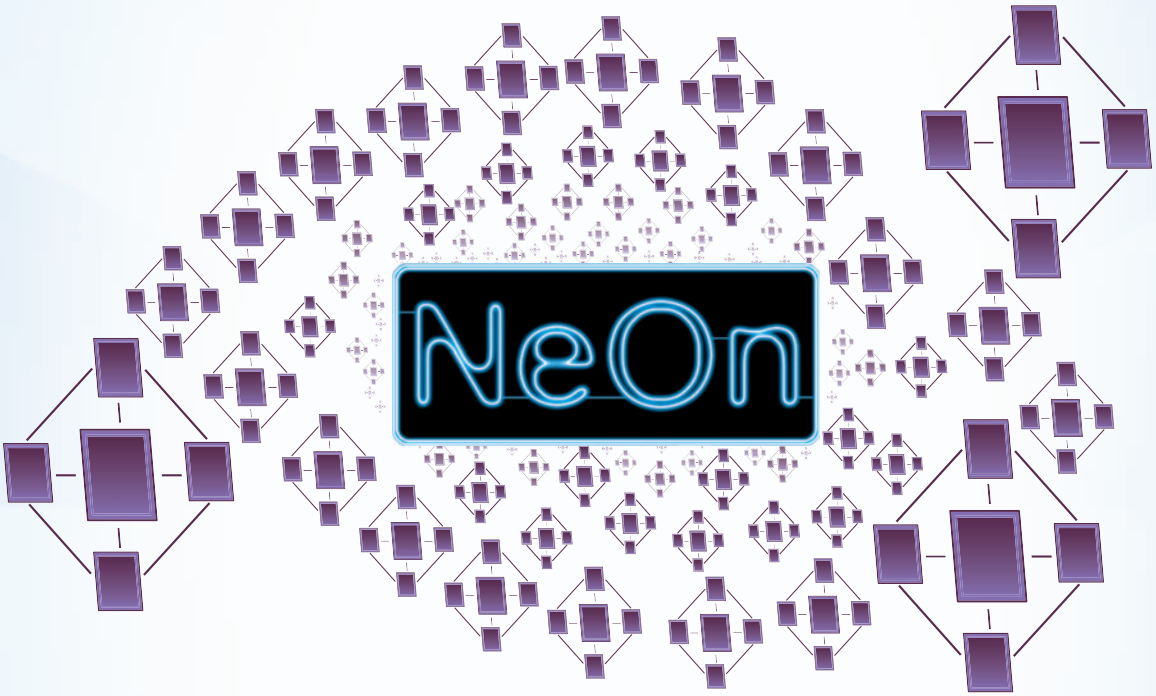
www.neon-project.org

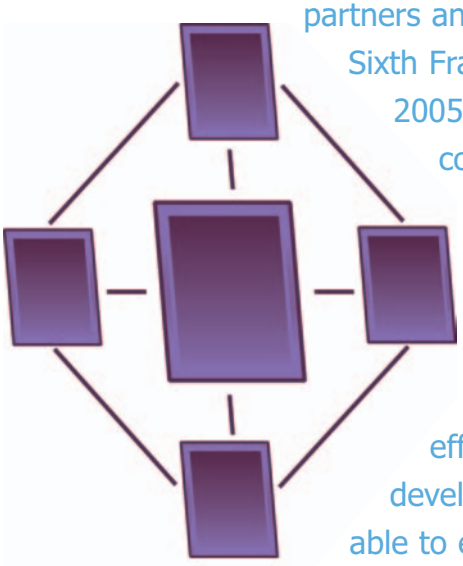


Lifecycle Support for Networked Ontologies

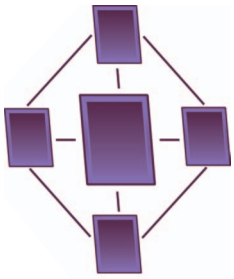
Annual Report 2009

Lifecycle Support for Networked Ontologies

NeOn is a 14.7 million Euros, four year project involving 14 European partners and co-funded by the European Commission's Sixth Framework Programme through grant no. IST-2005-027595. 14 partners from 6 European countries are involved in the project, which started in March 2006.



