INFORMATION SOCIETY TECHNOLOGIES
(IST)
PROGRAMME

Contract for:

Shared-cost RTD

Annex 1 - "Description of Work"

Project acronym: WonderWeb
Project full title: WonderWeb: Ontology Infrastructure for the Semantic Web
Proposal/Contract no.: *IST-2001-33052*
Related to other Contract no.: *n/a*

Date of preparation of Annex 1: November 10th, 2001

Operative commencement date of contract: *see Article 2.1 of contract*

WonderWeb:
Ontology Infrastructure for the Semantic Web

Contents

EUROPEAN COMMISSION
RESEARCH DIRECTORATES
GENERAL
SHARED COST
RTD CPF FORMS

EN C 1 FP5RTD

FOR COMMISSION USE ONLY

Project Acronym (2)	WonderWeb	Project No (3)	IST-2001-33052

A2. Project Summary (20)

Objectives (maximum 1000 characters)

The main objectives of the project are: the development of a family of ontology languages that extend existing Web standards while maintaining maximum backwards compatibility; the development of the comprehensive technical infrastructure and tool support that will be required for the development and deployment of ontologies; the development of a set of foundational ontologies covering a wide range of application domains, each providing a carefully crafted taxonomic backbone with a sound high level structure that can be used as the basis for the development of more detailed domain ontologies; the development of a framework of techniques and methodologies that will provide an engineering approach to the building and use of ontologies, dealing in particular with integration, migration, reconciliation and sharing.

Description of work (maximum 2000 characters)

The project objectives will be realised in the following work packages:

1. Language Architecture: A layered architecture of ontology languages will be developed based on existing Web standards. Members of the project consortium are already active in this area, and will cultivate existing contacts with non-European initiatives and the W3C, with a view to agreeing new Web standards.
2. Tools and Services: A technical infrastructure will be developed, with a component-based, extensible plug-in architecture providing persistent storage, reasoning services, versioning and APIs for ontology engineering and applications. To this will be added a suite of client tools supporting tasks such as ontology design and maintenance, extraction from legacy resources and integration.
3. Foundational Ontologies: The toolkit will be supplemented with a set of foundational ontologies, covering a wide range of application domains, that can be used as a basis both for constructing more detailed domain ontologies and for integrating existing ontologies.
4. Ontology Engineering: A framework of techniques and methodologies will be developed that provide an engineering approach to the building and use of ontologies, addressing such issues as modularisation, versioning, adaptation and reuse. Particular emphasis will be given to solving the problems associated with the integration, migration, reconciliation and sharing of ontologies.
5. Requirements Analysis and Assessment: Close links will be maintained with companies involved in the development of real life Semantic Web applications, both in order to inform the design and development process and to assess results.

Milestones and expected results (maximum 500 characters)

A requirements analysis and technology assessment. The development of a family of ontology languages. Contributions to the development of new Web standards and participation in standardisation efforts. Technical infrastructure, tools, and methodologies to support ontological engineering. A library of foundational ontologies. An evaluation and market report.

2 Objectives

The Context

We are on the brink of a new generation of World Wide Web (WWW) which, in his recent book ***Weaving the Web***, Tim Berners-Lee calls the ***Semantic Web***. Unlike the existing WWW, where data content is primarily intended for human consumption, the Semantic Web will provide data whose content is also machine processable. This will enable a wide range of intelligent services such as information brokers, search agents, information filters etc., a process that Berners-Lee describes as "Bringing the Web to its full potential". The importance of research in this area is indicated by the recently announced DAML initiative in the USA, under whose aegis projects aimed at developing the Semantic Web will receive DARPA funding totalling $70 million.[1]

The development of ***ontologies*** *will be central to this effort*. Ontologies are metadata schemas, providing a controlled vocabulary of terms, each with an explicitly defined and machine processable semantics. By defining shared and common domain theories, ontologies help both people and machines to communicate more effectively. They will therefore have a crucial rôle in enabling content-based access, interoperability and communication across the Web, providing it with a qualitatively new level of service: the Semantic Web. Examples of the use of ontologies to support content-based access and interoperability can already be seen in e.g., the American SHOE project[2] (in which HTML is being extended with ontology based semantic markup codes), the ONTOBROKER project[3] (in which ontologies are used to annotate and wrap Web documents and provides a query answering service) and the European IST-project ON-TO-KNOWLEDGE [4] (in which ontologies are being used to facilitate access to large intranets).

The Semantic Web will also be crucial to the development of Web applications such as e-commerce, providing users with much more sophisticated searching and browsing capabilities as well as support from intelligent agents such as shopbots (shopping "robots" that access vendor Web sites, compare prices etc.). Examples of the use of ontologies/taxonomies to support searching and browsing can already be seen at e.g., Yahoo Shopping (`http://shopping.yahoo.com/`) and amazon.com (`http://www.amazon.com`). Ontologies will also be important in B2B scenarios, for example in helping to solve the interoperability problem between companies wishing to exchange products.

The Aim

The importance of ontologies to the Semantic Web has prompted the development of *schema* extensions to existing Web standard languages: XML has been extended to give XML-Schema (XMLS), while RDF has been extended to give RDF-Schema (RDFS).[5] However, the language primitives provided by these standards are extremely basic when compared with those typically provided by ontology languages developed within the Knowledge Representation (KR) community, and efforts are already underway to develop *ontology* extensions of these standards. *The aim of the proposed project is to develop the infrastructure required for the large-scale deployment of ontologies as the foundation for the Semantic Web.* This will involve not only the establishment of Web standard ontology languages, but also the parallel development of the ontological engineering technology that will be required in order to "bring the web to its full potential".

[1] `http://www.daml.org/`
[2] `http://www.cs.umd.edu/projects/plus/SHOE/`
[3] `http://ontobroker.aifb.uni-karlsruhe.de`
[4] `http://www.ontoknowledge.org`
[5] `http://www.w3.org/`

The Consortium

It is ***essential to the future competitiveness of the EU*** in the emerging Web economy that European interests are represented in the development of new Web standards and technologies. Members of the project consortium are ideally placed to provide this representation: *they are already leading the effort to develop Web standard ontology languages, working in cooperation with the USA DAML initiative*[6] *and the W3C consortium*,[7] they are also leading developers of Web tools and technologies and they have been leading the development of the OIL Ontology Inference Layer,[8] an XML- and RDF-based ontology representation layer, developed in the IST-project ON-TO-KNOWLEDGE, which is already having a major impact on the American DAML programme. They are also leading contributors to the recently established IEEE "Standard Upper Ontology" initiative.[9] Given the importance of continuing these efforts, the proposed project should be considered as a very ***high priority***.

The consortium is further strengthened by an industrial advisory board whose membership includes some of the World's leading providers and users of Web technology, ranging from large multi-nationals to small and medium enterprises specialising in leading edge Web technologies. *The importance of this project is reflected in the list of organisations that have agreed to serve on the industrial advisory board.*

Main Objectives

The main objectives of the project are:

- The development of a ***family of ontology languages*** that extend existing Web standards while maintaining maximum backwards compatibility. The resulting layered architecture will provide the necessary flexibility (standardising on a single language is unrealistic in the Web environment) while maximising interoperability. This work has already begun with the development of OIL and DAML+OIL,[10] Web based ontology languages that extend and enrich RDFS. It is anticipated that a W3C standardisation working group will soon be established in order to develop a W3C ontology language recommendation based on DAML+OIL.

- The development of the ***comprehensive technical infrastructure and tool support*** that will be required both within the project and by real world applications in the Semantic Web. In particular, an ontology server architecture will be developed in order to link new and existing components in an integrated and extensible tool suite. This will include tools for editing, integrating and extracting ontologies as well as services such as persistent storage and reasoning support. The provision of reasoning support will be crucial to the development and deployment of ontologies, and has been a key consideration in the design of OIL and DAML+OIL.

- The development of a ***set of foundational ontologies*** covering a wide range of application domains. Each of these ontologies will provide a carefully crafted taxonomic backbone with a sound high level structure that can be used as the basis for the development of more detailed domain ontologies. The integration of existing ontologies with foundational ontologies will also provide a principled mechanism for the semantic integration of ontologies.

- The development of a ***framework of techniques and methodologies*** that will provide an engineering approach to the building and use of ontologies. In particular, techniques will be developed for the integration, migration, reconciliation and sharing of ontologies, issues that will be particularly important in the development of the Semantic Web.

Taken together, these components will constitute a ***complete methodology and toolkit*** supporting Web based ontological engineering.

---

[6] `http://www.daml.org`
[7] `http://www.w3.org/2001/sw/`
[8] `http://www.ontoknowledge.org/oil/`
[9] `http://ltsc.ieee.org/suo/refs.html`
[10] `http://www.daml.org/2000/12/daml+oil-index.html`

3 Participant list

WonderWeb participants						
Partic. Role*	Partic. no.	Participant name	Participant short name	Country	Date enter project	Date exit project
C	P1	University of Manchester	VUM	UK	1	30
P	P2	Vrije Universiteit Amsterdam	VUA	NL	1	30
P	P3	CNR-LADSEB	LADSEB	I	1	30
P	P4	University of Karlsruhe	AIFB	D	1	30

* C = Coordinator
P = Principal contractor

4 Contribution to programme/key action objectives

The Information Society builds on the convergence of information, communication and networking technologies, and takes advantage of infrastructures like the Internet and the Web. The aim of the IST Programme is "to realise the benefits of the information society for Europe both by accelerating its emergence and by ensuring that the needs of individuals and enterprises are met".

WonderWeb is concerned with one of the key issues in building the technical infrastructure for the information society: the next generation ***Semantic Web***. Semantic Web technology will be crucial to the information society for Europe, and ontologies are expected to be a key factor in this technology. This point is supported by the already high levels of activity by the World Wide Web Consortium (Semantic Web Activity) and the US Research community (DAML programme). It is ***essential to the future competitiveness of the EU*** in the emerging web economy that European interests are represented in the development of new Web standards and technologies.

The simple technology of the current web has contributed to its success, but it could also hamper further development: it already causes bottlenecks that hinder searching, extracting, maintaining, and generating information. Currently, computers are only used to post and render information; they do not have access to the actual content, and thus can only offer limited support in accessing and processing this information. Therefore, the main burden not only of accessing and processing information, but also of extracting and interpreting it, is placed on the human user. The Semantic Web technologies developed in WonderWeb will significantly improve this situation and therefore will help to achieve the goals of the IST program.

WonderWeb is a long-term research proposal with highly visionary elements, aiming for a Web based on machine-processable semantics. Therefore, we have submitted our proposal to ***Key Action VI.1.1 Future and Emerging Technologies*** (FET): "This specific activity on future and emerging technologies covers research that is of a longer term nature or involves particularly high risks—compensated by the promise of major advances and the potential for industrial and societal impact." (*Workprogramme 2001, page 70.*)

- The research on machine-processable semantics for data, enabling a new level of service for the web, is of longer term nature and is therefore relatively high risk;
- The new level of service in terms of information access and information processing promises major advances and the potential for industrial and societal impact in areas such as knowledge management and electronic commerce.

In this context, our proposal also significantly relates to Key Action II.1.4, "Exploratory high risk/long term research".

Other related key actions

Application areas of our technology are knowledge management and Electronic Commerce. Here, our project relates to issues raised by Key Actions II "New Methods of Work and Electronic Commerce".

Our project relates strongly to topics which will form part of the 2001 Action Line ***Semantic Web Technologies*** proposed under Area 4 (Information Access, Filtering, Analysis and Handling) of Key Action III (Multimedia Content and Tools) of the European Union's IST Programme.

In our context is also Key Action IV.3., asking for methods and tools for intelligence and knowledge sharing relying on sharable and reusable knowledge ontologies.

Our work on web language standardization also relates to Key action line VIII.1.3, "Channelling of Standardisation and Interoperability initiatives".

Finally, the thematic network OntoWeb,[11] (funded by Key action III), in which members of the consortium already run a SIG on Web language standardization, will provide a dissemination and awareness channel for WonderWeb.

[11] `http://www.ontoweb.org/`

5 Innovation

The large-scale deployment of ontologies will be fundamental to the development of the Semantic Web. *WonderWeb* will develop and demonstrate a range of innovative technologies that will provide the infrastructure necessary to support such a deployment. These technologies will encompass language architecture, tools and services, foundational ontologies and engineering methodologies.

