulting application containing runs on a Tomcat server. The application has three major outside modules as a core: INRIA API, OWL API and Jena 2. A JSP interface to make the application available for the final user and to realize the semi-automatic matching process was implemented.

All the mappings that are verified by a human via the implementation are stored in an OWL serialization in a publicly available place: http://align.deri.org:8080/people/mappings.owl. Therefore, usage and experiment with the online version of ontology alignment implementation will result in generation of human-verified data on matched ontology items that can be reused by Semantic Web applications.
The implementation is available for public testing and use at the URL: http://align.deri.org

The implementation is based on INRIA ontology alignment API and allows to

- select two ontologies for alignment via providing their web URIs or indicating a file, containing an ontology, on the local hard drive, select alignment method among the inbuilt methods and instantiate the alignment process
- browse through the proposals of the algorithm for ontology alignment and choose the acceptable ones
- save the chosen ontology mappings in common repository available on the web for everyone's reuse and receive an output containing the just chosen mappings in an OWL serialization

A screenshot of the user interfaces for the online ontology alignment tool is shown at Figure 21. Totally, the semi-automatic alignment process consists of three stages:

1) ontology selection (by inputting URI or a file from the local disk)
2) verification of the proposed ontology mapping suggestions (Figure 21)
3) generation/output and storage of the versified mappings available for reuse

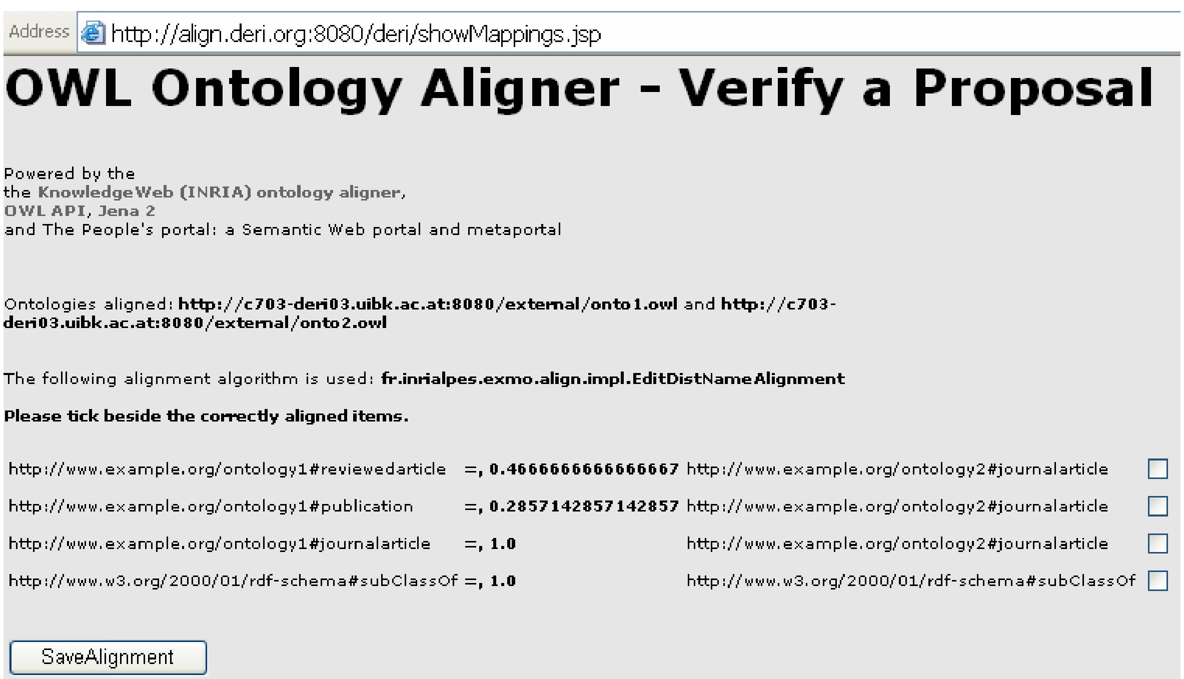

Figure 21: Verification of an Alignment Proposal

Further description and details of the interoperation and ontology alignment problem solution and plans for further work in these areas can be found in documents elsewhere [Euzenat04] [Zhdanova et al., 04].

6 Conclusion

In this deliverable, we reviewed the theories, prototypes and tools relevant for knowledge acquisition and modeling of the consensus process. The direction/framework of the consensus process modeling was described, and initial implementation infrastructure was outlined.

The future work of the activity is to deliver a prototype and report of a consensus making environment. The goal of this activity is the specification and implementation of a Semantic Web consensus making environment with a provision of dynamic and community/agents driven ontology construction, reaching agreement process support and ontology instantiation; dynamic ontology and ontology instance data alignment and aggregation; Semantic-based personalization, ontology views and targeted delivery of Semantic Web data and metadata; domain independent and domain dependent ontologies and ontology technologies widely applicable and appropriate for setting best practices on emerging Semantic Web. The Semantic Web consensus making environment will be applied to selected specific case studies such as expert environments, digital libraries and e-Tourism among B2B, B2C and C2C scenarios.

Acknowledgements

We would like to thank Francisco Martin-Recuerda (UIBK) for useful inputs and/or discussions regarding existing consensus making and collaboration tools and Katharina Siorpaes (UIBK) and Kathrin Prantner (UIBK) for help with the review of knowledge acquisition on the current Semantic Web portals.

References

[Aamodt and Plaza, 1994] Aamodt A., Plaza, E., 1994. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Communications, 7(1), pp. 39-59.

[ACE04] ACE, 2004. Annotation Guidelines for Entity Detection and Tracking (ETD). URI: http://www.ldc.upenn.edu/Projects/ACE/.

[Acker and Porter, 94] Acker, L., Porter, B., 1994. Extracting Viewpoints from Knowledge Bases. Proc. of AAAI'94.

[Agarwal et al., 03] Agarwal, S., Gonzalez, J., Hartmann, J., Hollfelder, S., Jameson, A., Klink, S., Ley, M., Rabbidge, E., Schwarzkopf, E., Shrestha, N., Stojanovic, N., Studer, R., Stumme, G., Walter, B., Weber, A., Lehti, P., Fankhauser, P., 2003. Semantic Methods and Tools for Information Portals. In Informatik03 - Jahrestagung der Gesellschaft für Informatik, pp. 116-131.