The aim of NeOn is to advance the state of the art in Ontology Engineering and Semantic Web technologies by providing effective methodological and tool support for developing a new breed of semantic applications, able to exploit effectively the large amounts of data, which are now becoming available on the Web.



We are currently in the fourth and final year of the project and most of our solutions are in place, including a comprehensive set of methodological guidelines for ontology engineering, a new ontology engineering platform (the NeOn Toolkit), which is compliant with the new OWL 2 standard, as well as a variety of leading edge methods and tools covering a wide spectrum of tasks in ontology engineering, including managing networks of ontologies, reusing best practices in ontology engineering, and locating, accessing and reusing content from the Semantic Web.

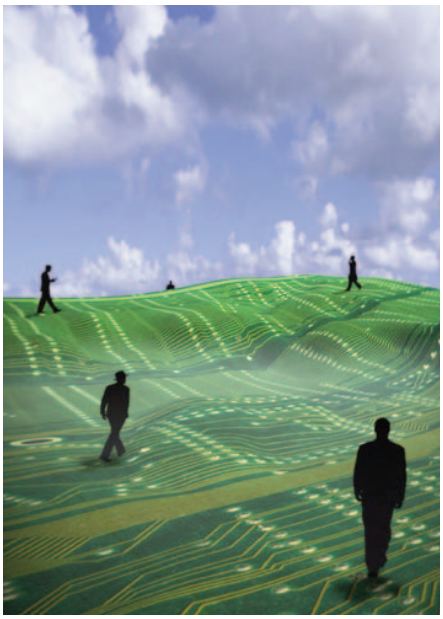
Our Ambitions

The Semantic Web already contains millions of semantic resources and this number is rapidly increasing. This unprecedented resource opens the way to a new class of semantic applications, which can exploit such large amounts of online data to provide new and advanced solutions for search, question answering, decision support and service integration. Hence, application design needs to reflect the fact that, given this new context, (i) newly developed *ontologies* are typically related to a network of pre-existing ones, and (ii) ontologies and metadata continuously evolve and need to be kept up-to-date within the changing application environments.

Our Scope For Innovation

The main challenge and innovation in NeOn is to push the boundaries of semantic technologies, in particular, in the emerging new context of the Semantic Web. The process here is analogous to what was witnessed in mainstream software development, which in the past 15 years has progressed from closed, relatively data-poor applications, to open, large-scale applications for accessing, integrating and re-using the vast amounts of information available on the Web, or in corporate intranets. A key premise of the NeOn project is that the current infrastructure for building semantic applications is not adequate to support this new generation of knowledge-based applications in the context defined by the Semantic Web. In particular, current ontology engineering environments do not yet fully support the new scenarios defined by the emerging semantic web. For instance, it is not easy to locate and link individual semantic data or definitions to an ontology under development and, more importantly, relatively weak support is provided to support the management of networks of ontologies, a key feature of this emerging new class of semantic web applications.

Therefore, to support the design and development of this new generation of semantically enriched applications, new methods, techniques and tools are needed. NeOn aims to provide efficient and scalable support for the entire lifecycle of networked ontologies. In particular, the project investigates methods and tools for managing the evolution of networked ontologies, for supporting the collaborative development of ontologies, and for the contextual adaptation of semantic resources. In addition, the project aims to ensure that not only are the solutions we develop good enough to tackle these problems, but that they are also cost-effective.



The NeOn Toolkit and the NeOn Methodology lie at the core of the NeOn vision, aiming to define a reference infrastructure and development process for creating and maintaining large-scale semantic applications.