Language Architecture

In order for ontologies to fulfil their rôle in the semantic integration of the Web, there will need to be some standardisation of Web ontology languages. As we have seen, the W3C is already moving in this direction with languages such as RDFS. However, in order to achieve the widest possible acceptability, these languages have deliberately been kept very simple and have relatively weak semantics. Much richer ontology specification languages will be needed in order support the design, sharing and integration of the complex ontologies that will be required in order to exploit the full potential of the Semantic Web. However, given the heterogeneity of the Web, it is unrealistic to suppose that a single language could satisfy all requirements or be accepted as a standard—efforts in this direction have largely been a failure even in the KR community (e.g., Ontolingua[12]) as it has proved impossible to persuade a sufficiently wide range of applications to commit to a single language standard.

A key idea behind *WonderWeb* is the development of a layered architecture of ontology languages that build on existing Web standards, extending them in a variety of directions while still maximising backwards compatibility. The bottom layer of this architecture will be RDFS. However, it should be possible to define an ontology specification layer on top of RDFS that provides the additional primitives used in most ontologies while still being general enough to be common to all extensions.

This approach has many advantages:

- Expressiveness is not limited a priori: the ontology layer can be extended in a number of directions (e.g., full first order logic expressiveness as provided by KIF[13]) in order to satisfy the requirements of different domains and applications.
- Ontologies can be shared by a wide range of applications with varying degrees of commitment to the language architecture: an application can use just that part of the ontology that is expressed within the language layer to which it is committed, and even applications that make no commitment beyond the RDFS standard can still make direct use of the class taxonomy information.
- The well defined semantics can be exploited by intelligent tools to provide reasoning support for ontological engineering (e.g., consistency checking, automatic taxonomy construction) and by web applications to leverage metadata using the rich content of deployed ontologies (e.g., by recognising instances of classes).

With this in mind, members of the consortium have already designed the OIL language (`http://www.ontoknowledge.org`), which extends RDFS with the basic ontological primitives found in typical frame languages, while at the same time endowing it with a well defined semantics. This language has now been adopted by the DAML programme (`http://www.daml.org/`), a USA DARPA funded research programme aiming at Semantic Web technology. The DAML programme has an $70M budget and runs for 3 years starting autumn 2000. Members of the *WonderWeb* consortium have been closely collaborating with the DAML programme to develop the DAML+OIL web-ontology language (`http://www.daml.org/2000/12/daml+oil-index.html`) which has now been adopted as the basis for the entire DAML programme. A joint EU/US committee on agent markup languages has also been established in order to coordinate research efforts in Europe and the USA (in particular, the DAML programme).[14] Several members of the *WonderWeb* consortium serve on this

[12] `http://ontolingua.stanford.edu/`
[13] Knowledge Interchange Format. See `http://meta2.stanford.edu/kif/kif.html`
[14] `http://www.daml.org/committee/`

committee. A key component of WonderWeb will be the continued development of the DAML+OIL ontology language layer, and members of the consortium will work closely with the OntoWeb thematic network to coordinate participation in the anticipated W3C ontology language standardisation working group.

Tools and Services

Realizing the Next-Generation-Web will require the development of a comprehensive technical infrastructure and extensive tool support. This infrastructure will be required both to support research activities in the WonderWeb project and to enable real world applications in the Semantic Web. Wonder Web will will satisfy these requirements by:

- Developing ONTOSERVER, a component-based, extensible plug-in architecture that will serve as the main organisational unit and infrastructure kernel.
- Connecting and adapting existing clients and developing new clients that hook up to ONTOSERVER to provide a range of services including support for ontology development and maintenance, migration, sharing and integration.

The design of ONTOSERVER will be based on our experiences with different semantic portals, our experiences of ontology-based knowledge management and our experiences of ontology structures for ontology learning. The design will also take into account language development efforts (WP1), ontological engineering methodologies (WP4), experiences gained when connecting clients (WP2), and input from members of the industrial advisory board (WP5).

ONTOSERVER will be tightly connected to several core components: an ontology and instance repository for persistent storage based on RDF as a foundation for the other services, inference engines for offering reasoning services, and application programmer interfaces (API's) for ontology engineering (e.g., versioning, maintenance, migration and integration) and Semantic Web applications. The challenge of this task is the provision of capable, but open interfaces that allow the welding together of components that already exist (e.g. Guha's RDFDB, Melnik's RDF-API, etc.), are to be developed in WonderWeb, or will be provided by the wider World Wide Web community. The aim is to be able to integrate ONTOSERVER into the OntoWeb portal[15] such that other members of OntoWeb will be able to provide their tools for use by the Semantic Web community.

The utility of ONTOSERVER's component-based architecture will be demonstrated by connecting and adapting several existing client-based services. These will include an RDF triple engine, one or more ontology development environments (e.g., OntoEdit[16] and OilEd[17]), the SiLRi inference engine,[18] and the FaCT inference engine.[19]

As well as adapting and connecting existing components, WonderWeb will also extend the range of available services by developing innovative new ONTOSERVER clients. These will include:

- A client that extracts ontologies from legacy resources. This client will help to address the emerging question of how the current web is to be transformed into the Semantic Web. A huge effort has already been invested in the development of schema structures, such as XML-DTDs, XML-Schema, and relational database schemata. The new client will ease the transition to the Semantic Web by semi-automatically extracting light ontologies from such legacy resources.
- An inference engine that will provide full reasoning support for the ontology language layer. Existing inference engines will not provide complete reasoning support for the ontology language layer. In particular, they do not support reasoning with XML Schema datatypes[20] (e.g., integers,

[15] http://www.ontoweb.org
[16] http://ontoserver.aifb.uni-karlsruhe.de/ontoedit/
[17] http://ontoserver.aifb.uni-karlsruhe.de/ontoedit/
[18] http://www.aifb.uni-karlsruhe.de/~sde/rdf/
[19] http://www.cs.man.ac.uk/~horrocks/FaCT/
[20] http://www.w3.org/2000/10/XMLSchema

and strings) and with extensionally defined classes (e.g., the class of primary colours defined as the set {red, orange, yellow, green, blue, indigo, violet}). Theoretical work has already been completed that shows how reasoning with datatypes and existential classes can be integrated with existing inference engines. We will use these results to develop a new inference engine based on FaCT that will support reasoning with all of the language primitives in the ontology layer.

Foundational Ontologies

Experience in software and database engineering tells us that the refinement and extension of existing designs is likely to be much more rapid and less error prone than the development of wholly new designs. The same will surely be true for the development of ontologies. Moreover, the sharing and integration of ontologies will be facilitated if they conform to a coherent high level design.

The ONTOSERVER client toolkit will therefore be supplemented with a set of "minimal foundational ontologies" (roughly of the size of a few hundred concepts) covering a wide range of application domains. Each of these ontologies will provide a carefully crafted taxonomic backbone with a sound high level structure that can be used as the basis for:

- the development of more detailed domain ontologies;
- the semantic integration and harmonisation of existing ontologies and metadata standards.

The integration of existing ontologies with foundational ontologies will also provide a principled mechanism for the semantic integration work in WP4.

Formal links and mappings will be maintained between the various foundational ontologies, in order to guarantee a coherent framework. Altogether, these foundational ontologies will constitute the core of an upper-level reference library intended to influence (and exploit) existing standardization initiatives in the field. The need of a standardized upper level ontology is well recognized by the scientific and industrial community, as testified by the recent IEEE Standard Upper Ontology initiative,[21] and by other initiatives aiming at content harmonisation of metadata standards.[22] These efforts build on a number of existing large ontological/linguistic resources, like WordNet, CYC, and others.[23] Rather than duplicating these approaches, which attempt at establishing a monolithic upper-level ontology, we shall leverage on them, carefully isolating the fundamental ontological options and their formal relationships and making them available to the users in a modular way, keeping our own ontological commitment to the very minimum.

A document addressing the methodological guidelines for the design of well-founded general ontologies will also be produced. This document will focus on the ontological choices to be made, independently of the representation language adopted, and will inform the engineering methodologies developed in WP4 as well as the development/acquisition of industrial ontologies in WP5.2. Together with the description of a selected set of foundational ontologies, such a document will hopefully result in a publication on ontology design that would complement existing books on knowledge engineering methodology, such as CommonKads. The urgent need of such a publication is widely recognized by the ontology community.

The foundational ontologies will be formulated in the formal language(s) developed in WP1.

Sharing and Integration of Ontologies

Developing ontologies is a costly and time-consuming activity, and this may prove a major barrier to the development of the Semantic Web. *WonderWeb* will address this problem by developing a framework of techniques and methodologies that provide an engineering approach to the building and

[21] `http://ltsc.ieee.org/suo/`

[22] `http://www.indecs.org/`, `http://www.mdcinfo.org/`, `www.ilrt.bris.ac.uk/discovery/harmony`, `http://www.iso18876.org/`

[23] See `http://ltsc.ieee.org/suo/refs.html` for a partial list.

use of ontologies. This framework will feed into the design of ONTOSERVER and the client toolkit, and when combined with the foundational ontologies will facilitate the rapid development of high quality Web based ontologies.

The sharing, integration, reconciliation and maintenance of ontologies are important issues in the development of such a framework. Problems associated with these tasks include:

- ontologies that use different representation languages (languages may differ in expressiveness, semantics, representational constructs and syntax);
- ontologies that use different paradigms or modelling styles;
- ontologies that encode (possibly conflicting) knowledge about overlapping domains;
- ontologies that represent the same knowledge at different levels of granularity;
- ontologies that need to be refined and extended for new tasks and domains; and
- ontologies that evolve over time;

Experience in software engineering shows that large programs can only be produced based on versioning, modularization, adaptation and reuse support. *WonderWeb* will address the above issues by developing these methodologies for use in ontology development, and will demonstrate their utility in the *WonderWeb* ontology engineering environment.

Versioning and cooperative development Large ontologies are often developed by several persons and continue to evolve over time. Both the development and the maintenance of ontologies require advanced versioning methods and guidelines for cooperative development. *WonderWeb* will develop a versioning and development framework for ontologies that will be supported by the ONTOSERVER tools and services. The framework will provide methods to disambiguate the interpretation of concepts for users of ontology revisions, and it will make explicit the compatibility of various revisions and their impact on conforming data.

Modularization Ontological modules provide a mechanism for packaging coherent sets of concepts, relations and axioms, and a means for reusing these sets in new environments. To enable the reuse of such modules, a description of the competence of their components is necessary, e.g., language choices and commitments to paradigms and modelling styles. Modularization also requires the specification of mechanisms to construct new ontologies from modules, e.g., inclusion mechanisms, mapping rules and redefinition methods. As an ontology may be built up from other ontologies written in different representation languages, the characterization of modelling primitives in languages may also be necessary. In *WonderWeb* we will develop mechanisms to *define* and to *describe* ontology modules and to *construct* new ontologies from such modules. Criteria for effective modularization will also be provided.

Adaption and reuse Building large ontologies from scratch is ineffective and costly. *WonderWeb* will therefore develop mechanisms for reusing existing ontologies and deriving new ontologies by refinement and modification. In particular, we will develop techniques for constructing domain and task-specific ontologies by refining generic ontologies (such as the foundational ontologies developed in WP3), and by restructuring ontologies (such as those acquired from WP5). Integration mechanisms for ontologies that describe the same domain to a different level of detail will also be developed. This work will start with a detailed requirements analysis undertaken in cooperation with members of the industrial advisory board, and will build on the methodological guidelines developed in WP3.

6 Community added value and contribution to EU policies

The Semantic Web is the next generation Web—the next generation ubiquitous distributed information environment. It will support a vast pool of machine-processable metadata, an extensive range of automated services dependent on these metadata, and intelligent agents able to exploit these metadata. Examples of services and agents include electronic commerce, digital libraries and knowledge management, both within enterprises (e.g., corporate memory) and community wide (e.g. e-science).

The increasing importance of the emerging Web economy makes it essential to the future competitiveness of the EU that European interests are represented in the development of new Web standards and technologies. The Semantic Web crucially depends on ontologies and ontology infrastructure to provide the semantics that it carries.

One of the objectives of EU policies is to strengthen the competitiveness of European industry in the global economy (e.g., Council Decision of 25 June, 1996). WonderWeb significantly contributes to this objective generally by improving business effectiveness through harnessing corporate knowledge assets, and specifically by increasing EU competitiveness in e-commerce and enterprise integration. It does this by fundamental research into ontology technologies, methodologies and tools, and practical demonstration of that research in the deployment of the Semantic Web.