[Aggarwal et al., 02] Aggarwal, C., Philip S. Yu, P. S., 2002. An Automated System for Web Portal Personalization. VLDB 2002, Hong Kong, China, pp. 1031-1040.

[Alexaki et al. 2000] Alexaki, S., Christophides, V., Karvounarakis, G., Plexousakis, D., Tolle, K., 2000. The RDFSuite: Managing Voluminous RDF Description Bases. Technical Report, Institute of Computer Science FORTH, Heraklion, Greece. URL: http://www.ics.forth.gr/proj/isst/RDF/RSSDB/rdfsuite.pdf.

[Al-Khatib et al., 99] Al-Khatib, W., Day, Y.F., Ghafoor, A., Berra, P.B., 1999. “Semantic modeling and knowledge representation in multimedia databases”. IEEE Transactions on Knowledge and Data Engineering, 11(1): 64–80.

[Arkes and Blumer, 1985] Arkes, H. R., Blumer C., 1985. The psychology of sunk cost. *Organizational Behavior and Human Performance,* 35, 129-140.

[ARPA93] Advanced Research Projects Agency, 1993. *Proceedings of the Fifth Message Understanding Conference (MUC-5)*, Morgan-Kaufmann, California.

[Arthur, 1988] Arthur, W. B., 1988. Self-reinforcing mechanisms in economics. In P. W. Anderson, A. Arrow and P. Tines (eds.), The economy as an evolving complex system, pp. 9-10, Reading, MA: Addison-Wesley.

[Arthur, 1989] Arthur, W. B., 1989. Competing technologies, increasing returns and lock in by historical events. The economic journal, 99, 116-131.

[Assfalg et al., 02] Assfalg, J., Berlini, M., Del Bimbo, A., Nunziat, W., Pala, P., 2002. “Soccer Highlights Detection and Recognition using HMMs”, ICME, pp. 825-828.

[Benitez et al., 00] Benitez, A.B., Smith, J.R., Chang, S. F., 2000. “MediaNet: A Multimedia Information Network for Knowledge Representation”, IS&T/SPIE-2000, vol. 4210, Boston, MA.

[Berners-Lee et al., 01] Berners-Lee, T., Hendler, J., Lassila, O., 2001. The Semantic Web. Scientific American 284(5), pp. 34-43.

[Berry, 1996] Berry D. A., 1996. Statistics - A Bayesian Perspective. Duxbury Press, Belmont, California, 1996.

[Bikel et al., 99] Bikel, D., Schwartz, R., Weischedel, R. M., 1999. An Algorithm that Learns What's in a Name. Machine Learning, Special Issue on Natural Language Learning, 34 (1-3).

[Bishop, 1995] Bishop Ch.M., 1995. Neural Networks for Pattern Recognition. Clarendon Press, Oxford, 1995.

[Boser et al., 1992] Boser B.,Guyon I., and Vapnik V., 1992. A training algorithm for optim margin classifier. In Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory, Pages 44-52. ACM.

[Bowden et al., 96] Bowden, P. R., Halstead, P. and Rose, T.G., 1996. Extracting conceptual knowledge from text using explicit relation markers. Proc. of 9th European Knowledge Acquisition Workshop (EKAW-96), Nottingham, UK, 147-162.

[Bowen, 1987] Bowen, M. G., 1987. The escalation phenomenon reconsidered: Decision dilemmas or decision errors? Academy of Management Review, 12, 52-66.

[Bowen and Power, 1993] Bowen, M. G., Power, F. C., 1993. The moral manager: Communicative ethics and the Exxon Valdez disaster. Business Ethics Quarterly, 3, 97-115.

[Brockner, 1992] Brockner, J., 1992. The escalation of commitment to a failing course of action: Toward theoretical progress. Academy of Management Review, 17, 39-61.

[Calic and Izquierdo, 02] Calic, J., Izquierdo, E., 2002. “A Multiresolution Technique for Video Indexing and Retrieval”, IEEE Int. Conf. On Image Processing, ICIP2002, Rochester, New York, September.

[Carré and Dekker, 90] Carré, B., Dekker, L., 1990. The point of view notion for multiple inheritance. Proc. of ECOOP/OOPSLA’90, Ottawa, Canada.

[Carroll et al., 04] Carroll, J.J., Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., Wilkinson, K., 2004. Jena: Implementing the Semantic Web Recommendations. Proc. of the Thirteenth International World Wide Web Conference (WWW 2004), Alternative track papers and posters, pp.74-83.

[Chan et al., 02] Chan, S.S.M., Qing, L., Wu, Y. and Zhuang., Y., 2002. “Accommodating hybrid retrieval in a comprehensive video database management system”. IEEE Trans. on Multimedia, 4(2): 146–159, June.

[Chang and Sundaram, 00] Chang, S., Sundaram, H., 2000. “Structural and Semantic Analysis of Video”, IEEE International Conference on Multimedia and Expo (II).

[Chang02] Chang, S.F., 2002. "The holy grail of content-based media analysis". IEEE Multimedia, 9(2): 6–10, April-June.

[Chapelle et al., 99] Chapelle, O., Haffner, P., Vapnik, V., 1999. "Support Vector Machines for Histogram-Based Image Classification", IEEE Trans. on Neural Networks, Vol.10, No. 5, September.

[Chen01] Chen, W., Chang., S.F., 2001. "VISMap: an interactive image/video retrieval system using visualization and concept maps". In Proc. IEEE Int. Conf. on Image Processing, volume 3, pages 588–591.

[Chein and Mugnier, 92] Chein, M., Mugnier, M.-L., 1992. Conceptual graphs: fundamental notions. RIA, 6(4): 365-406.

[Chinchor92] Chinchor, N., 1992. MUC-4 Evaluation Metrics. In Proceedings of the Fourth Message Understanding Conference, 22-29.

[Chklovski03] Chklovski, T., 2003. LEARNER: A System for Acquiring Commonsense Knowledge by Analogy. In *Proceedings of Second International Conference on Knowledge Capture (K-CAP 2003)*. October 2003.

[Christel et al., 02] Christel, M., Hauptmann, A., Wactlar, H. and Tobun, D.Ng, 2002. "Collages T. Ng, "Collages as Dynamic Summaries for News Video", Proceedings of ACM Multimedia 2002, Juan-les-Pins, France, December 1-6.