Important Areas Of Our Work

Ontology dynamics

- Formally developing the notion of networked ontologies
- Developing methods for managing the evolution of networked ontologies
- Developing methods for reasoning with inconsistencies in networked ontologies

Collaborative aspects

- Investigating community-centred ontology design, ontology design patterns, and design rationale capture and management
- Developing effective methods for automatically selecting and integrating ontologies and their modules in response to application needs

Context awareness

- Formalizing and reasoning with the notion of context in semantic applications
- Developing methods for ontology alignment which do not require global consistency but can work with contextualized notions of local consistency

Human-ontology interaction

- Investigating and understanding user needs in current ontology engineering practice
- Customizing, personalizing & adapting networked ontologies to different user needs
- Supporting multilingual user interaction in the NeOn Testbeds
- Developing new methods for visualizing and navigating ontologies which make it easier to make sense of the content of ontologies

Deployment and product development

- An improved management of the Fisheries knowledge base at the Food and Agriculture Organisation of the United Nations (FAO) to support expert decision making and fish stock depletion assessment
- An infrastructure for e-Invoice exchange and management in the Pharmaceutical sector in the context of the PharmaInnova Cluster
- A platform for integrating and sharing information about pharmaceutical products in the context of Atos Origin's Semantic Nomenclator application

Our Achievements To Date

In the past 12 months we have made further progress in all areas of the project, producing a new OWL 2 compliant version of the NeOn Toolkit, trialling and evaluating the NeOn methodology, further improving our applications testbeds, developing new methods and tools for managing networks of ontologies, supporting collaborative ontology engineering, and customizing ontologies in accordance with user needs. We have also continued our outreach activities, showcasing NeOn technologies in a variety of international events, targeting both academic and industrial communities. In particular, these include ISWC 2009 and SemTech 2009 in the USA, and ESTC2009 in Vienna, as well as other events targeted at specific scientific sectors, such as the pharmaceutical sector. We have also targeted specific scientific communities at the forefront of the use of technology in standardization processes, such as the biodiversity and the biomedical ontologies communities.

The NeOn Toolkit

NeOn Toolkit available in open source

The NeOn Toolkit is our solution to the need for a new-generation ontology engineering environment to support the design and development of web-aware, semantically enriched applications. The NeOn Toolkit is designed around an open and modular architecture, which includes infrastructure services, such as registry and repository, and supports distributed components for ontology management, reasoning and collaboration in networked ontologies.

An evaluation version of Neon Toolkit V2.3 was released in November 2009. It is based on the OWL-API (v3) to better support the full functionality of the OWL 2 specification. The NeOn Toolkit is freely available in open source (under the Eclipse Public License, EPL) as the reference implementation of the NeOn architecture. Building on the Eclipse platform, the NeOn Toolkit offers a development environment familiar to software engineers, as well as an open framework for plugin developers - - www.neon-toolkit.org

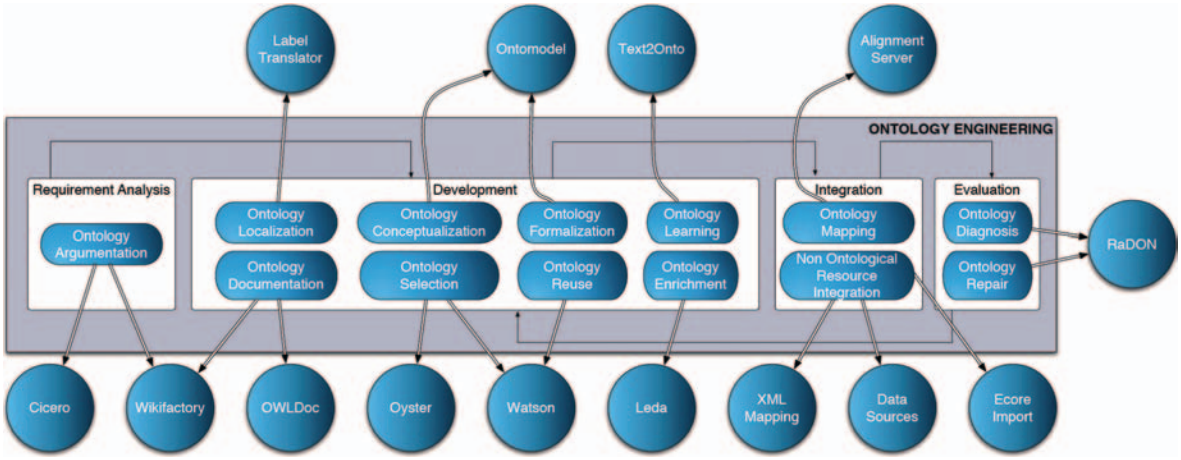


Figure 1. A selection of NeOn Toolkit plugins supporting ontology engineering activities along the lifecycle of networked ontologies