Europe has a unique opportunity to exploit ontology-based technology, which would play a leading role in the emerging area of e-commerce and knowledge-management. WonderWeb brings together the leading European research institutions and companies that are developing services exploiting the emerging technologies of the Semantic Web, especially ontologies. This is an area in which Europe in general, and this project consortium in particular, is currently preeminent. Up until recently, proposals and recommendations of the World Wide Web Consortium (W3C) that push the Semantic Web idea have been dominated by US companies and institutions. However, the European institutions' leading edge in providing formal underpinning and inference services for emerging ontology specification languages has seen them take centre stage in the development of the practical yet theoretically sound languages needed by both W3C and the DARPA Agent Markup Language program. Thus, European technology and know-how has been recognised as essential to W3C standards. WonderWeb is an essential initiative to capitalise on this moment and further strengthen the European position in world wide activities.

Leveraging the Data Grid for e-Science The Grid has received 9.8 million Euro of funds from the IST programme in January 2001. Researchers working on the Grid hope to produce a system, "which is able to cope better with the increasing flow of information in cyberspace. The high speed network would link people with collaborative tools and supercomputers, processor farms, disks, major databases and informatics." However, this is the data and computational grid alone—on top of this layer are the information grid (the management and exchange of information rather than the streaming of data) and the knowledge grid. These upper layers can be thought of as e-science knowledge assets in the same sense as corporate knowledge assets (knowledge about people, processes, past experiences, memos, reports etc.). The technologies to be developed by WonderWeb will be directly applicable to the upper layers of the Grid, and will intersect with Grid activities by delivering the required technologies just when the scientific community will recognise their value.

Increasing EU competitiveness in E-Commerce At the same time, competitiveness of European companies will not only depend on how they access their knowledge, but also on the products and services they offer. The growth of a wide range of e-commerce services, both to individuals and between businesses, and the demands of users of these services, are contributing to the increasing international reality of information access and trading of products and services. Suppliers and information providers want to reach as many different customers as possible, independently of their nationality, whilst many end users prefer to be offered a huge range of products and services at lower cost of time and in their own language. The ability to find, interrogate and exchange knowledge is fundamental for B2B and B2C e-commerce. In COM(97)0157 the Commission recognises the importance of e-commerce and set as an objective was "to promote the technology and infrastructure needed to put in place". The advertisement, discovery and negotiation of services requires a shared understanding and exchange of metadata, explicitly represented in a machine processable way. *WonderWeb* will provide the technologies required for the promotion of e-commerce in Europe. Ontologies will play a crucial role in enabling content-based access, interoperability, multilinguality, integration information and communication.

Establishing critical mass in human and financial terms *WonderWeb* includes a high profile partner (Stanford) and major industrial collaborators (e.g., Boeing, GlaxoSmithKline, Lucent, Sun) in the USA, giving Europe the opportunity to influence standardisation activities and industrial take-up in ontology content and languages. Moreover, it gives the European community the framework by which to influence the strategic direction of the W3C vision of a Semantic Web, and to propose the appropriate technologies and their deployment to realise the vision. Europe has a lead in the combination of theory and practice of ontology-based solutions and *WonderWeb* will capitalise on and consolidate that lead.

The participants range across European countries, combining skills in ontology theory, tool development, knowledge representation systems, information integration, information retrieval, knowledge acquisition, and theorem proving. The industrialists on the industrial advisory board are from a range of application domains including drug discovery, bio-medicine, e-commerce, and manufacturing. This brings together a critical mass in technical and financial terms that will enable the successful realisation of the project. The mix of industrial and academic participants gives the project an appropriate balance of theory and practice.

WonderWeb directly responds to the following EC policy challenges

- Helping in the aim of creating a user-friendly information society by providing non-discriminatory access to services by individuals and businesses. It will help in the finding and extraction of information from an exponentially growing Semantic Web.
- Stimulating the dissemination of results and applications in all areas, with special emphasis in Web based applications, e-commerce, knowledge management and information integration.
- The scientific community and companies will clearly benefit from this globalisation of services, as having a better access to ongoing research, resources and available technology.
- The work envisaged by *WonderWeb* must be undertaken at a European level, as the skills in ontological engineering, knowledge management, tool development and e-commerce are widely dispersed in European laboratories and companies.
- *WonderWeb* also provides significant contributions to EU policies towards SMEs. European SMEs will be able to exploit emerging ontology-based technologies for automated services from the very beginning, and will therefore gain competitive advantage in these very fast developing markets. New business opportunities can be generated by exploiting these techniques and new products and services can be created, developed and brought to the market place ahead of other competitors.
- *WonderWeb* will provide the technical know-how for and a demonstration of a real Semantic Web.

In summary, *WonderWeb* will be the European focal point for building an example of the Semantic Web using innovative ontology-based methods and tools. This area is considered as one of the most promising business areas for the future. *WonderWeb* will significantly enhance the competitive position of Europe in this emerging business field and will thus contribute to the overall further development of the information society.

7 Contribution to Community social objectives

WonderWeb contributes in a number of different ways to the Community social objectives, and will be a key element in achieving the "democratisation" of the Web.

1. Making online information services accessible for all.

   The increasing importance of the Web in all aspects of business and social life make it vital that it is accessible to all members of society. It is estimated that 30 million households will have internet-connected PCs in the EU, while IDC 1998 forecasts 40 million TV sets connected to the Internet by 2002. This represents a population of consumers of services that are not just commercial but also cultural and entertainment. Moreover, many services will become only accessible online as e-commerce becomes prevalent. Unless we lower the threshold of complexity for accessing these services, we are in danger of disenfranchising communities of people. The Semantic Web specifically tries to address a variety of users, ranging from skilled IT workers to the computer layperson (who is likely to be an expert in some other domain). The Semantic Web will help to hide as much as possible complex, technological issues from information users, thus widening the range of users who will be able to configure semantic information services for particular communities, schools, businesses, non-IT academic communities, etc. Thus the objectives of WonderWeb are a concrete example of the wider goal to raise the level of support on the global information infrastructure, by hiding the technological complexity of the underlying system. For example, ontologies enable searching and ranking of web documents based on content rather than the current error-prone approach of formulating complex keyword-based queries.

2. Lowering the costs for information access.

   Unless a consumer knows exactly what to buy and where in an e-commerce site, the purchasing of a product can take an unnecessarily long time. Given that Forrester Research predicts e-commerce sites of 17,000+ pages, this problem is likely to escalate. The key to improved precision when seeking appropriate information is improved understanding of the information held on the site. This requires that the information migrates from being merely readable to being readily interpretable by automated software services such as intelligent search engines, agents and filters. WonderWeb's purpose is to build and demonstrate technologies that will make accurate information seeking possible.

3. Lowering the cost of information provision.

   We will help to lower the costs for providing information that is timely, consistent and adaptable to the specific needs of different user groups. More generally, Europe's international position will highly depend on its ability to develop and use the new information infrastructure based on online information services. This area of technology development (i.e., improving services in electronic networks) is currently one of the most competitive areas, and the position of Europe in domains such as web commerce, electronic business, and knowledge management will significantly influence its labour market, social affairs, its wealth and international influence.

4. Conservation of scarce resources

   E-commerce is seen as an effective mechanism to reduce paper resources and travel congestion, and thus conserve natural resources, reduce global warming and other side effects such as improved transport safety. E-commerce is predicated on the provision and unambiguous understanding of information as provided by a Semantic Web of the kind implemented by WonderWeb.

5. Quality of life, health and safety of the citizens

   The information management technologies developed and implemented by WonderWeb are not restricted to e-commerce. They are equally relevant to e-education, e-government, e-science and e-culture; in fact any endeavour based on the publication of information by an information provider and the seeking of that information by a consumer. Thus WonderWeb will be a key concrete exemplar to the development of standards of information exchange in genomics, medicine and digital libraries and illustrated by our participants. By improving the mechanisms for describing web content through ontologies, greater gains can be made in the protection of citizens from erroneous, untrustworthy, malicious or damaging information—for example the protection of minors from pornography.

6. Supporting the livelihoods of European citizens

   The Web will exert a strong influence over the jobs and livelihoods of increasing numbers of European citizens. Europe's economic future will significantly depend on staying at the leading edge of research into and the application of intelligent solutions in the area of electronic information access. WonderWeb will make an important contribution to European technological expertise, which will be vital to jobs and to the future health of the European economy in general.

8 Economic development and S&T prospects

As well as providing an important influence over the development of new Web standards and technologies, *WonderWeb* will also develop valuable infrastructure, tool support and related methodologies for the development of large-scale ontologies. In order to ensure effective exploitation, the *Wonder Web* consortium includes Interprice Technologies, GmbH, a leading supplier of technology support for e-business and e-commerce. They will assume the main responsibility for the exploitation task, assisting the other partners in identifying relevant market segments and potential customers. This will include the provision of a report describing the requirements for industry-standard ontology engineering environments and reviewing state-of-the-art ontology editors and environments (D23), and a report on the evaluation of *WonderWeb* technologies, with an assessment of the potential market (D25).

Furthermore, the project's industrial advisory board includes some of Europe's leading providers and users of Web technology. As well as keeping the consortium informed about current industrial practice and requirements, they will provide test sites, case studies and additional exploitation pathways.

The emergence of the Internet and the "World Wide Web" (WWW) as a universal medium for the access to and processing of information, and increasingly for the conducting of business transactions, is providing tremendous opportunities to businesses. However, as is clear from the Semantic Web initiative, the exploitation of the internet as a means of conducting business will only reach its full potential when the representation of information to be exchanged between business partners becomes formalised and standardised. The development and adoption of large scale, domain specialised ontologies is a means to this end.

According to Forrester Research, online business trade will soar from $43 billion in 1998 to $1.3 trillion in 2003. According to another study by Ovum Corp. of the UK, the world market in the related areas of Knowledge Management (KM) software will undergo a similar rapid growth from $285 million in 1998 to $1.6 billion in 2002. KM services on the other hand will grow from $1.5 billion to more than $5 billion. Looking at the information industry itself—providers of information and content—the market is expected to balloon to more than $1 billion a year according to a Gartner Group estimate.

The trend in the future development of both knowledge management and online business trade is towards more knowledge driven (semantic) approaches. Thus, for example, whereas Web Portals were up to relatively recently considered central elements of the WWW, they are already being challenged by "Vortals" (Vertical pORTALS), their knowledge-enhanced descendants that address a specific sector (vertical domain) such as the chemical industry, and allow more focused retrieval of relevant information. A specific type of Vortal is an internet marketplace, where typically goods from multiple vendors within a particular market niche are made accessible to a consumer. The proliferation of Vortals and niche on-line marketplaces is expected to reach 10,000 worldwide sometime in 2002.

More information-enabled knowledge management and online business, as mentioned above, relies on the development and adaptation of ontologies. In such B2C and B2B systems, as indeed in areas such as agent based systems, there is an underlying need and trend towards increased specialisation and knowledge content and an ability to structure and process information in context specific ways. Ontologies are a central underlying technology especially as the need for electronic resources to interact and exchange information increases. In short, ontologies aid communication through the provision of a standard and agreed upon set of concepts for information exchange between communicating parties, such as business partners in an online marketplace. However, despite the growing significance of ontologies, the currently available technology for ontology development, maintenance and deployment is widely regarded as being deficient in such areas as: automated ontology construction, modularisation and reuse of ontologies, versioning of ontologies, support for multiple views of an ontology, support for multi-user ontology development, etc. It can therefore be seen that *WonderWeb* plays a key role in driving the development of those technologies required to fulfil the imminent demand for industrial strength ontology development environments.

Technology alone is, however, only part of the solution story. The ability to generate large scale ontologies that will be widely adopted requires the development of industry-focussed, academically developed and industry tested ontology engineering methodologies, based on the technology tools. This is also part of the scope of *WonderWeb* and will help enable the growth of a new service industry in using and developing ontologies.

Work on ontologies also seems to offer significant potential in terms of scientific and technological prospects. At a first level of application, ontologies are increasingly being used not only as the basis for more accurate information structuring and retrieval, but also to support new approaches to natural language (NL) understanding that are based not so much on statistical methods as in the past, but on an understanding of NL texts on a semantic level. Ontologies also offer solutions in areas where large amounts of data (sometimes only loosely structured) must be integrated and queried by users who may be domain experts, but may not understand the structure of the data resources, e.g., in bioinformatics, medical informatics etc. In this context, ontologies can be used to describe resources and to assist in query formulation.