[Cohen et al., 1972] Cohen, M. D., March, J. G., & Olsen, J. P., 1972. A garbage can model of organizational choice. Administrative Science Quarterly, 17, 1-25.

[Corcho et al, 03] Corcho, O., Gomez-Perez, A., Lopez-Cima, A., Lopez-Garcia, V., Suarez-Figueroa, M., 2003. ODESeW. Automatic Generation of Knowledge Portals for Intranets and Extranets. In: Fensel, D. et al. (Eds.), Proceedings of the Second International Semantic Web Conference; Springer, LNCS 2870, pp. 802-817.

[Cunningham et al., 02] Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V., 2002. GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. In Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics.

[Curtis92] Curtis, P., 1992. Mudding: Social Phenomena in Text-based Virtual Realities. In Proceedings of the 1992 Conference on the Directions and Implications of Advanced Computing, Berkeley. Also available as Xerox PARC technical report CSL-92-4. URL ftp://ftp.parc.xerox.com/pub/MOO/papers/DIAC92.ps.

[Daft et al. 1984] Daft, R.L. e Weick, K.E., 1984. Toward a Model of Organizations as Interpretation Systems. *Academy of Management Review,* 9 (2): pp. 284-295.

[Dasiopoulou et al., 04] Dasiopoulou, S., V. K. Papastathis, V. Mezaris, I. Kompatsiaris, M. G. Strintzis, "An Ontology Framework For Knowledge-Assisted Semantic Video Analysis and Annotation", Proc. 4th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot 2004) at the 3rd International Semantic Web Conference (ISWC 2004), November 2004, Hiroshima, Japan.

[Davidson et al., 98] Davidson, L., Kavanagh, J., Mackintosh, K., Meyer, I. and Skuce, D., 1998. Semi-Automatic Extraction of Knowledge-Rich Contexts from Corpora. Proc. of 1st Workshop on Computational Terminology (Computerm '98), COLING-ACL, Montreal, Canada, 50-56.

[Decker et al., 99] Decker, S., Erdmann, M., Fensel, D., Studer, R., 1999. “Ontobroker: Ontology based access to distributed and semi-structured information,” in *In R. Meersman et al, editor, Database Semantics: Semantic Issue in Multimedia Systems, Proceedings TC2/WG 2.6 8th Working Conference on Database Semantics (DS-8),* Rotorua, New Zealand, January.

[Diederich et al., 88] Diederich, J., Ruhmann, I., May, M., 1988. KRITON: a knowledge acquisition tool for expert systems. In B. R. Gaines and J. H. Boose (eds.), Knowledge-based Systems: Vol 2: Knowledge Acquisition Tools for Expert Systems, Academic Press, London.

[Dieng and Hug, 98a] Dieng, R., Hug, S., 1998. Comparison of "personal ontologies" represented through conceptual graphs. In H. Prade ed, Proc. of the 13th European Conference on Artificial Intelligence (ECAI'98), Wiley & Sons, p. 341-345, Brighton, UK.

[Dieng and Hug, 98b] Dieng, R., Hug, S., 1998. MULTIKAT, a Tool for Comparing Knowledge from Multiple Experts. In M. L. Mugnier, M. Chein eds, Conceptual Structures: Theory, Tools and Applications, Proc. of the 6th Int. Conference on Conceptual Structures (ICCS'98), Montpellier, Springer-Verlag, LNAI 1453.

[Dieng et al., 94a] Dieng, R., Corby, O., Labidi, S., 1994. Agent-based knowledge acquisition. In L. Steels, G. Schreiber, and W. Van de Velde, editors, A Future for Knowledge Acquisition: Proc. of the 8th European Knowledge Acquisition Workshop, EKAW'94, pp. 63-82. Springer-Verlag, Hoegaarden, Belgium, September. EKAW'94.

[Dieng et al., 94b] Dieng, R., Corby, O., and Labidi, S., 1994. Expertise conflicts in knowledge acquisition. In Proc. of the 8th International Workshop on Knowledge Acquisition for Knowledge Based Systems (KAW-94), volume 2, pages 23-1-23-19, Banff, Canada, January 30 - February 4.

[Dieng94] Dieng, R., 1994. Agent-based method for building a cooperative knowledge-based system. In Proc. of FGCS'94 Workshop on Heterogeneous Cooperative Knowledge-Bases, Tokyo, Japan, December . FGCS'94.

[Dieng et al., 96] Dieng, R., Giboin, A., Amergé, C., Corby, O., Després, S., Alpay, L., Labidi, S., Lapalut, S., 1996. Building of a Corporate Memory for Traffic Accident Analysis. Proceedings of the 10th Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'96), Banff, Canada, November p. 35.1-35.20. KAW'96.

[Dieng et al., 98] Dieng, R., Giboin, A., Amergé, C., Corby, O., Després, S., Alpay, L., Labidi, S., Lapalut, 1998. Building of a Corporate Memory for Traffic Accident Analysis. AI Magazine, Vol. 19, n.4, p. 80-100, Winter

[Dieng et al., 01] Dieng, R., Corby, O., Gandon, F., Giboin, A., Golebiowska, J., Matta, N., Ribière, M., 2001. Méthodes et outils pour la gestion des connaissances : une approche pluridisciplinaire pour le "knowledge management", 2ème édition,DUNOD, INFORMATIQUES, Série Stratégies et systèmes d'information

[Dill et al., 03] Dill, S., Eiron, N., Gibson, D., Gruhl, D., Guha, R., Jhingran, A., Kanungo, T., Rajagopalan, S., Tomkins, A., Tomlin, J. A., Zien, J. Y., 2003. SemTag and Seeker: Bootstrapping the semantic web via automated semantic annotation. In Proceedings of WWW'03.

[Domingue98] Domingue, J., 1998. Tadzebao and WebOnto: Discussing, Browsing, and Editing Ontologies on the Web. 11th Knowledge Acquisition for Knowledge-Based Systems Workshop, April 18th-23rd. Banff, Canada.

[Domingue et al., 04] Domingue, J., Dzbor, M. and Motta, E., 2004. Collaborative Semantic Web Browsing with Magpie. First European Semantic Web Symposium (ESWS 2004), Heraklion, Crete.