Plugins for the NeOn Toolkit

A typical NeOn Toolkit plugin consists of a set of Eclipse extensions that encapsulate a coherent set of ontology engineering functionalities. The plugin mechanism makes the NeOn platform very flexible, as not only does it support predefined extensions that fit into certain slots, but also allows for future extensions of current functionalities. The NeOn Toolkit also supports the integration of distributed plugins which include, for example, plugins based on web service standards to access remote components of the NeOn infrastructure. At the time of writing, 45 plugins are available for the NeOn Toolkit which can be found at – www.neon-toolkit.org/wiki/Neon_Plugins

Selected Research Highlights

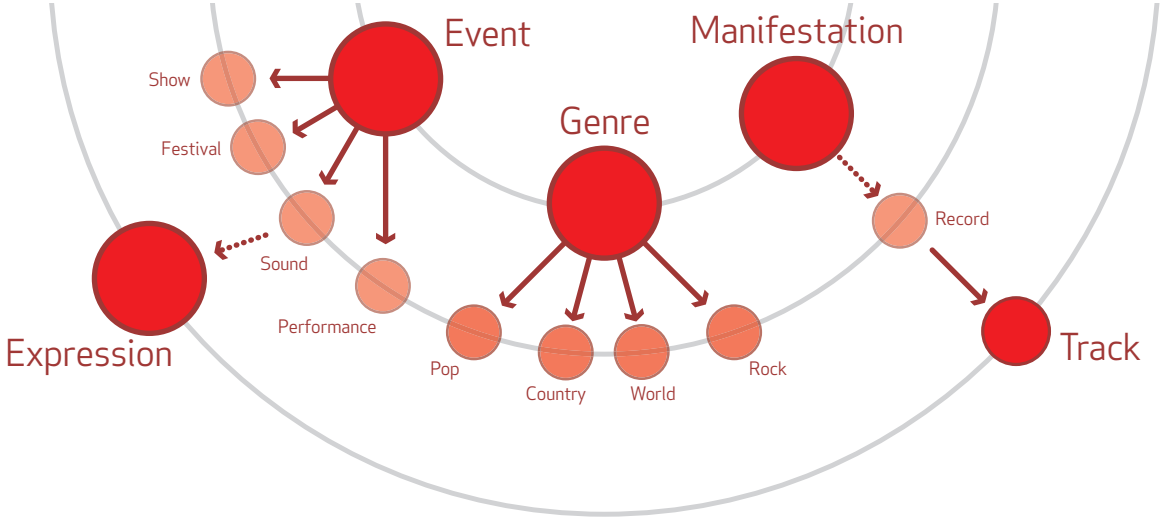
NeOn is contributing to the body of knowledge in several disciplines e.g. ontology engineering, semantic web, web science, and human-computer interaction. Of particular value are our methods and tools for working with evolving and contextually grounded networked ontologies. Examples of key activities and outputs to date include:

Ontology Design Patterns

While ontology engineering is often considered an art, the availability of a library of ontology design patterns is an important step towards achieving the ultimate goal of turning ontology design into a structured and reproducible engineering process. At the time of writing, the ODP portal, found at http://ontologydesignpatterns.org, contains 67 patterns divided into six main categories: Structural, Correspondence, Content, Reasoning, Presentation, and Lexico-Syntactic. The portal defines a focal point for the ontology engineering community, and provides support for submitting design patterns, retrieving them from the repository, posting modelling issues, discussing and reviewing existing solutions, and finding educational material on ontology design. Aldo Gangemi and Valentina Presutti serve as Editors-in-Chief, while a Quality Committee ensures the quality of the library. The integration between the ODP portal and the NeOn Toolkit is provided by the XD plugin, which implements the eXtreme Design method, supporting pattern-based design in ontology engineering. XD knows good practices and provides them to an ontology developer in response to a specific modelling issue. XD runs like a wizard, using dialogue and graphical interfaces, and also assists a developer in adding annotations and publishing the resulting ontology in a repository.