9 Workplan

9.1 General description

The aim of the *WonderWeb* project is to develop and demonstrate the infrastructure required for the large-scale deployment of ontologies as the foundation for the Semantic Web. This will be achieved by developing:

1. A layered architecture of ontology languages that extend existing Web standards.
2. Infrastructure and tools to support Web based ontological engineering.
3. A set of foundational ontologies providing a carefully crafted taxonomic backbone that will facilitate extension and integration.
4. Ontological engineering methodologies supporting versioning, modularization, adaptation and reuse.

Taken together, these components will constitute a complete methodology and toolkit supporting Web based ontological engineering and demonstrating its practical application.

The above components will be developed in five technical workpackages, each of which will be described in detail below (a sixth workpackage deals with project management).

The five workpackages have significant interdependencies, yet are sufficiently independent that their individual success is not wholly reliant on the results of other workpackages. Moreover, the workplan has been designed so that preliminary versions of important components will be delivered early in the project, e.g., the ontology language (month 9), the ONTOSERVER architecture (month 6) and prototype (month 12), the preliminary ontology library (month 6) and the versioning framework (month 8). This will allow progress to be monitored, problems to be detected early, and contingency plans to be implemented if required.

Early delivery of the above components is made possible by the fact that members of the consortium are already leading developers of Web technologies. Each of the partners is an acknowledged world leader in their field, and the complementary areas of expertise within the group cover all relevant areas of Web technology including automated reasoning, (Web) ontology languages, Web applications, formal ontologies, knowledge engineering, conceptual modelling, ontology engineering, Web tools and systems, databases, data management, digital libraries, semi-structured data, e-business and e-commerce.

As well as generating a critical mass for research in a new and highly interdisciplinary area, the consortium will also have the necessary international standing to exert a significant influence over the development of the Semantic Web. Important cooperations have already been established between the partners and these links will be further extended in the project workpackages, facilitating the development of larger clusters, for example with projects in the IST Key Action II and III programmes.

The consortium is further strengthened by an industrial advisory board whose membership includes some of the World's leading providers and users of Web technology, ranging from large multi-nationals to small and medium enterprises specialising in Web technologies. These links with industry will ensure that the consortium is well informed about current industrial practice and requirements and will facilitate ongoing assessment and evaluation of results based on real life applications. They will also provide important dissemination and exploitation pathways for the new tools and technologies developed in the project.

9.2 Workpackage list

WonderWeb Workpackage list

Work-package No	Workpackage title	Lead contractor No	Person-months	Start month	End month	Deliv-erable No
1.1	Ontology Language	P1	9.0	0	9	1
1.2	Language Extensions	P1	18.0	6	30	2, 3
1.3	Language Standardisation	P1	1.0	0	30	4
2.1	Ontology Server	P4	12.0	0	24	5, 6, 7
2.2	Connecting Clients	P4	29.0	6	24	8, 9, 10
2.3	LiFT Client	P4	7.0	6	30	11, 12
2.4	Reasoning Engine	P1	18.0	12	30	13, 14
3.1	Ontology Roadmap	P3	13.0	0	27	15, 16
3.2	Foundational Ontologies	P3	20.0	3	24	17, 18
3.3	Content Standardisation	P3	1.0	0	30	19
4.1	Ontology Versioning	P2	7.0	0	8	20
4.2	Ontology Modularisation	P2	10.0	6	18	21
4.3	Ontology Integration	P2	10.0	12	30	22
5.1	Requirements and Assessment	P1	7.0	0	30	23
5.2	International Conference	P1	4.0	12	24	24
5.3	Technology Evaluation	P1	4.0	18	30	25
6.1	Project Management	P1	19.0	0	30	26, 27, 28, 29
	TOTAL		189.0			

In addition to the above resources, partners P1 (VUM), P2 (VUA) and P4 (AIFB) will contribute a total of 21 person months of permanent staff time to the project. This will mostly be in a managerial and supervisory capacity, directing the research being carried out in the workpackages. The following table shows how this effort will be distributed amongst the workpackages.

Work-package	Workpackage title	Person months per participant		
		P1 (VUM)	P2 (VUA)	P4 (AIFB)
1	Language Architecture	2.0	1.0	1.0
2	Tools and Services	2.0		3.0
3	Foundational Ontologies			
4	Ontology Engineering		3.0	
5	Assessment, Dissemination and Evaluation	2.0	1.0	1.0
6	Project Management	3.0	1.0	1.0
	TOTALS	9.0	6.0	6.0

9.3 Workpackage descriptions

WP1: Language Architecture

In this workpackage we will develop the layered language architecture described in Section 5. This will involve continued close cooperation with the DAML programme (e.g., via the joint EU/US committee on agent markup languages[24]) and with the W3C (e.g., via participation in the anticipated W3C standardisation process).

By providing resources for ongoing research and development, this workpackage will complement the activities of the OntoWeb thematic network.[25] The OntoWeb network will coordinate standardisation efforts and will be used to forge additional links with other projects and with industry so as to further inform the language design process and to disseminate results as widely as possible.

This workpackage will consist of 3 main parts: the design of the ontology language layer; the design of layered language extensions; participation in language standardisation efforts. These parts will now be described in detail.

WP1.1: Design of Ontology Language Layer

The objective of this part of the workpackage is to complete the design of the ontology language layer. This layer will extend existing standards (in particular RDFS), providing the basic ontological primitives that are common to most knowledge representation paradigms.

Tasks:

- Design and development of the ontology language layer. This language will support common ontological paradigms and provide a common base layer for additional language extensions. The work will be carried out in cooperation with the Joint EU/US committee on agent markup languages, which has been established to coordinate research efforts in Europe and the USA (in particular, the DAML programme).

Deliverables:

D1 *Ontology Language*: a technical report and tutorial describing the ontology language layer in detail, including syntax, semantics and ontology examples, both didactic and realistic, the latter being derived from WP3, WP4 and WP5.

WP1.2: Language Extensions

The ontology layer will be sufficient for many applications, but language extensions will be required by other applications. The requirement for an extension of the ontology layer to include some form of rules language has already been identified,[26] and even this may not be adequate for all applications and/or application domains: in some cases expressive power up to and perhaps exceeding that of first order logic may be required. The objective of this part of the workpackage is to:

- design a rule language extension, cooperating with the relevant W3C interest group which, it is anticipated, will be set up alongside the ontology layer standardisation working group;
- identify relevant extensions and develop designs that exploit the potential of the ontology and rules layers.

Tasks:

[24] `http://www.daml.org/committee/`
[25] `http://www.ontoweb.org/`
[26] See for example `http://www.dfki.uni-kl.de/ruleml/`

- Design and development of the rule language layer in cooperation with the anticipated W3C interest group.
- Work with users (including members of the industrial advisory board and of the OntoWeb network) to identify important language extensions on the basis of application/domain requirements, and design suitable language extensions layered on top of the ontology and/or rule languages.

Deliverables:

D2 *Rule Language*: a technical report and tutorial describing the rule language layer in detail, including syntax, semantics and ontology examples, both didactic and realistic, the latter being partly derived from WP3, WP4 and WP5.

D3 *Language Extensions*: a requirements report and designs for additional language extensions.

WP1.3: Language Standardisation

The objective of this part of the workpackage is to work with the OntoWeb network in order to coordinate participation in language standardisation efforts, in particular the World Wide Web Consortium (W3C) standardisation process.

Tasks:

- Participate in the World Wide Web Consortium (W3C) standardisation process for web ontology languages.
- Monitor other language standardisation efforts (e.g., ISO, IEEE), contributing and participating where relevant.

Deliverables:

D4 *Standardisation Report*: A report on the progress of the W3C standardisation process and any other relevant standards.

WP2: Tools and Services

In this workpackage we will develop the technical infrastructure and tool support that will be required by real world applications in the Semantic Web, for example to support ontology development and maintenance, migration, sharing and integration. The design of the tools and services developed in this workpackage will reflect the languages and methodologies developed in WP1 and WP4, and will also support the ontology engineering activities described in the other work packages.

The main targets of this workpackage are to design and build ONTOSERVER, the main organisational unit and infrastructure kernel, to connect and adapt existing clients and to develop new clients that hook up to ONTOSERVER in order to provide additional tools and services. This work will consist of 4 main parts: developing ONTOSERVER; connecting/adapting existing clients; developing a new client to semi-automatically extract light ontologies from legacy resources; developing a new inference engine client to support the ontology layer. These parts will now be described in detail.

WP2.1: ONTOSERVER

ONTOSERVER will be realized as a component-based, extensible plug-in architecture tightly connected to several core components: an ontology and instance repository for persistent storage based on RDF as a foundation for other services, inference engines for offering reasoning services, and application programmer interfaces (API's) for ontology engineering (e.g., versioning, maintenance, migration and integration) and Semantic Web applications.

Tasks:

- Develop a generic and extensible internal model based on the RDF data model.
- Design a generic API and a component-based architecture using this core model.
- Assess the capabilities and performance of ONTOSERVER, in particular when used with the clients developed in WP2.2, WP2.3 and WP2.4. These tests will be performed with respect to realistic applications, in particular those suggested by members of the industrial advisory board, and realistic ontologies, for example those derived from WP3, WP4 and WP5.

Deliverables:

D5 ONTOSERVER *architecture*: A report describing the ONTOSERVER architecture and API.

D6 ONTOSERVER *prototype*: A prototype of ONTOSERVER with sample plug-ins, in particular a prototype of the RDF triple engine being developed in WP2.2.

D7 ONTOSERVER *demonstrator*: A fully functional demonstration version of ONTOSERVER, including an assessment of ONTOSERVER's capabilities and performance (e.g. throughput in terms of queries with regard to different ontologies and inference engines).

WP2.2: Connecting/Adapting Existing Clients

As showcase for the suitability of the component-based architecture of ONTOSERVER, we will connect and adapt several existing client-based services. This will require adaptation and evolution of the tools following input from both WP1 and WP4, and their inclusion into the ONTOSERVER architecture.

Tasks:

- Connect an RDF triple engine to provide an ontology and instance repository as a foundation for other services.
- Connect an ontology engineering environment (e.g. OntoEdit).
- Connect the SiLRi inference engine.
- Connect the FaCT inference engine and OilEd editor.

Deliverables:

D8 *Triple Client*: a demonstration of the RDF triple engine working as an ONTOSERVER client and providing storage for different ontology libraries (e.g. for foundational ontologies from WP3 and industrial ontologies from WP5).

D9 *SiLRi and OntoEdit Clients*: a demonstration of the SiLRi inference engine and OntoEdit ontology engineering environment working as ONTOSERVER clients.

D10 *FaCT and Oiled Clients*: a demonstration of the FaCT inference engine and OilEd editor working as ONTOSERVER clients.

The final project report (D28) will also include a section on the practical experiences gained when integrating plug-ins into the architecture.

WP2.3: LiFT Client for Legacy Schemas

A huge effort has been invested in the development of schema structures, such as XML-DTD, XML-Schema, and relational database schemata. The LiFT tool will ease the transition from the current web to the Semantic Web by semi-automatically extracting light ontologies from such legacy resources.

Tasks:

- Develop prototype of LiFT that provides techniques for manual manipulation of legacy schema data.
- Integrate semi-automatic extraction techniques into prototype.
- Evaluate performance of LiFT with legacy resources. The evaluation will be performed in cooperation with members of the industrial advisory board and will exploit their legacy resources.

Deliverables:

D11 *LiFT Prototype*: First prototype of LiFT, including techniques for manual manipulation of legacy schema data and full documentation.

D12 *LiFT Demonstrator*: Demonstration version of LiFT incorporating semi-automatic extraction techniques, including an evaluation of semi-automatic techniques in comparison with human performance.

WP2.4: New Reasoning Engine

We will extend FaCT inference engine so that it supports reasoning with all of the language primitives in the ontology layer, in particular datatypes (e.g., integers, and strings) and extensionally defined classes (e.g., the class of primary colours defined as the set {red, orange, yellow, green, blue, indigo, violet}).

Tasks:

- Develop a datatype reasoner capable of deciding consistency problems for sets of datatype definitions and values.
- Integrate datatype reasoner with existing FaCT inference engine.
- Extend FaCT to deal with extensionally defined classes.

Deliverables:

D13 *Reasoner Prototype*: A prototype of the new reasoner supporting reasoning with datatypes.

D14 *Reasoner Demonstrator*: A fully functional demonstration version of the new reasoner, including an assessment of its performance with different ontologies, in particular those derived from WP3 and WP5.