[Ekin et al., 03] Ekin, A., Tekalp, A.M., Mehrotra, R. , 2003. “Automatic Soccer Video Analysis and Summarization”, IEEE Trans. on Image Processing, vol.12, no.7, July.

[Erdmann et al., 00] Erdmann, M., Schnurr, H.P., Maedche A. and Staab, S., 2000. From Manual to Semi-automatic Semantic Annotation: About Ontology-based Text Annotation Tools. In P. Buitelaar and K. Hasida, editors, Proc. COLING 2000 Workshop on Semantic Annotation and Intelligent Content, Luxembourg, August 5-6.

[Euzenat95] Euzenat, J., 1995. Building consensual knowledge bases: context and architecture, in N. Mars (ed.), Towards very large knowledge bases, IOS press, Amsterdam (NL), pp143-155, 1995.

[Euzenat96] Euzenat, J., 1996. Corporate memory through Cooperative Creation of Knowledge-Based Systems and Hyper-Documents. Proceedings of the 10th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'96), Banff, Canada, November, (CA)pp(36)1-18.
URI: http://www.inrialpes.fr/sherpa/papers/euzenat96b/euzenat96b.html

[Euzenat et al., 97] J. Euzenat, C. Chemla, B. Jacq, A., 1997. Knowledge base for *D. melanogaster* gene interactions involved in pattern formation, Proc. 5th international conference on intelligent systems for molecular biology, Halkidiki (GR), pp108-119, 1997.

[Euzenat97] Euzenat, J., 1997. A protocol for building consensual and consistent repositories, Research report 3260, INRIA Rhône-Alpes, Grenoble (FR), September 1997, 46p.

[Euzenat04] Euzenat, J., 2004. An API for Ontology Alignment. In *Proc. of the International Semantic Web Conference*, Hiroshima, Japan, November 2004.

[Fakotakis and Sgarbas, 1998] Fakotakis, N. D. and Sgarbas, K. N., 1998. Machine Learning in Human Language Technology In Lecture Notes in Computer Science Georgios Paliouras Vangelis Karkaletsis Constantine D. Spyropoulos (Eds.)Pages 267-273

[Falquet and Mottaz Jiang, 00] Falquet, G. and Mottaz Jiang, C.-L., 2000. Conflict Resolution in the Collaborative Design of Terminological Knowledge Bases. In R. Dieng & O. Corby eds, Knowledge Engineering and Knowledge Management: Methods, Models and Tools, Proc. of the 12th International Conference, EKAW'2000, Springer-Verlag, LNAI 1937, Juan-les-Pins, p. 156-171.

[Farquhar et al., 97] Farquhar, A., Fikes, R., & Rice, J., 1997. The Ontolingua Server: Tool for Collaborative Ontology Construction. IJHCS, 46(6), 707 -- 728.

[Fayyad et al., 1996] Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P., 1996. From data mining to knowledge discovery: An Overview. In Advances in Knowledge Discovery and Data Mining, U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, eds., MIT Press, Cambridge, Mass., 1-36. 1996

[Flach, 1998] Flach, P. A., 1998. The logic of learning: a brief introduction to Inductive Logic Programming. In Proceedings of the CompulogNet Area Meeting on Computational Logic and Machina Learning, Pages 1-17. University of Manchester, 1998.

[Florian et al., 04] R. Florian, H. Hassan, A. Ittycheriah, H. Jing, N. Kambhatla, X. Luo, N. Nicolov, S. Roukos, 2004. A Statistical Model for Multilingual Entity Detection and Tracking. In Proc. of HLT/NAACL 2004, Boston, MA.

[Fonseca et al., 2002] Fonseca, F., M. Egenhofer, C. Davis and G. Câmara, 2002. Semantic Granularity in Ontology-Driven Geographic Information Systems. *Annals of Mathematics and Artificial Intelligence - Special Issue on Spatial and Temporal Granularity,* 36.1-2, 121-51.

[Frawley et al., 1992] Frawley W.J, Piatetsky-Shapiro G., and Matheus C. J., 1992. Knowledge discovery in databases - an overview. AI Magazine, Pages 13-57, 1992.

[Gabel et al., 04] Gabel, T., Sure, Y., Voelker, J., 2004. KAON – Ontology Management Infrastructure. SEKT deliverable D3.1.1.a. http://km.aifb.uni-karlsruhe.de/projects/sekt/index_html

[Garcia96] Garcia, C., 1996. Cooperative building of an ontology within multi-expertise framework. Proc. of the 2nd International Conference on Cooperative Systems (COOP'96), Juan-les-Pins, June, p. 435-454.

[Garigue, 2003] Garigue, R., 2003. Managing Ontological Ambiguity: Extending the Knowledge Management Framework. 2nd Annual Knowledge Summit.

[Gioia and Chittipetti, 1991] Gioia, D. A., Chittipetti, K., 1991. Sensemaking and Sensegiving in Strategic Change Initiation. Strategic Management Journal. 12, 433-448.

[Gomez et al. 2000] Gomez P.Y, Jones B., 2000. An Interpretation of Deep Structure in Organizations, *Organization Science*, Vol (11) 6, 696-708.

[Greco et al., 04] Greco, G., Greco, S., Zumpano, E., 2004. Web Communities: Models and Algorithms. World Wide Web: Internet and Web Information Systems, 7, pp. 59-82.

[Grefenstette94] Grefenstette, G., 1994. Explorations in Automatic Thesaurus Discovery, Kluwer Academic Publishers.

[Gruber93] Gruber, T. R., 1993. A translation approach to portable ontologies. Knowledge Acquisition, 5(2), pp.199-220.

[Hahn and Schnattinger , 98] Hahn, U., Schnattinger, K., 1998. Towards text knowledge engineering. In Proc 15th National Conference on Artificial Intelligence (AAAI-98), MIT Press, Menlo Park, CA, 524-531.

[Haemmerlé95] Haemmerlé, O., 1995. CoGITo: une plate-forme de développement de logiciel sur les graphes conceptuels. PhD Thesis, Université de Montpellier II.

[Han et al., 1996] Han J., Fu Y., Wang W., Chiang J., Gong W., Koperski K., Li D., Lu Y., Rajan A., Stefanovic N., Xia B., and Zaianae O. R., 1996. DBMiner: A System for Mining Knowledge in Large Relational. Proc. Int. Conf. on Data Mining and Knowledge Discovery (KDD'96)

[Hantula et al, 1999] Hantula, D. H., DeNicolis Bragger, J. L., 1999. The effects of feedback equivocality on escalation of commitment: An empirical investigation of decision dilemma theory. *Journal of Applied Social Psychology,* 29, 424-444.