KC-Viz

The user studies carried out at the beginning of the NeOn project clearly indicated that the user interaction metaphors used in ontology engineering toolkits are largely inadequate, especially for those users with limited experience. Hence, a key challenge in our work on Human-Ontology Interaction concerns overcoming these problems and developing novel interactive frameworks for visualising and navigating large and complex ontologies.



To this purpose we designed and implemented an innovative solution, based on the idea of identifying 'key concepts' in ontologies and using them as starting points for exploring and making sense of large ontologies. Key concepts can be seen as descriptive ontology elements that best summarise what a particular ontology is about. In our work we defined this notion precisely, defining an algorithm for key concept extraction which combines theories from the cognitive sciences with formal methods from linguistics and ontology engineering. As a result, it becomes possible for a user to get a quick understanding of an ontology of, say, one thousand concepts, simply by being presented with 15-20 key concepts. This work has been received very well by the Semantic Web community to the extent that, when first reported at the 2008 Asian Semantic Web Conference, it earned its authors a best paper award.

This approach to ontology summarisation has provided the basis for developing an innovative approach to visualising and navigating ontologies. In particular, it enables 'middle-out ontology browsing', where it becomes possible to move through complex information spaces from the most valuable nodes (i.e., key concepts) and then to unfold larger chunks of the ontological graph to inspect specific sub-parts of an ontology. This approach is similar to map-based visualisation and navigation in Geographical Information Systems, where, e.g., major cities are displayed more prominently than others, depending on the current level of granularity. Another innovation, called 'conceptual zooming', offers the user the opportunity to calculate ontology summaries, taking into account both the content and the topology of the underlying model. Drawing upon a visual metaphor familiar to Web 2.0 users, where tags with greater popularity are depicted more prominently, our KC-Viz framework allows the user to distinguish between several layers of key concept importance, thus realising the notion of key concept at different levels of granularity.

KC-Viz has been implemented as a NeOn Toolkit plugin and can be downloaded using the 'update' feature on the NeOn Toolkit. Once deployed, it can process either OWL ontologies loaded locally in the NeOn Toolkit or remote ontologies known to the Watson ontology search engine and identified by their URIs.

Cupboard: sharing and reusing networked ontologies

Ontology engineers need a place to host their ontologies, as well as tools that can allow them to manage them and, ultimately, to make them available to others. As ontology users, we need tools that can help us to locate ontologies that are relevant to our task. In order to be able to make an informed choice, users also need information about the quality, provenance and key characteristics of the ontologies available for reuse. Finally, application developers need infrastructure components, which can support the exploration and querying of both standalone and networked ontologies.

Cupboard is an online ontology hosting system that intends to address the needs of these three categories of

users. Each user of Cupboard is provided with his/her own Ontology Space, where ontologies and alignments can be uploaded and stored. An ontology space provides a summary of the networked ontologies it includes, as well as means to review and attach rich metadata to them. Moreover, advanced search mechanisms are provided so that users can easily find, inspect and explore ontologies available online. In addition to these user-oriented elements, an Ontology Space also provides a virtual infrastructure for Cupboard users to build applications exploiting ontologies and alignments, as well as ontology ratings and metadata.

A number of services and APIs are deployed to handle tasks such as ontology search, exploration and querying. The Cupboard plugin for the NeOn Toolkit supports ontology development by reuse, in particular by exploiting the ratings users have given to ontologies. Experiments have shown that this plugin effectively facilitates the tasks of finding, selecting and integrating existing formal knowledge structures in an ontology development project.