WP3: Foundational Ontologies

In this workpackage we will develop a library of “minimal foundational ontologies” (roughly of the size of a few hundred concepts) that will supplement the toolkit developed in WP2. These ontologies will provide a carefully crafted taxonomic backbone with a sound high level structure that can be used as the basis for the development of more detailed domain ontologies (e.g., in WP5.1) and for the semantic integration and harmonisation of existing ontologies and metadata standards. This will improve both the speed and accuracy of ontology development as well as providing a principled mechanism for the semantic integration task addressed in WP4.

The foundational ontologies will use the layered language architecture developed in WP1, possibly integrated with KIF comments. They will be designed using the tools and services developed in WP2.

A document addressing the methodological guidelines for the design of well-founded general ontologies will also be produced. This document will focus on the ontological choices to be made, independently of the representation language adopted. Together with the description of a selected set of foundational ontologies, such document will hopefully result in a publication on ontology design that would complement existing books on knowledge engineering methodology, such as CommonKads. The urgent need for such a publication is widely recognized by the ontology community.

The work will consist of 3 main parts: the development of ontology design methodologies that will contribute both to the engineering methodologies developed in WP4 and transitioning and development of ontologies in WP5.1; the development of the top-level foundational ontologies; participation in ontology standardisation efforts. These parts will now be described in detail.

WP3.1: State of the Art, Roadmap and Methodology

The objective of this part of the workpackage is to provide a review of the current state of the art, a roadmap and a methodology for ontology design.

Tasks:

- Review the state of the art of content-oriented ontology design methodologies.
- Develop a detailed roadmap of major ontological design choices.
- Develop a comprehensive set of methodological guidelines for designing well-founded ontologies, focusing on fundamental ontological choices that are independent of the representation language adopted.

Deliverables:

D15 *Ontology Roadmap*: A report on the current state of the art in ontology design methodologies and a roadmap of major ontological design choices.

D16 *Ontology Methodologies*: A report on methodological guidelines for designing well-founded ontologies.

WP3.2: Top-level Foundational Ontologies

The objective of this part of the workpackage is the development of a library of foundational ontologies that will supplement the toolkit developed in WP2. Input from WP5.1 and from members of the industrial advisory board will be used to help determine the requirements for these ontologies.

Tasks:

- Requirements analysis (in cooperation with WP5.1 and members of the industrial advisory board)
- Development of minimal ontologies, e.g., of objects and events and of information and information processing.

Deliverables:

D17 *Ontology Library (preliminary)*: A preliminary version of the ontology library including at least a minimal ontology of objects and events.

D18 *Ontology Library (final)*: A final version of the ontology library, including ontologies of objects and events and of information and information processing.

WP3.3: Participation in Content Standardization Process

The objective of this part of the workpackage is to to work with the OntoWeb network in order to coordinate participation in content standardisation efforts, in particular the IEEE SUO and World Wide Web Consortium (W3C) activities.

Tasks:

- Participate in the IEEE SUO activities.
- Participate in W3C content standardization activities.

Deliverables:

D19 *Standardisation Report*: A report on the progress of the IEEE SUO and W3C standardisation activities.

WP4: Ontology Engineering

Ontologies play a major role in supporting information exchange processes in various areas, and may help to overcome current bottlenecks of information access. However, by providing a solution for semantically enriched information access and exchange, ontologies may also become a new bottleneck for the further development of the Semantic Web. Developing ontologies is costly and time-consuming activity, and this may cause a major barrier for their use in intelligent information integration.

Experience in software engineering shows that large programs can only be produced based on versioning, modularization, and support for adaptation and reuse. WonderWeb will address the above issues by developing these methodologies for use in ontology development. Taken together with the other workpackages, this will constitute a complete methodology and toolkit supporting web based ontological engineering. The utility of the methodology and toolkit will be demonstrated/evaluated by applying it to the development of real-world ontologies (e.g., ontologies derived from WP5).

WP4.1: Versioning and Cooperative Development

Large ontologies are often developed by several persons and continue to evolve over time. The objective of this part of the workpackage is to support such development with advanced versioning methods and guidelines for cooperative development.

Tasks:

- Develop a versioning and development framework for ontologies that will be integrated into the design of ONTOSERVER (WP2).

Deliverables:

D20 *Versioning Framework*: A report that describes the versioning framework and evaluates its utility in ontology development (e.g., in WP3 and WP5).

WP4.2: Modularization for ontologies

Ontological modules provide a mechanism for packaging coherent sets of concepts, relations and axioms and a means for reusing them in new environments.

This objective of this part of the workpackage is to provide mechanisms to *define* and to *describe* ontology modules and to *construct* new ontologies from such modules. Criteria for effective modularisation will also be provided.

Tasks:

- Develop a modularisation mechanism for ontologies.
- Develop construction mechanisms for ontology modules.
- Decompose and reconstruct a real-world ontology (e.g., derived from WP5), using the developed mechanisms.

Deliverables:

D21 *Modularisation Mechanisms*: A report that describes the modularisation and construction mechanisms and evaluates their utility in the decomposition and reconstruction of a real-world ontology (e.g., an ontology derived from WP5).

WP4.3: Refinement, Modification and Integration Mechanisms

Building large ontologies from scratch is ineffective and costly. The objective of this part of the workpackage is to provide means for reusing ontologies. This may involve both the derivation of new ontologies by refinement and modification (e.g., of generic ontologies such as those developed in WP3), and the restructuring of existing ontologies (e.g., of industrial ontologies derived from WP5). Integration mechanisms for ontologies that describe the same domain to a different level of detail will also be developed.

The work will start with a detailed requirements analysis for refinement and modification mechanisms undertaken in cooperation with members of the industrial advisory board.

Tasks:

- Analysis of the requirements for refinement, modification and integration mechanisms;
- Development of the appropriate mechanisms.
- Demonstrate the utility of these mechanisms by refining and integrating real-world ontologies (e.g., derived from WP5).

Deliverables:

D22 *Refinement Mechanisms*: A report describing the refinement, modification and integration mechanisms, and evaluates their utility in refining and integrating real-world ontologies.

WP5: Assessment, Dissemination and Evaluation

The aim of the *WonderWeb* project is to develop the infrastructure required for the large-scale deployment of ontologies as the foundation for the Semantic Web. If this aim is to be realised, it is vital to maintain close links with companies involved in the development of real-life Semantic Web applications, both in order to inform the design and development process and to assess results. This will be achieved by close cooperation between members of the consortium and the industrial advisory board to provide both information about industrial requirements and assessments based on real-life applications. Regular meetings and workshops will be organised in order to facilitate this cooperation, and to keep the advisory board informed as to the progress of the project.

Participation in project workshops will normally be restricted to members of the consortium and the industrial advisory board. In order to ensure that the results of the project are disseminated as widely as possible, a major international conference on the Semantic Web will also be organised. The conference speakers and location will be carefully selected so as to appeal to both researchers and industrialists, providing an opportunity for the results of the project to be presented to both of these groups. Professionally published conference proceedings will provide an additional dissemination pathway.

As well as providing an important influence over the development of new Web standards and technologies, WonderWeb will also develop valuable infrastructure, tool support and related methodologies for the development of large-scale ontologies. The potential for the commercial exploitation of these technologies will become clearer as the project progresses. If deemed appropriate after the first year project review, and taking into account input from the reviewers and the EU project officer, a detailed evaluation and market report will be commissioned from industry experts with the necessary skills (such experts to be selected with the prior approval of the project officer). The project budget includes 4 person-months of subcontracting to cover such a commission.

WP5.1: Industrial Requirements and Assessment

The objective of this part of the workpackage is to provide information about industrial requirements and assessments of the utility of WonderWeb technologies and methodologies in real life industrial applications. This will be achieved by close cooperation and coordination between members of the consortium and the industrial advisory board.

Tasks:

- Regular project meetings and workshops. Regular meetings and workshops will be organised in order to disseminate results to and receive feedback and assessment from the members of the industrial advisory board.
- Coordinating with members of the industrial advisory board on the development/acquisition of industrial ontologies for testing and assessment purposes.

Deliverables:

D23 *Assessment Report* A report reviewing the results of assessments of WonderWeb technologies and methodologies in real life industrial applications.

WP5.2: International Conference

The objective of this part of the workpackage is to organise a major international conference on the Semantic Web. The conference will take place towards the end of year 2 of the project so that mature results from WonderWeb can be presented. Several high profile speakers from both the industrial and research communities will be invited to the conference, and the timing and conference venue will also be carefully selected so as to appeal to both communities. The conference will also have published proceedings, both in order to attract high quality contributions from outside the WonderWeb consortium and to provide an additional dissemination pathway.

Tasks:

- Select international programme committee.
- Select attractive conference venue and appropriate timing.
- Invite high quality and high profile speakers from both industry and academia.
- Arrange for professional publication of proceedings.
- Organise and run conference.

Deliverables:

D24 *International Conference*

WP5.3: Evaluation and Market Report

The objective of this part of the workpackage is to evaluate the utility of WonderWeb technologies and methodologies in relation to competitor technologies, and to produce a market assessment. The bulk of this work will be subcontracted to appropriate industry experts.

Tasks:

- Evaluating the utility of WonderWeb technologies and methodologies by comparing them with competitor technologies.
- Assess the market for WonderWeb technologies and provide guidance to their further development in response to evolving industrial requirements.

Deliverables:

D25 *Evaluation and Market Report*: A report on the evaluation of WonderWeb technologies and an assessment of the potential market, including guidance as to directions for further development in response to evolving industrial requirements.

WP6: Project Management

The objectives of this workpackage are:

- To provide administrative management, coordination of activities, planning and control.
- To ensure that the project objectives are met and to represent the project to the Commission and the IST community.
- To produce project reports and organise project reviews.
- To coordinate the dissemination and exploitation of project results.

The project management structures and activities are described in detail in Section C5.

Tasks:

- Project management and coordination. This will be carried out by the project coordinator in cooperation with the workpackage leaders and with members of the industrial advisory board. Workpackage leaders will be responsible for the content and timely availability of project deliverables.
- Project reports and reviews. Early in the project we will produce a project presentation as an important awareness action and to disseminate the vision of the project, and a dissemination and use plan that describes plans for the dissemination and exploitation of the knowledge gained during the work. Towards the end of the project we will produce a technical implementation plan. In addition, annual project reports and a final report will be produced.
- Dissemination and exploitation of project results. Cooperation with members of the industrial application board will facilitate this task.

Deliverables:

D26 *Project Presentation* A presentation of the project to increase awareness and disseminate the vision of the project.

D27 *Dissemination and Use Plan*

D28 *Technical Implementation Plan*

D29 *Final Report*

Annual project review reports will also be produced.

One-page description of each work package

Language Architecture

Workpackage number:	1	**Start date or starting event:**	month 1

Participant:	VUM	VUA	LADSEB	AIFB
Person-months per participant:	16	6		6

Objectives

The development of a family of ontology languages that extend existing Web standards while maintaining maximum backwards compatibility. The resulting layered architecture will provide the necessary flexibility (standardising on a single language is unrealistic in the Web environment) while maximising interoperability.

Description of work

In this workpackage we will develop a layered language architecture. This will involve continued close cooperation with the DAML programme (e.g., via the joint EU/US committee on agent markup languages) and with the W3C (e.g., via participation in the anticipated W3C standardisation process). The workpackage will consist of 3 main parts:

- Complete the design of the ontology language layer.
- Develop language extensions, in particular a rule layer.
- Participate in language standardisation efforts.

Deliverables

D1 *Ontology Language*: a technical report and tutorial describing the ontology language layer.
D2 *Rule Language*: a technical report and tutorial describing the rule language layer.
D3 *Language Extensions*: a requirements report and designs for additional language extensions.
D4 *Standardisation Report*: A report on the progress of the W3C standardisation process and any other relevant standards.

Milestones and expected result

The expected result of this workpackage is a family of ontology languages.

Milestones:
MS1 Ontology language (month 9).
MS6 Rule language (month 18).

Tools and Services

Workpackage number:	2	**Start date or starting event:**	month 0	
Participant:	VUM	VUA	LADSEB	AIFB
Person-months per participant:	24			42

Objectives

To develop the technical infrastructure and tool support that will be required by real world applications in the Semantic Web, for example to support ontology development and maintenance, migration, sharing and integration. The tools and services developed in this workpackage will also support the activities described in the other work packages.