[Harris68] Harris, Z.S., 1968. Mathematical Structures of Language, Wiley (Interscience), New York.

[Hatch, 1997] Hatch, M. J., 1997. Organization Theory: Modern, Symbolic, and Postmodern Perspectives. Oxford University Press.

[Haykin99] Haykin, S., 1999. "Neural Networks: A comprehensive foundation", Second Edition, Prentice Hall.

[Hearst, 1997] Hearst, M. A., 1997. Text data mining: Issues, techniques, and the relationship to information access. Presentation notes for UW/MS workshop on data mining, 1997

[Hirschman et al., 75] Hirschman, L., Grishman, R. and Sager, N., 1975. Grammatically Based Automatic Word Class Formation, Information Processing and Retrieval (11), 39-57.

[Hollink et al., 03] Hollink, L, Schreiber, A.Th., Wielemaker, J., Wielinga, B., 2003. "Semantic Annotation of Image Collections", in proc. of KCAP'03 Workshop, Florida.

[Hunter01] Hunter, J., 2001. "Adding Multimedia to the Semantic Web: Building an MPEG-7 ontology", SWWS 2001: 261-283.

[Hyvönen et al., 02] Hyvönen, E., Styrman, A., Saarela, S., 2002. "Ontology-Based Image Retrieval", in Proceedings of XML Finland 2002 Conference, Towards the semantic web and web services, HIIT Publication 2002-03, Helsinki, Finland. pp. 15-27.

[Instone04] Instone, K., 2004. An Information Architecture Perspective on Personalization. In: Karat, C.-M. et al. (Eds.), Designing Personalized User Experiences in eCommerce, the Netherlands, Kluwer.

[Jaimes et al., 03] Jaimes, A., Tseng, B.L., Smith, J.R., 2003. "Modal Keywords, Ontologies and Reasoning for Video Understanding", In Proc. Intl. Conf on Image and Video Retrieval, July 24-25, Urbana, IL, USA.

[Jones and Rehg, 98] Jones, M.J., Rehg, J.M., 1998. "Statistical Color Models with application to Skin Detection", Cambridge Research Laboratory – Technical Report Series, CRL 98/11 – December.

[Karakoulas and Semeraro, 1998] Karakoulas G., and Semeraro G., 1998. Machine Learning for Intelligent Information Access. In Lecture Notes in Computer Science G. Paliouras, V. Karkaletsis, and C.D. Spyropoulos (Eds.): ACAI '99, LNAI 2049, pp. 274-280, Springer-Verlag Berlin Heidelberg 2001

[Khan and MacLeod, 00] Khan, L., McLeod, D., 2000. “Audio structuring and Personalised Retrieval Using Ontologies”, *in proc. IEEE Advances in Digital Libraries (ADL 2000)*, pp. 116, Washington, DC, May.

[Kijak et al., 03] Kijak, E., Granier, G., Gros, P., Oisel, L., Bimbot, F., 2003. “HMM Based structuring of Tennis Videos using Visual and Audio Cues”, ICME.

[Kingston94] Kingston, J., 1994. Linking knowledge acquisition with CommonKADS knowledge representation, AIAI-TR-156, University of Edinburgh.

[Knublauch et al., 04] Knublauch, H., Fergerson, R.W., Noy, N.F., Musen, M.A., 2004. The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications. In *Proc. of the International Semantic Web Conference,* November 2004, Hiroshima, Japan, pp.229-243.

[Kodratoff, 1998] Kodratoff, Y., 1998. Comparing Machine Learning and Knowledge Discovery in DataBases: An Application to Knowledge Discovery in Texts. In Lecture Notes in Computer Science Georgios Paliouras Vangelis Karkaletsis Constantine D. Spyropoulos (Eds.) Pages 1-21

[Kompatsiaris et al., 04a] Kompatsiaris, I., Mezaris, V. and Strintzis, M.G., 2004. “Multimedia content indexing and retrieval using an object ontology”, Multimedia Content and Semantic Web Methods, Standards and Tools, Editor G. Stamou, Wiley, New York.

[Kompatsiaris et al., 04b] Kompatsiaris, I., Avrithis, Y., Hobson P. and Strintzis, M.G., 2004. "Integrating Knowledge, Semantics and Content for User-Centred Intelligent Media Services: the aceMedia Project", Proc. of Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS ’04), Lisboa, Portugal, April 21-23.

[Konstan and Riedl, 03] Konstan J.A., Riedl, J.T., 2003. Recommender Systems for the Web, in “Visualizing the Semantic Web” Geroimenko, V. and Chen, C. (Eds.), Springer, pp. 151-167.

[Kosko92] Kosko, B., 1992. “Neural networks and fuzzy systems: a dynamical approach to machine intelligence”, Prentice Hall, Englewood Cliffs.

[Kouzani03] Kouzani, A.Z., 2003. “Locating human faces within images”, Elsevier Computer Vision and Image Understanding 91, pp 247-279.

[Labidi96] Labidi, S., 1996. Supporting Inter-Experts Cooperation in Knowledge Acquisition Stage. In Proc. of the Second International Conference on Cooperative Systems (COOP'96), Juan-les-Pins, 12-14 June, p. 207-222.

[Lin and Lee, 95] Lin, C.T., Lee, C.S., 1995. “Neural fuzzy Systems: A neuro-fuzzy synergism to intelligent systems”, Prentice-Hall, Englewood Cliffs, NJ.

[Luo and Savakis, 01] Luo, J., Savakis, A., 2001. "Indoor vs Outdoor Classification of Consumer Photographs Using Low-Level and Semantic Features", In Proc. IEEE Int. Conf. on Image Processing (ICIP01).

[Luo et al., 04] Luo, J., Singhal, A., Etz, S.P., Gray, R.T., 2004. "A computational approach to determination of main subject regions in photographic images", Image and Vision Computing, Volume 22, Issue 3, 1 March, Pages 227-241.

[Maedche, 2002]. Maedche, A., 2002. Ontology Learning for the Semantic Web. Kluwer Academic Publishers, 2002.