Ontologies the modular way

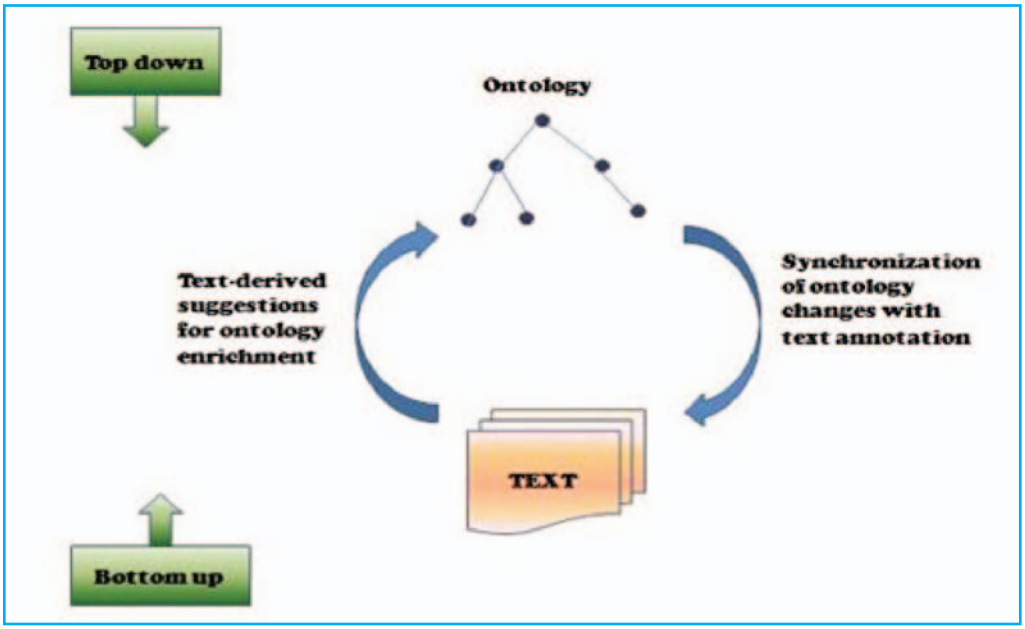
The NeOn Toolkit is the only ontology engineering environment that includes a complete framework for the creation and manipulation of modular ontologies. First, following best practices in software design, ontologies are encapsulated in components that specify their interfaces, i.e., the elements they expose for other modules to reuse, and the ones they reuse from other modules. Modular ontologies can then be completely specified in terms of these ontology components. In addition, the NeOn Toolkit also includes algorithms for decomposing ontologies into a set of modules, and also for extracting specific modules from ontologies, on the basis of users' specifications. Finally, whether they are manually specified or automatically extracted, modules can be combined using simple set-based operations. For instance, new modules can be created simply by merging two modules or by removing a set of definitions from a module.

The set of modularisation plugins included in the NeOn Toolkit provides all the necessary support to adopt a modular approach to ontology design and reuse. However, such an approach requires some experience and knowledge on the part of the ontology engineer. For this reason, the NeOn Methodology also provides guidelines for the modularisation activity, thus facilitating the construction of modular ontologies by relatively inexperienced ontology engineers.

GATE and The NeOn Toolkit

GATE, one of the most widely used and well known natural language processing (NLP) architectures, is now interoperable with the NeOn ToolKit ontology change management methodology, so that ontology change can be propagated in two directions.

Our top-down approach enables changes made to the ontology to be propagated to textual metadata in the form of annotations so that as little information as possible is lost, while our bottom-up approach enables an existing ontology to be augmented on the basis of textual evidence from NLP techniques, which provide the linguistically motivated basis for new links between unstructured text and an ontological representation.



This two-way mechanism (illustrated above) ensures that the dynamic relations between ontologies and texts are maintained, an especially important functionality in corporate semantic webs. Text annotations are fully synchronized with new versions of ontologies (whereas the link with older versions is maintained), and new textual evidence offers suggestions for ontology enrichment for expert evaluation within the Toolkit. A change log management system ensures full interoperability between GATE and the NeOn Toolkit and enables collaborative distributed ontology modification via both architectures

Sector And Community-Level Contributions

NeOn is applying its scientific and technological outcomes to a number of innovative test beds in two early adopter sectors, namely, in the Pharmaceutical and the Agriculture/Fisheries sectors. In Year Four, we have progressed the software development and deployment of three applications in these domains, which act as use cases for assessing the performance of NeOn solutions in concrete real-world scenarios.