Description of work

Design and build ONTOSERVER, the main organisational unit and infrastructure kernel, connect and adapt existing clients and develop new clients that hook up to ONTOSERVER in order to provide additional tools and services. This work will consist of 4 main parts:

- developing ONTOSERVER;
- connecting/adapting existing clients;
- developing a new client to semi-automatically extract light ontologies from legacy resources;
- developing a new inference engine client to support the ontology layer.

Deliverables

D5 ONTOSERVER *architecture*: A report describing the ONTOSERVER architecture and API.
D6 ONTOSERVER *prototype*: A prototype of ONTOSERVER.
D7 ONTOSERVER *demonstrator*: A fully functional demonstration version of ONTOSERVER.
D8 *Triple Client*: a demonstration of the RDF triple engine working as an ONTOSERVER client.
D9 *SiLRi and OntoEdit Clients*: a demonstration of the SiLRi inference engine and OntoEdit ontology engineering environment.
D10 *FaCT and OilEd Clients*: a demonstration of the FaCT inference engine and OilEd clients.
D11 *LiFT Prototype*: First prototype of LiFT client, using manual extraction.
D12 *LiFT Demonstrator*: Demonstration version of LiFT including semi-automatic extraction.
D13 *Reasoner Prototype*: A prototype of the new reasoner supporting reasoning with datatypes.
D14 *Reasoner Demonstrator*: A fully functional demonstration version of the new reasoner.

Milestones and expected result

The expected result of this workpackage is an integrated collection of tools and services supporting real-world applications in the Semantic Web.

Milestones:
MS5 ONTOSERVER prototype (month 12).
MS6 Triple client (month 18).
MS7 Reasoner prototype (month 21).

Foundational Ontologies

Workpackage number:	3	**Start date or starting event:**	month 0	
Participant:	VUM	VUA	LADSEB	AIFB
Person-months per participant:		3	31	

Objectives

To develop a library of "minimal foundational ontologies" (roughly of the size of a few hundred concepts) that will supplement the toolkit developed in WP2. These ontologies will provide a carefully crafted taxonomic backbone with a sound high level structure that can be used as the basis for the development of more detailed domain ontologies and for the semantic integration and harmonisation of existing ontologies and metadata standards. This will improve both the speed and accuracy of ontology development (e.g., in WP5.2) as well as providing a principled mechanism for the semantic integration task addressed in WP4.

Description of work

The work will consist of 3 main parts:

- A review of the current state of the art of ontology design methodologies.
- The development of a library of foundational ontologies that will supplement the toolkit developed in WP2. Input from WP5.2 and from members of the industrial advisory board will be used to help determine the requirements for and utility of these ontologies.
- Participation in ontology standardisation efforts, in particular the IEEE SUO and World Wide Web Consortium (W3C) activities.

Deliverables

D15 *Ontology Roadmap*: A report on the current state of the art in ontology design methodologies and a roadmap of major ontological design choices.

D16 *Ontology Methodologies*: A report on methodological guidelines for designing well-founded ontologies.

D17 *Ontology Library (preliminary)*: A preliminary version of the ontology library including at least a minimal ontology of objects and events.

D18 *Ontology Library (final)*: A final version of the ontology library, including ontologies of objects and events and of information and information processing.

D19 *Standardisation Report*: A report on the progress of the IEEE SUO and W3C standardisation activities.

Milestones and expected result

The expected result of this workpackage is a library of "minimal foundational ontologies" that will supplement the toolkit developed in WP2.

Milestones:

MS2 Preliminary ontology library (month 6).

MS8 Final ontology library (month 24).

Ontology Engineering

Workpackage number:	4	**Start date or starting event:**	month 0	
Participant:	VUM	VUA	LADSEB	AIFB
Person-months per participant:		24	3	

Objectives

To develop methodologies for use in ontology development based on techniques from software engineering, in particular versioning, modularization, and support for adaptation and reuse. Taken together with the other workpackages, this will constitute a complete methodology and toolkit supporting web based ontological engineering.

Description of work

The work will consist of 3 main parts:

- The development of advanced versioning methods and guidelines for cooperative development.
- The development and evaluation of modularisation techniques, including the provision of mechanisms to define and to describe ontology modules and to construct new ontologies from such modules.
- The development and evaluation of techniques for reusing ontologies, including the derivation of new ontologies by refinement and modification, as well as the restructuring of existing ontologies. Integration mechanisms for ontologies that describe the same domain to a different level of detail will also be developed.

Deliverables

D20 *Versioning Framework*: A report that describes and evaluates the versioning framework.
D21 *Modularisation Mechanisms*: A report that describes the modularisation and construction mechanisms and evaluates their utility in the decomposition and reconstruction of a real-world ontology.
D22 *Refinement Mechanisms*: A report describing the refinement, modification and integration mechanisms, and evaluates their utility in refining and integrating real-world ontologies.

Milestones and expected result

The expected result of this workpackage is a set of methodologies for use in ontology development based on techniques from software engineering, and a demonstration of their utility in real-world applications.

Milestones:
MS3 Versioning framework (month 8).
MS6 Modularisation mechanisms (month 18).

Assessment, Dissemination and Evaluation

Workpackage number:	5	**Start date or starting event:**	month 0	
Participant:	VUM	VUA	LADSEB	AIFB
Person-months per participant:	9	3	3	

Objectives

To maintain close links with companies involved in the development of real-life Semantic Web applications, both in order to inform the design and development process and to assess results; to ensure that the results of the project are disseminated as widely as possible by organising a major international conference on the Semantic Web; to evaluate the potential for the commercial exploitation of WonderWeb technologies

Description of work

Assessment, dissemination and evaluation of WonderWeb technologies and methodologies.

- The organisation of regular meetings and workshops in order to disseminate results to and receive feedback and assessment from the members of the industrial advisory board.
- Coordinating with members of the industrial advisory board on the development/acquisition of industrial ontologies for testing and assessment purposes.
- The organisation of a major international conference on the Semantic Web at which Wonder Web results can be presented to an international audience of researchers and industialists.
- The evaluation of WonderWeb technologies and methodologies in relation to competitor technologies, and an assessment of the market for WonderWeb technologies.

Deliverables

D23 *Assessment Report*: A report reviewing the results of assessments of WonderWeb technologies and methodologies in real life industrial applications.

D24 *International Conference*: A major international conference on the Semantic Web.

D25 *Evaluation and Market Report*: A report on the evaluation of WonderWeb technologies and an assessment of the potential market, including guidance as to directions for further development in response to evolving industrial requirements.

Milestones and expected result

The expected results of this workpackage are regular project workshops and an assessment roport, the organisation of a major international conference on the Semantic Web, and an evaluation and market report.

Milestones:

MS9 Evaluation and market report (month 30).

Project Management

Workpackage number:	6	**Start date or starting event:**	month 0		
Participant:		VUM	VUA	LADSEB	AIFB
Person-months per participant:		12	3	2	2

Objectives

- To provide administrative management, coordination of activities, planning and control;
- To ensure that the project objectives are met and to represent the project to the Commission and the IST community.
- To produce project reports and organise project reviews.
- The dissemination and exploitation of project results.

Description of work

- Project management and coordination, carried out by the project coordinator in cooperation with the workpackage leaders and with members of the industrial advisory board.
- Project reports and reviews.

Deliverables

Annual project review reports.
D26 *Project presentation*.
D27 *Dissemination and Use Plan*.
D28 *Technical implementation plan*.
D29 *Final report*.

Milestones and expected result

The expected results of this workpackage is the successful and timely completion of project.

Milestones:
MS1 Project presentation (month 3).
MS9 Successful and timely completion of project (month 30).

9.4 Deliverables list

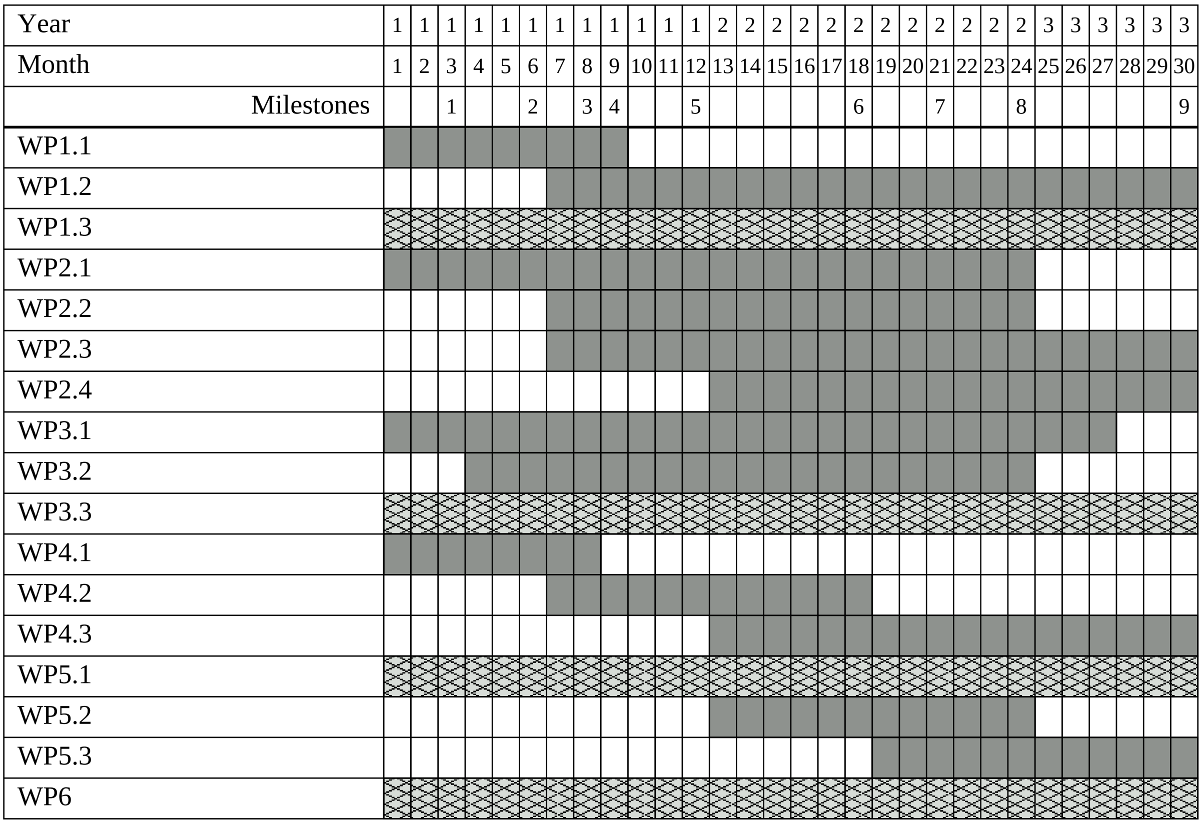
WonderWeb deliverables list

Del. no.	Deliverable name	WP no.	Lead particip-ant	Estimated person-months	Del. type	Security*	Delivery (proj. month)
1	Ontology Language	1.1	VUM	9.0	R	Pub.	9
2	Rule Language	1.2	VUM	9.0	R	Pub.	18
3	Language Extensions	1.2	VUM	9.0	R	Pub.	30
4	Standardisation Report	1.3	VUM	1.0	R	Int.	30
5	OntoServer Architecture	2.1	AIFB	3.0	R	Int.	6
6	OntoServer Prototype	2.1	AIFB	3.0	P	Pub.	12
7	OntoServer Demonstrator	2.1	AIFB	6.0	D	Pub.	24
8	Triple Client	2.2	AIFB	10.0	D	Pub.	18
9	SiLRi and OntoEdit Clients	2.2	AIFB	13.0	D	Pub.	20
10	FaCT and OilEd Clients	2.2	VUM	6.0	D	Pub.	20
11	LiFT Prototype	2.3	AIFB	3.0	P	Int.	12
12	LiFT Demonstrator	2.3	AIFB	4.0	D	Pub.	30
13	Reasoner Prototype	2.4	VUM	9.0	P	Int.	21
14	Reasoner Demonstrator	2.4	VUM	9.0	D	Pub.	30
15	Ontology Roadmap	3.1	LADSEB	6.0	R	Pub.	12
16	Ontology Methodologies	3.1	LADSEB	7.0	R	Pub.	27
17	Ontology Library (prelim.)	3.2	LADSEB	6.0	P	Int.	6
18	Ontology Library (final)	3.2	LADSEB	14.0	D	Pub.	24
19	Standardisation Report	3.3	LADSEB	1.0	R	Int.	30
20	Versioning Framework	4.1	VUA	7.0	R	Pub.	8
21	Modularisation Mechanisms	4.2	VUA	10.0	R	Pub.	18
22	Refinement Mechanisms	4.3	VUA	10.0	R	Pub.	30
23	Assessment Report	5.1	VUM	7.0	R	Int.	30
24	International Conference	5.2	VUM	4.0	O	Pub.	24
25	Evaluation and Market Report	5.3	VUM	4.0	R	Int.	30
26	Project Presentation	6.1	VUM	2.0	R	Pub.	3
27	Dissemination and Use plan	6.1	VUM	2.0	R	Int.	6
28	Technical Implementation Plan	6.1	VUM	3.0	R	Int.	30
29	Final Report	6.1	VUM	3.0	R	Int.	30