[Maedche et al., 03] Maedche, A., Staab, S., Stojanovic, N., Studer, S., Sure, Y., 2003. SEmantic portAL - The SEAL approach. In D. Fensel, J. Hendler, H. Lieberman, W. Wahlster (eds.), Spinning the Semantic Web, pp. 317-359. MIT Press, Cambridge, MA., 2003.

[Manning and Schütze, 1999] Manning, H. and Schütze C., 1999. Foundations of Statistical Natural Language Processing, The MIT Press.

[Marcaillou et al., 93] Marcaillou, S., Coulette, B., Vo, D., 1993. An approach to viewpoint modelling. Proc. of TOOLS'93.

[Marino et al, 90] Marino, O., Rechenmann, F. and Uvietta, P., 1990. Multiple Perspectives and Classification Mechanism in Object-Oriented Representation. Proc. of ECAI'90, Stockholm, p. 425-430.

[Martin-Recuerda et al., 04] Martin-Recuerda, F., Harth, A., Decker, S., Zhdanova, A.V., Ding, Y., Stollberg, M., 2004. *Deliverable D2.1 "Report on requirements analysis and state-of-the-art"* within WP2 "Ontology Management" of the DIP project. URI: http://bscw.dip.deri.ie/bscw/bscw.cgi/0/3012

[Maynard and Ananiadou, 99] Maynard, D.G. and Ananiadou, S., 1999. A Linguistic Approach to Context Clustering, Proc.of Natural Language Processing Pacific Rim Symposium (NLPRS)}, Beijing, China, 346-351.

[Maynard et al., 02] Maynard, D., Tablan, V., Cunningham, H., Ursu, C., Saggion, H., Bontcheva, K., Wilks, Y., 2002. Architectural Elements of Language Engineering Robustness, Journal of Natural Language Engineering -- Special Issue on Robust Methods in Analysis of Natural Language Data, 8 (2/3), 257—274.

[Maynard and Cunningham, 03] Maynard, D., Cunningham, H., 2003. Multilingual Adaptations of a Reusable Information Extraction Tool, Proceedings of the Demo Sessions of EACL'03, Budapest, Hungary.

[Maynard et al., 04] Maynard, D., Yankova, M., Aswani N. and Cunningham, H., 2004. Automatic Creation and Monitoring of Semantic Metadata in a Dynamic Knowledge Portal. Proceedings of the 11th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA 2004), Varna, Bulgaria.

[McCarthy01] McCarthy, J. F., 2001. The Virtual World Gets Physical: Perspectives on Personalization. IEEE Internet Computing 5(6), pp. 48-53.

[Mentzas, 2002] Mentzas, G., 2002. *Knowledge asset management: beyond the process-centered and product-centered approaches*. London ; New York: Springer.

[Meyer01] Meyer, I., 2001. Extracting knowledgerich contexts for terminography. Recent Advances in Computational Terminology, John Benjamins, 279-302.

[Mezaris et al., 04a] Mezaris, V., Kompatsiaris, I., Boulgouris, N.V. and Strintzis, M.G., 2004. "Real-time compressed-domain spatiotemporal segmentation and ontologies for video indexing and retrieval", IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Audio and Video Analysis for Multimedia Interactive Services, vol. 14, no. 5, pp. 606-621, May.

[Mezaris et al., 04b] Mezaris, V., Kompatsiaris, I. and Strintzis, M.G., 2004. "Region-based Image Retrieval using an Object Ontology and Relevance Feedback", EURASIP Journal on Applied Signal Processing, Special Issue on Object-Based and Semantic Image and Video Analysis, vol. 2004, no. 6, pp. 886-901, June.

[Michie et al., 1994] Michie D, Spiegelhalter D. J. and Taylor C. C., 1994. Machine Learning, Neural and Statistical Classification. Editors: D. Michie, D.J. Spiegelhalter, C.C. Taylor, 1994. Pages 1-5.

[Mintzberg, 1994] Mintzberg, H., 1994. The Rise and Fall of Strategic Planning. New York, NY: The Free Press.

[Mitchell, 1997] Mitchell, T., 1997. Machine Learning. McGraw-Hill Co..

[Montemagni and Vanderwende, 93] Montemagni, S., Vanderwende, L., 1993. Structural Patterns vs. String Patterns for Extracting Semantic Information from Dictionaries, In Natural Language Processing: the PLNLP Approach, Jensen, K., Heidhorn, G.E. and Richardson, S.D.(eds.), Kluwer Academic Publishers, Dordrecht, 149-159.

[Motik et al., 02] Motik, B., Oberle, D., Staab, S., Studer, R., Volz, R., 2002. KAON SERVER Architecture. WonderWeb Deliverable D5. 2002. http://wonderweb.semanticweb.org

[Motta98] Motta, E., 1998. *Reusable Components for Knowledge Models*. PhD Thesis. Knowledge Media Institute. The Open University, UK, 1998.

[Mueller and Dieng, 00] Mueller, H. J., Dieng, R., editors, (2000), Computational Conflicts: Conflict Modeling for Distributed Intelligent Systems, Springer-Verlag, ISBN 3-540-66799-7, 242 pages.

[Mueller and Dieng, 00] Mueller, H. J., Dieng, R., 2000. On Conflicts in General and their Use in AI in Particular. In H. J. Mueller and R. Dieng Editors, Computational Conflicts: Conflict Modeling for Distributed Intelligent Systems, Springer-Verlag,

[Naphade et al., 02] Naphade, M.R., Kozintsev, I.V., Huang, T.S., 2002. "A factor graph framework for semantic video indexing". IEEE Trans. on Circuits and Systems for Video Technology, 12(1):40–52, January.

[Noy and Musen, 04] Noy, N.F., Musen, M.A., 2004. Specifying Ontology Views by Traversal. In *Proc. of the International Semantic Web Conference*, November 2004, Hiroshima, Japan, pp. 713-725.

[Oakes and Paice, 01] Oakes, M.P., Paice, C.D., 2001. Term extraction for automatic abstracting. Recent Advances in Computational Terminology, John Benjamins, 353-369.

[Palous, 2002] Palous, J., 2002. Machine Learning and Data Minino. Series of Research Reports, Report No: GL 157/02. 2002

[Pingali et al., 00] Pingali, G., Opalach, A., Jean, Y., 2000. "Ball Tracking and Virtual Replays for Innovative Tennis Broadcasts", ICPR, vol 4.