Applying NeOn to the Pharmaceutical Domain

NeOn's activity in the Pharmaceutical domain is focused on two different segments of the market. First, we aim at cataloguing information about pharmaceutical products in order to enable pharmacies in Spain and also Europe-wide associations of pharmacy professionals access to homogenized information repositories on such products. NeOn technologies enable the maintainers of the distributed repositories to leverage integration of their heterogeneous pharmaceutical content. Information about chemicals and their regulations for human consumption is highly distributed and pharmacies do not have real-time access to it. As a solution to this problem, NeOn has produced a global vade mecum where distributed databases and regulations can be integrated and kept up-to-date with the databases containing the official and approved information about medicines.

NeOn is also tackling the financial side of the co-operation in the Pharmaceutical sector. Since a European directive in 2002 authorised the use of digitally signed invoices for commercial transactions, the use of electronic invoices has grown exponentially. However, the main obstacle to broader uptake of this cost-saving business interaction is the heterogeneity of the schemas to represent and exchange invoice information, as well as limited uptake of invoicing standards by the main players in the sector. The technologies provided by NeOn enable the different peers involved in a commercial transaction to automatically process arbitrary invoices by abstracting the underlying information from the particular data streams, representation formats and proprietary styles championed by different ERP systems.

Considerable progress has already been made in recent months and, in particular, work has begun on integrating the invoicing prototype with the infrastructure used by the PharmaInnova cluster of pharmaceutical companies in Spain. Upon completion, PharmaInnova members will use NeOn technology for automatic invoice exchange in the PharmaInnova cluster.

Ontologies for the Fisheries use case

The Food and Agriculture Organization (FAO) of the United Nations leads international efforts to defeat hunger. Serving both developed and developing countries, FAO acts as a forum where nations meet to negotiate agreements and debate policies. FAO is a very knowledge-intensive organization and one of its key goals is to provide information to users in the most efficient and accurate manner.

Extending the capability of computers to understand information better and deliver the best results to the users is fundamental to this goal. FAO participates in NeOn to enhance the semantics that underpins the technical information contained in the FAO intranet; to make it easier for its stakeholders to gather and analyze the data they need. FAO's expertise in global Fisheries information systems contributes to the NeOn project in the context of developing a decision support system to aid in assessing possibly depleted fish stocks.

This case study is used to test a new approach to compiling, sharing and disseminating electronic information. Progress on this case study includes a set of medium to large-scale ontologies developed by reusing and reengineering the existing Fisheries classification systems and the prototype of a Fish Stock Depletion Assessment System (FSDAS), to test the use of semantic technologies to provide access to a variety of information sources. All ontologies are available from the FAO website at http://www.fao.org/aims/neon.jsp.

Finding Information About NeOn

The NeOn website www.neon-project.org is the primary showcase of the project's outputs. These include technologies and publications, in the form of project reports, papers presented at scientific conferences, lectures delivered at industrial or academic events and articles published in scientific journals and magazines, as well as news about the project and the sector in general.
The NeOn community portal www.neon-toolkit.org plays a central role in the community building activities around the NeOn Toolkit. It is intended to serve as the focal point for distributing and for accessing information about the NeOn Toolkit (be it for the end users, or plugin developers). The portal is equipped with the following features and functionalities:

- Download area for releases and updates of the NeOn Toolkit
- Resource area with documentations, tutorials, FAQs
- Mailing lists, forums and wikis for users and plugin developers
- Bug reporting and management

Further Information

Coordinating partner:
The Open University (United Kingdom)

Partners:

- Universität Karlsruhe TH (Germany)
- Universidad Politécnica de Madrid (Spain)
- Software AG (Germany)
- Intelligent Software Components SA (Spain)
- Institute 'Jozef Stefan' (Slovenia)
- University of Sheffield (United Kingdom)
- Institut National de Recherche en Informatique et en Automatique (France)
- Universität Koblenz-Landau (Germany)
- Ontoprise GmbH (Germany)
- Consiglio Nazionale delle Ricerche (Italy)
- Food and Agriculture Organization of the United Nations (Italy)
- Laboratorios Kin (Spain)
- Atos Origin SAE (Spain)

For More Information

Email Project Administrator: contact@neon-project.org
Visit project web site: www.neon-project.org