* *Int.* *Internal circulation within project (and Commission Project Officer if requested)*
Pub. *Public document*

9.5 Project planning and timetable

Gantt Chart



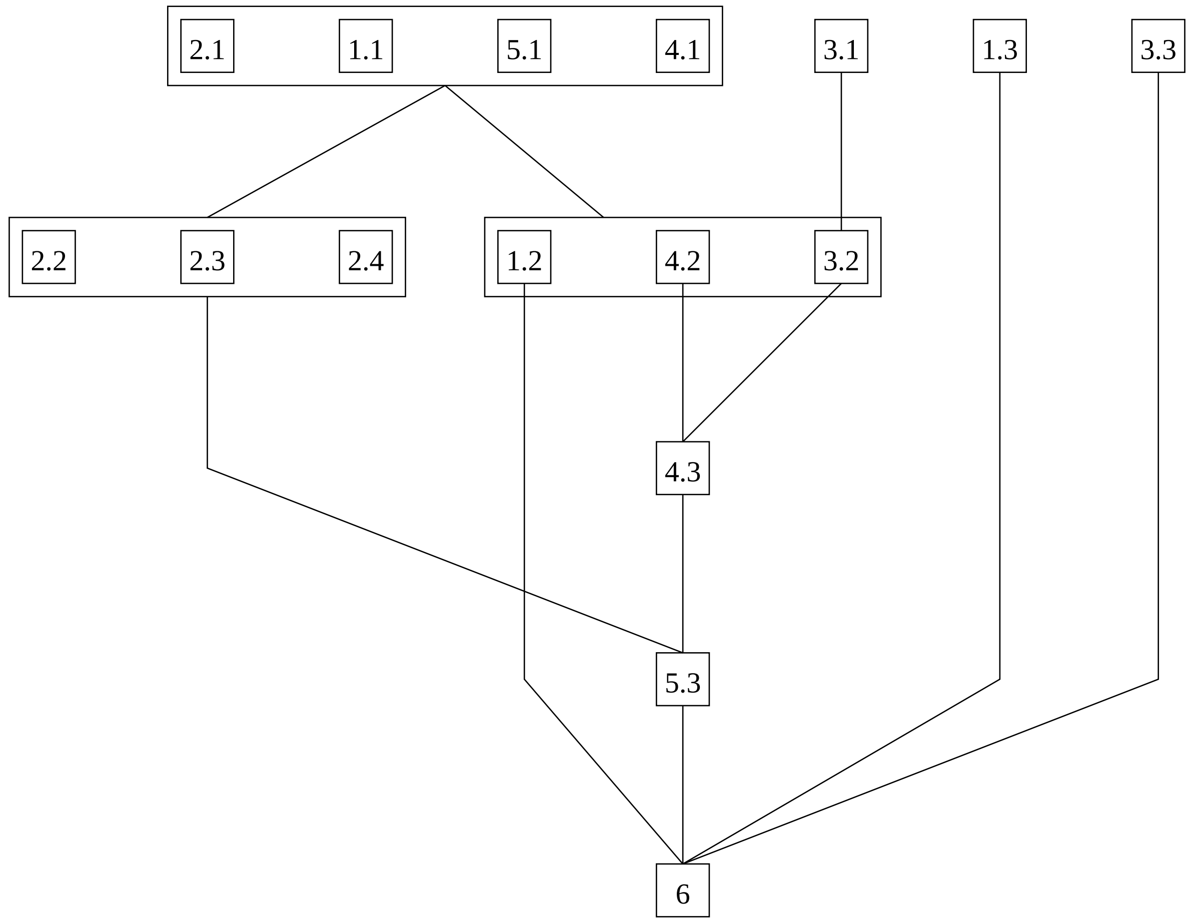
Key: = Normal activities

= Ongoing/periodic activities

The above gantt chart illustrates the project plan, showing the timing of the different WPs and their components.

9.6 Graphical representation of project components

Pert diagram



The above pert diagram illustrates the dependencies between the various tasks of the work package list. A downward line indicates results from the top-box are a necessary prerequisite for the bottom box.

9.7 Project management

The *WonderWeb* project will be lead by a Project Coordinator and a Project Management Board made up of the Project Coordinator, the workpackage leaders and members of the industrial advisory board.

Project Coordinator

The Project Coordinator will organise the timely production of deliverables and act as liaison with IST and other EU bodies as necessary. The Victoria University of Manchester (VUM) will take the role of Coordinating Partner, and the Project Coordinator will be Dr. Ian Horrocks. The Computer Science Department of VUM has the experience and resources necessary for successful project coordination. It has a long and distinguished research record and considerable experience in both national and European projects (e.g., the EC AIM programme "GALEN" and "GALEN-IN-USE" projects, the Esprit "Foundations of Data Warehouse Quality" project and the IST Thematic Network "OntoWeb"). The staff of the Computer Science Department have experience with the management of EU projects, having coordinated both the GALEN and GALEN-IN-USE projects.

In addition, Dr. Dieter Fensel of the Vrije Universiteit Amsterdam will assist the project coordinator in an advisory capacity, and will be a member of the project management board. He will bring to bear his considerable experience of EU projects (e.g. IBROW, On-to-knowledge, AgentLink, OntoWeb) and project management (e.g., On-to-knowledge, OntoWeb).

Work package Leaders

Each work package will have a designated leader from the organisation managing the task. The work package leaders will be responsible for the content and timely availability of project deliverables. Their responsibilities will consist of:

- Leading the technical development of the work-packages.
- Coordinating activities for tasks within the work packages.
- Ensuring work package timings, resources, and costs are maintained.
- Preparing presentations and demonstrations for technical reviews of milestones.

Project Management Board

The Project Management Board (PMB) is responsible for the overall success of the project. The PMB will comprise the project coordinator, the work package leaders, and selected members of the industrial advisory board. It will meet at least two times a year and will be responsible for running the project, formulating and revising its strategic objectives, and for conflict resolution. The PMB has two main functions: administrative and executive.

The PMB's administrative function will consist of:

- Preparing and distributing non-technical reports (Management Reports, Progress Reports).
- Maintaining accurate consolidated records of costs, resources, and time-scales.
- Preparing and submitting to the Commission the cost statements of all participants.
- Maintaining close contacts with the Commission and the Project Officer.
- Communicating with other projects and coordinating presentations.

The PMB's executive function will consist of:

- Coordinating the integration of results from the various workpackages in accordance with the workplan.
- Coordinating the preparation and distribution of all major technical deliverables.
- Ensuring technical consistency and maximising synergy between the workpackages.
- Reviewing project progress.
- Ensuring that the project maintains its objectives and relevance within the IST programme.
- Coordinating the exploitation of project results.
- Resolving any technical, administrative or contractual issues.

The boards will meet regularly and will be responsible for the progress of the project. The Project Management Board will select a small Executive Project Management Board (lead by the Project Coordinator) that will run daily business between overall meetings.

Exploitation

The main responsibility for exploitation will be assumed by the Coordinating Partner who, with assistance from subcontracted industry experts, will assist the other partners in identifying relevant market segments and potential customers. This is in addition to the exploitation pathways offered by members of the industrial advisory board and the academic partner's existing outward links to applied/industrial projects in the domain. Moreover, the members of the consortium are also partners in the OntoWeb network, and this will provide further exploitation pathways.

Resolution of Conflicts

The Consortium will follow a "collaborative" approach for avoiding conflicts. The consortium has already successfully used this approach during the preparation of the short proposal and in earlier cooperations. In the unlikely case that a conflict should arise, it will be dealt with by the Project Management Board, and if necessary, input from the EC Project Officer will be sought.

10 Clustering

Members of the consortium will attend workshops and conferences organised by the European Commission.

11 Other contractual conditions

- The potential for the commercial exploitation of *WonderWeb* technologies will become clearer as the project progresses. If deemed appropriate after the first year project review, and taking into account input from the reviewers and the EU project officer, a detailed evaluation and market report will be commissioned from industry experts with the necessary skills. Such experts will be selected with the prior approval of the project officer. The project budget includes a figure to cover an estimated 4 person-months of subcontracting at typical industry rates for this level of person.
- The "other specific project costs" in the budget relate solely to the provision of audit certificates as required by the commission.

Appendix A: Consortium description

The project consortium is made up of leading research groups from four European countries (the *full partners*). In addition, there will be an industrial advisory board that includes some of the World's leading providers and users of Web technology.

Full Partners The research and development work described in the work plan will be carried out by the full partners. Each of the partners is an acknowledged world leader in their field, and the complementary areas of expertise within the group cover all relevant areas of Web technology:

[P1] University of Manchester, UK (VUM) automated reasoning, ontology languages and applications (*project coordinator*).

[P2] Vrije Universiteit Amsterdam, Netherlands (VUA) web applications and ontologies.

[P3] Consiglio Nazionale delle Ricerche, Instituto di Sistemistica e Bioingegneria, Italy (LADSEB) formal ontologies, knowledge engineering and conceptual modelling.

[P4] Institute AIFB, University of Karlsruhe, Germany (AIFB) ontology engineering, web tools, systems and applications.

As well as generating a critical mass for research in a new and highly interdisciplinary area, the consortium has the necessary international standing to exert a significant influence over the development of the Semantic Web. In particular, they have direct links into the DAML programme (VUM and AIFB are DAML sub-contractors and are invited to attend DAML PI workshops) and the W3C (VUA are members of W3C, VUM's Ian Horrocks is a W3C "invited expert", and both VUM and VUA will be participating in the W3C Web Ontology Language working group which has recently been chartered).

Important cooperations have already been established between the partners (for example in the ON-TO-KNOWLEDGE project, in the OntoWeb network, in the development of the OIL language and in the DAML project). These links will be further extended in the project workpackages and will facilitate the development of larger clusters, for example with projects in the IST Key Action III programme.

Extensive exploitation pathways will be offered by members of the industrial advisory board and the academic partner's existing outward links to applied/industrial projects in the domain. Further advice and assistance in exploiting the results of the project will be provided by experts from industry working as subcontractors to the coordinating partner (should this be deemed appropriate by the EU project officer).

Industrial advisory board Members of the board will attend regular workshops and meetings during the course of the project, the purpose of which will be:

- to keep the consortium informed about current industrial practice and requirements;
- to review the progress of the project and to guide the research in directions that will be of maximum relevance to industry;
- to keep the industrial advisory partners informed about new Web technologies and standards, and to disseminate other results of the research and development work undertaken in the project;
- to provide case studies, evaluation, and exploitation pathways for the new tools and technologies developed in the project.

P1 University of Manchester (VUM)

The University of Manchester Department of Computer Science has a long and distinguished research record dating back as far as 1948 when they developed the worlds first stored-program digital computer. Currently, the department has a strong reputation for research across a wide area of Computer Science and in particular in the area of information management.

The Information Management Group (IMG) carries out research in different aspects of data intensive application development. Current research seeks to extend the functionality of database systems, to exploit Description Logics (DLs) in advanced applications, and to make advanced information management systems easier to use. IMG is one of the leading international centres for research in the theory and practical application of DLs, having been responsible both for the implementation of THE state of the art DL reasoner (FaCT) and its use in a range of innovative applications including the OilEd ontology editor.

Members of IMG have wide experience in national and international projects in areas such as Managing Semi-Structured Information (DWQ, CAMELOT, STARCH, COHSE), Object Data bases (ODESSA, DOQL, POLAR) and User Interfaces to Data Intensive Systems (TEALLACH, KALEIDOQUERY, INFOLENS). IMG maintains close links at Manchester University with the Bioinformatics Group and also has a range of projects in this area (TAMBIS, GIMS, RASH, IRBANE). IMG also benefits from close links with the Formal Methods Group, the Medical Informatics Group and the Advanced Interfaces Group.