[Pinto et al., 2001] Pinto H, Han J, Pei J., Wang K., Chen Q, and Dayal U., 2001. Multi-Dimensional Sequential Pattern Mining, Proc. 2001 Int. Conf.\ on Information and Knowledge Management (CIKM'01), Atlanta, GA, Nov. 2001.

[Pinto et al., 04] Pinto, S., Staab, S., Sure, Y., Tempich, C., 2004. OntoEdit Empowering SWAP: a Case Study in Supporting DIstributed, Loosely-Controlled and evolvIng Engineering of oNTologies (DILIGENT). 1st European Semantic Web Symposium (ESWS 2004), May 2004, Heralicon, Crete, Greece. Springer, LNCS 3053, pp. 16-30.

[Potter01] Potter, S, 2001. A Survey of Knowledge Acquisition from Natural Language. AKT project report Task 1.1.2,
URL: http://www.aiai.ed.ac.uk/project/akt/work/stephenp/TMA of KAfromNL.pdf

[Quinlan, 1996] Quinlan J.R., 1986. Induction of Decision Trees. Machine Learning, Pages 1- 8, 1986.

[Reidsma et al., 03] Reidsma, D.¸ Kuper, J., Declerck, T., Saggion, H. and Cunningham, H., 2003. "Cross Document Ontology based Information Extraction for Multimedia Retrieval", Supplementary proceedings of the ICCS03, Dresden.

[Renman et al., 03] Renman, H., Rea, N.¸ Kokaram, A. , 2003. "Content-Based Analysis for Video from Snooker broadcasts", Elsevier Computer Vision and Image Understanding, pp 178.

[Ribière98] Ribière, M., 1998. Using viewpoints and CG for the representation and management of a corporate memory in concurrent engineering. In M. L. Mugnier, M. Chein eds, Conceptual Structures: Theory, Tools and Applications, Proc. of the 6th Int. Conference on Conceptual Structures (ICCS'98), Montpellier, August 10-12, p. 94-108

[Ribière99] Ribière, M., 1999. Représentation et gestion de multiples points de vue dans le formalisme des graphes conceptuels, PhD Thesis, Université de Nice - Sophia Antipolis, 19 April.

[Ribière and Dieng-Kuntz, 02] Ribière, M., Dieng-Kuntz, R., 2002. A Viewpoint Model for Cooperative Building of an Ontology, Conceptual Structures : Integration and Interfaces, Proceedings of the 10th International Conference in Conceptual Structures (ICCS'2002), Springer-Verlag, LNCS 2393, U. Priss, D. Corbett, G. Angelova (Eds.), p. 220-234, Borovetz, Bulgaria, 15-19 July.

[Russel and Norvig, 1995] Russell S., Norvig P. 1995. Artificial Intelligence: A Modern Approach. Prentice Hall, Englewood Cliffs, NJ.

[Sager72] Sager, N., 1972. Syntactic Formatting of Scientific Information, Proc. of 1972 Fall Joint Computer Conference, (41), Montvale, NJ, 791-800.

[Schiaffino and Amandi, 04] Schiaffino, S., Amandi, A., 2004. User-interface agent interaction: personalization issues. Int. J. Human-Computer Studies 60, pp. 129-148.

[Schreiber et al., 01] Schreiber, A.Th., Dubbeldam, B., Wielemaker, J., Wielinga, B. , 2001. "Ontology-Based Photo Annotation", *IEEE Intelligent Systems*, May-June.

[Shaw et al., 88] Shaw, M. L. G., Gaines, B. R., 1988. KITTEN: knowledge initiation and transfer tools for experts and novices. In B. R. Gaines and J. H. Boose (eds.), Knowledge-based Systems: Vol 2: Knowledge Acquisition Tools for Expert Systems, Academic Press, London.

[Sidenbladh et al., 00] Sidenbladh, H., Black, M.J., Fleet, D.J., 2000. "Stochastic Tracking of 3D Human Figures using 2D image motion", European Conf. on Computer Vision, ECCV.

[Sigal et al., 00] Sigal, L., Sclaroff, S., Athitsos, V., 2000. "Estimation and Prediction of Evolving Color Distributions for Skin Segmentation under Varying Illumination", IEEE Conf. on Computer Vision and Pattern Recognition (CVPR).

[Simoudis, 1996] Simoudis, E., 1996. Reality check for data mining. IEEE Expert, 11(5).

[Sommer, 1994] Sommer, E., 1994. FENDER: an approach to theory restructuring (extended abstract), in: Machine Learning: ECML-95 (N. Lavra c, S. Wrobel, Eds.), vol. 912 of LNAI, Springer-Verlag, 356--359.

[Sowa84] Sowa, J., 1984. Conceptual Graphs: Information Processing in Mind and Machine. Reading, Addison Wesley.

[Stamou and Tzafestas, 01] Stamou, G.B., Tzafestas, S.G., 2001. "Resolution of composite fuzzy relational equations based on Archimedean triangular norms", Fuzzy Sets and Systems, 120, pp. 395-407.

[Stollberg et al., 04] Stollberg, M., Lausen, H., Lara, R., Ding, Y., Sung-Kook, H., Fensel, D., 2004. Towards Semantic Web Portals. In Proceedings of WWW2004 Workshop on Application Design, Development and Implementation Issues in the Semantic Web, 18 May 2004, New York, USA.

[Stone et al., 1996] Stone, M.M., Brush. C.G., 1996. Planning in Ambiguous Contexts: The Dilemma of Meeting Needs for Commitment and Demands for Legitimacy, Strategic Management Journal 17: 633-652.

[Sure et al., 02] Sure, Y., Angele, J., Staab, S., 2002. OntoEdit: "Guiding Ontology Development by Methodology and Inferencing", Springer-Verlag.

[Sure et al., 02a] Sure, Y., Erdmann, M., Angele, J., Staab, S., Studer, R., Wenke, D., 2002. *OntoEdit: Collaborative Ontology Engineering for the Semantic Web.* In Proceedings of the First International Semantic Web Conference 2002 (ISWC 2002), June 9-12 2002, Sardinia, Italia, Springer, LNCS 2342, pp. 221-235.

[Tansley98] Tansley, R., 1998. "The Multimedia Thesaurus: An Aid for Multimedia Retrieval and Navigation", Master Thesis, Computer Science, University of Southhampton, UK.