Carole Goble is a Professor in Computer Science, and co-leads the Information Management Group. She has been on the faculty since 1985, and has worked in the Multimedia and Medical Informatics groups. She was/is an investigator on a number of projects using DLs as modelling and retrieval languages for: medical information systems (PEN&PAD, GALEN), semantic hypermedia systems (COHSE), picture archives (STARCH), mediating disparate bioinformatic information sources (TAMBIS, TAMBIS-II), and improved protein function prediction using ontologies (Irbane). She is co-investigator in a basic research project on DL-based ontology servers (CAMELOT). She has over 40 publications in the area and has served on many conference committees on databases and multimedia. Her current interests lie in conceptual modelling, the use of ontologies and thesauri in a range of metadata and data-intensive applications, including multimedia, hypermedia, medical informatics and bioinformatics.

Ian Horrocks is a Lecturer in Computer Science. He graduated in Computer Science from Manchester in 1982, winning the prize for most outstanding graduate. After working in industry he returned to Manchester to complete a PhD in 1997. His FaCT system revolutionised the design of DL systems, redefining the notion of tractability and establishing a new standard for implementations. He designed the OIL language, is an editor of the DAML+OIL language and is a member of the Joint EU/US Committee on Agent Markup Languages. He is the lead researcher on several EPSRC projects and was a leading researcher on the DWQ project (Esprit 22469). He has published widely in major journals and conferences, winning the best paper prize at KR'98. He is a member of the programme/editorial committees of several international conferences, workshops and journals. His current research interests include knowledge representation, automated reasoning, ontological engineering and the Semantic Web.

P2 Vrije Universiteit Amsterdam (VUA)

The VU Division W&I, the Division of Mathematics and Computer Science of the Free University Amsterdam, brings to the project its extensive experience in the areas of knowledge management, ontological engineering, and distributed systems and intelligent agents analysis and design. For example, VU Division W&I researchers have delivered key contributions to internationally renowned methodologies, frameworks and languages for knowledge engineering and management such as CommonKADS, DESIRE, and (ML)2 . Related to this, the VU is well known for its recent work in ontology development (e.g. (KA)2 , PhysSys) and intelligent and multi-agent systems (e.g. HomeBots).

The VU Division W&I represents a long-standing international experience in direct as well as consortium-based industry-university collaborations, in addition to the fact that it is rated as the best academic research site in computer science in The Netherlands. Apart from direct joint work with industry, VU researchers have been active over the years in a wide range of European efforts, ranging from EU Esprit basic research (e.g. REFLECT, DRUMS, IBROW), large RTD projects (e.g. KADS-II, OLMECO, On-to-knowledge), networks of excellence (e.g. AgentLink), to industry take-up and user leveraging projects (e.g. TRACKS).

Dr. Dieter Fensel studied Social Science at the Free University of Berlin and Computer Science at the Technical University of Berlin. In 1993 he was awarded a Doctor's degree in economic science (Dr. rer. pol.) at the University of Karlsruhe and in 1998 he received his Habilitation in Applied Computer Science. From 1996 to 1999 he worked as a senior scientist at the institute AIFB, University of Karlsruhe, and since October 1999, he has worked as a Associate Professor at the Vrije Universiteit Amsterdam. He is involved in several national and internal research projects including IBROW (IST-1999-19005) and On-to-Knowledge (IST-1999-10132), where he is project coordinator. He is also coordinating the thematic network proposal OntoWeb (`www.ontoweb.org`). He has published over 150 papers and several books, he has organised numerous workshops and conferences and he edits a number of international journals including *IEEE Intelligent Systems*. His current research interests include knowledge engineering, knowledge management, intelligent information integration and electronic commerce.

Frank van Harmelen (1960) studied mathematics and computer science in Amsterdam. In 1989, he was awarded a PhD from the Department of AI in Edinburgh for his research on meta-level reasoning. After his PhD research, he moved back to Amsterdam where he was involved in various ESPRIT projects (REFLECT, KADS), and where he lead the development of the $(ML)^2$ language for formally specifying Knowledge-Based Systems. In 1995 he joined the AI research group at the Vrije Universiteit Amsterdam, where he holds a senior lectureship. He is currently involved in a number of Fifth Framework projects that are concerned with applying Knowledge Representation Techniques the the Web (IBROW, On-To-Knowledge). He is author of a book on meta-level inference, editor of a book on knowledge-based systems, has published over 60 research papers, and is one of the most-cited Dutch AI scientists.

P3 Consiglio Nazionale delle Ricerche, Instituto di Sistemistica e Bioingegneria (LADSEB)

CNR-LADSEB is an Institute of the Italian National Research Council, depending on the Italian Ministry of University and Scientific and Technological Research. The Institute has a staff of about 20 people and a similar number of external collaborators, mainly from the University of Padova. The research group on Conceptual Modelling and Knowledge Engineering (`http://www.ladseb.pd.cnr.it/infor/ontology/ontology.html`) has been active since 1991, in close cooperation with the Department of Philosophy of the University of Padova. The group performs basic and applied research on the ontological foundations of knowledge engineering and conceptual modelling, exploring the role of ontologies in knowledge representation, database design, information retrieval, natural-language processing, multi-agent systems, software engineering, and information systems in general. It is characterised by a strong interdisciplinary approach that combines together Computer Science, Philosophy, and Linguistics, and relies on Logic as a unifying paradigm. On the application side, the main emphasis is on the use of ontologies for electronic commerce, enterprise integration, knowledge management, and information access to the Web. Since the organisation of the *First International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation* in 1993, the group has played a leading role in promoting a well-founded ontological approach within the Computer Science community: we may mention for instance the special issue of *Data and Knowledge Engineering* on modelling parts and wholes, the special issue of the *International Journal of Human-Computer Studies* on formal ontology in the information technology, the organisation of the first workshop on *Product Knowledge Sharing for Integrated Enterprises*, and the organisation of a series of conferences on *Formal Ontology in Information Systems* (FOIS'98 and FOIS2001).

CNR-LADSEB has been the coordinator of two CNR national projects on formal ontology, conceptual modelling and information integration (SOLMC and ONTOINT) and participated in a project funded by Italian Space Agency on ontology-based database integration and access. It is currently involved in a large EUREKA project on ontology-driven Intelligent Knowledge Fusion (IKF) and is leader of a workpackage on ontology-based harmonisation of metadata standards in the OntoWeb thematic network.

CNR-LADSEB has been a member of the Italian STEP promotion centre, CeSTEP, has been an active participant in the ANSI-X3T2 Ad Hoc Group on Ontology Standard, and is now involved in the IEEE "Standard Upper Ontology" initiative. The institute has also contributed to pre-standardisation activities of the Foundation for Intelligent Physical Agents (FIPA) concerning ontology services.

Nicola Guarino (1954) is a senior research scientist at CNR-LADSEB. He graduated in Electrical Engineering at the University of Padova in 1978. In 1979–1984 he was responsible of the data acquisition and monitoring system of a large nuclear fusion experiment. He then joined CNR-LADSEB to work on knowledge representation issues. He has been active in the ontology field since 1991, and has played a leading role in the AI community in promoting the study of ontological foundations of knowledge engineering and conceptual modelling. His current research activities regard ontology design, ontology-driven information integration and retrieval, and conceptual modelling. He will be conference chair of FOIS2001, is associate editor of the Semantic Web area of the *Electronic Transactions on Artificial Intelligence* and of the *International Journal of Human-Computer Studies*, has published about 60 research papers, and has been guest editor of several journal special issues related to formal ontology and information systems.

P4 Institute AIFB, University of Karlsruhe (AIFB)

The Institute AIFB, University of Karlsruhe has grounded experience[27] in the fields of knowledge management, intelligent information integration, web-based broker architectures, knowledge engineering and knowledge discovery.

The Ontobroker and GETESS projects (funded by the German Ministry of Research and Technology) provide advanced information access to information sources at the web. Ontologies are used to define semantic access to semistructured information and free texts. As part of these projects environments offering sophisticated means for ontology engineering and learning from natural language are developed.

The institute is involved in many industry-university cooperations, both on a European and a national level, including a number of intelligent web systems case studies. Currently, the institute AIFB is participating in the EU basic research project IBROW and the IST project On-to-Knowledge. The Institute AIFB is subcontractor in the OnToAgents project[28] (part of the DARPA DAML initiative), where the goal is to establish an agent infrastructure on the WWW or WWW-like networks.

Prof. Dr. Rudi Studer is the head of the knowledge management research group, and obtained a Diploma in Computer Science at the University of Stuttgart in 1975. In 1982 he was awarded a Doctor's degree in Mathematics and Computer Science (Dr. rer. nat.) at the University of Stuttgart and 1985 he obtained his Habilitation in Computer Science. From July 1985 to October 1989 he was project leader and manager at IBM Germany, Institute of Knowledge Based Systems. Since November 1989 he has been a full professor in "Applied Computer Science", University of Karlsruhe. Since 1993, he has been the speaker of the Special Interest Group "Knowledge Management" of the Gesellschaft fuer Informatik (GI) in Germany. His current research interests include knowledge management, knowledge engineering, intelligent web brokers, and knowledge discovery.

Dr. Steffen Staab is Assistant Professor for Applied Computer Science at Karlsruhe University. He studied Computer at Erlangen University, Germany. In 1993 and 1994 he was a Fulbright scholar at the University of Pennsylvania, Philadelphia, PA, where he earned a M.S.E. in Computer and Information Science. He received a scholarship from the German National Science Foundation for joining a graduate program in Cognitive Science (Computational Linguistics) that led up to his PhD in Computer Science from Freiburg University, Germany, in 1998. Steffen Staab is project manager in the German government-funded GETESS project on ontology-based information extraction. He has organised several international workshops on ontologies and ontology-based applications. He has published in the fields of information extraction, knowledge representation and reasoning, knowledge management, knowledge discovery, and intelligent systems for the web.

[27]Further information at `http://www.aifb.uni-karlsruhe.de/WBS`
[28]Further information available at `http://WWW-DB.Stanford.EDU/OntoAgents/`

Industrial advisory board

The industrial advisory board is made up of representatives from some of Europe's and the World's leading providers and users of Web technology. The inclusion of non-European/multi-national members in the advisory board reflects the international nature of the Web; it will also boost the international standing of the project and help maximise our influence over the development of the Semantic Web.

Membership of the advisory board will be as follows:

Dr. V. Richard Benjamins, **iSOCO**, Barcelona, Spain
Dr. J. Bullock, **Canon Research Centre Europe**, Guildford, UK
Richard Chen, **InGenuity Systems**, Alviso, USA
Peter Crowther, **Network Inference**, Manchester, UK
Ian Davis, **Photonica**, London, UK
Dr. John Davies, **BT Advanced Communication Technology Centre**, Ipswich, UK
Robert Engels, **CognIT**, Asker, Norway
Dr. Einar H. Fredriksson, **IOSPress**, Amsterdam, Netherlands
Masahiro Hori, Ph.D., **IBM Tokyo Research Laboratory**, Kanagawa, Japan
Ian Lang, **Assistum (High Level Systems)**, Liphook, UK
Alain Léger, **France Telecom**, Cesson Sévigné, France
Robin McEntire, **GlaxoSmithKline Pharmaceuticals**, King of Prussia, PA, USA
Drs. J.J.E. van der Meer, **aidministrator**, Amersfoort, Netherlands
Dr. H. J. Müller, **T-Nova Deutsche Telekom Innovationsgesellschaft**, Darmstadt, Germany
Peter F. Patel-Schneider, **Lucent Technologies**, Murray Hill, NJ, USA
Dr. A. Persidis, **biovista**, Athens, Greece
Ulrich Reimer, **Swiss Life**, Zürich, Switzerland
Hans-Peter Schnurr, **ontoprise**, Karlsruhe, Germany
Massimo Soroldoni, **Nomos Systema**, Milan, Italy
Prof. Austin Tate, **ATK project, AIAI**, Edinburgh, UK
Arthur J. Thomas, Ph.D., **BioWisdom**, Cambridge, UK
Dr. Luca Toldo, **Merck**, Darmstadt, Germany
Dr. Michael Uschold, **The Boeing Company**, Seattle, WA, USA
Guido Vetere, **IBM Rome Tivoli Labs**, Rome, Italy
Matthew West, **Shell Services International**, London, UK
Mario Wolczko, **Sun Microsystems**, Palo Alto, CA, USA