[Tennison and Shadbolt, 98] Tennison, J., Shadbolt, N., 1998. *APECKS: A tool to support living ontologies*. In Proceedings of the 11th Knowledge Acquisition Workshop (KAW 98). Banff, Canada, April 1998.

[Thaler, 1994] Thaler, R.H., 1994. Quasi rational economics. Russell Sage Foundation.

[Thanitsukkarn et al., 98] Thanitsukkarn, T. and Finkelstein, A., 1998. A Conceptual Graph Approach to Support Multiperspective Development Environments. Proceedings of the 11th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop (KAW'98).

[Town and Sinclair, 03] Town, C., Sinclair, D., 2003. "A Self-referential Perceptual Inference Framework for Video Interpretation", In Proc. of the International Conference on Vision Systems, volume 2626, pp. 54-67.

[Troncy03] Troncy, R., 2003. "Integrating Structure and Semantics into Audio-visual Documents", ISWC 2003, LNCS 2870, pp. 566-581.

[Tsechpenakis et al., 02] Tsechpenakis, G., Akrivas, G., Andreou, G., Stamou, G. and Kollias, S.D., 2002. "Knowledge-Assisted Video Analysis and Object Detection",. In Proc. European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems (Eunite02), Algarve, Portugal, September.

[Tzouvaras et al., 03] Tzouvaras, V., Stamou, G. and Kollias, S.D., 2003. "Knowledge Refinement Using Fuzzy Compositional Neural Networks", Proceedings of International Conference of Artificial Neural Networks, Istanbul.

[Uchihashi et al., 99] Uchihashi, S., Foote, J., Girgensohn, A. and Boreczky, J., 1999. "Video Manga: Generating Semantically Meaningful Video Summaries", ACM Multimedia '99, Orlando, FL, November.

[Vapnik98] Vapnik, V., 1998. "Statistical Learning Theory", Springer-Verlag, New York.

[Volz et al., 03a] Volz, R., Oberle, D., Studer, R., 2003. Implementing Views for Light-Weight Web Ontologies. 7th International Database Engineering and Applications Symposium (IDEAS 2003), 16-18 July 2003, Hong Kong, China. IEEE Computer Society 2003, pp. 160-169.

[Volz et al., 03b] Volz, R.; Oberle, D.; Staab, S.; Motik, B., 2003. *KAON SERVER - A Semantic Web Management System.* In Alternate Track Proceedings of the Twelfth International World Wide Web Conference, WWW2003, Budapest, Hungary, 20-24 May 2003. ACM, 2003.

[Wallace et al., 03] Wallace, M., Akrivas, G., Mylonas, P., Avrithis, Y. and Kollias., S.D., 2003. "Using Context and Fuzzy Relations to Interpret Multimedia Content". In Proc. 3rd International Workshop on Content-Based Multimedia Indexing, CBMI'03, Rennes, France, September 22-24.

[Webb, 2000] Webb, G. I., 2002. Integrating Machine Learning with Knowledge Acquisition. In Leondes, C. T. (Ed.), Expert Systems, volume 3, pages 937-959. San Diego, CA: Academic Press.

[Weick, 1995] Weick, E. K., 1995. Sensemaking in Organizations. Sage Publications Inc.

[Weick, 2000] Weick, E. K., 2000. Making sense of Organizations. Blackwell.

[White and Sleeman, 1998] White S., and Sleeman D. H., 1998. Providing Advice on the Acquisition and Reuse of Knowledge Bases in Problem Solving. In Proceedings of the 11th Banff Knowledge Acquisition Workshop. SRDG Publications, Calgary, Canada. PpS-21.

[Whyte, 1986] Whyte, G., 1986. Escalating commitment to a course of action: A reinterpretation. Academy of Management Review, 11, 311-321.

[Won02] Won, K., 2002. Personalization: Definition, Status, and Challenges Ahead. J. of Object Technology (1) 1, pp. 29-40.

[Wrobel et al., 2003] Wrobel, A., Wuermli, I., Joller, J., Cheung Hui, S., 2003. Data Mining for Ontology Building, IEEE Intelligent Systems Magazine.

[Wu et al., 2002] Wu S.H, Hsu W.L., 2002. SOAT: A Semi-Automatic Domain Ontology Acquisition Tool from Chinese Corpus. In the 19th International Conference on Computational Linguistics, Howard International House and Academia Sinica, Taipei, Taiwan.

[Xie et al., 04] Xie, L., Xu, P.¸ Chang, S.F., Divakaran, A., Sun, H., 2004. "Structure Analysis of Soccer Video with Domain Knowledge and HMM", Elsevier Pattern Recognition Letters 25, pp 767-775.

[Yang01] Yang, J., 2001. "Thesaurus-Aided Approach for Image Browsing and Retrieval", *in proc. IEEE ICME 2001*, pp. 313-316, Tokyo, Japan, August.

[Yangarber et al., 00] Yangarber, R., Grishman, R., Tapanainen, P. and Huttunen, S., 2000. Unsupervised Discovery of Scenario-level Patterns for Information Extraction. In Proceedings of ANLP-NAACL'00, Seattle, WA.

[Yangarber02] Yangarber, R., 2002. Acquisition of Domain Knowledge. Proceedings of SCIE 2002, Frascati, Italy, 1-28.

[Yoshitaka et al., 94] Yoshitaka, A., Kishida, S., Hirakawa, M. and Ichikawa, T., 1994. "Knowledge-assisted content-based retrieval for multimedia databases", IEEE Multimedia.

[Yoshitaka and Ichikawa, 99] Yoshitaka, A., Ichikawa, T., 1999. "A survey on content-based retrieval for multimedia databases". IEEE Transactions on Knowledge and Data Engineering, 11(1): 81–93, Jan/Feb.

[Zhdanova04] Zhdanova, A.V., 2004. The People's Portal: Ontology Management on Community Portals. In proceedings of the 1st Workshop on Friend of a Friend, Social Networking and the Semantic Web (FOAF'2004), 1-2 September 2004, Galway, Ireland, pp. 66-74.

[Zhdanova et al., 2004] Zhdanova, A.V., de Bruijn, J., Zimmermann, K., Scharffe, F., 2004. Ontology Alignment Solution v2.0. EU-IST Esperonto project deliverable (http://esperonto.semanticweb.org).