NeOn Highlights

www.neon-project.org

EU-IST-2005-027595 SIXTH FRAMEWORK PROGRAMME

From the Co-ordinator

Welcome to NeOn Highlights – our window into the NeOn project. Our aim is to advance the state of the art in ontology engineering and Semantic Web technologies. More specifically, we want to provide 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 available on the Web.

Our research agenda closely follows the emerging open context of the Semantic Web. As shown by the two NeOn case studies described in this newsletter, we are witnessing a trend away from relatively data-poor, closed systems, towards open semantic applications, which are able to locate, integrate and utilize large amounts of distributed data. These are usually derived from a variety of sources, are heterogeneous, and may exhibit different degrees of formalization and quality. In such new scenarios a number of issues to do with supporting contextual viewpoints, collaboration and data integration become crucial and indeed these topics define key areas of work in the project. Another important aspect that influences the trajectory of the project is the increased level of integration between the Semantic Web and Web 2.0 technologies. Much like Web 2.0 environments emphasize distributed content production, in NeOn we are developing methods and tools for managing knowledge, which is distributed, heterogeneous, contextualized, and developed collaboratively.

Our rich research agenda is further motivated by the fact that existing ontology engineering environments only partially support the novel applications described above. Indeed, as we detail in this newsletter, a user study conducted early in the project highlighted limitations in the level of support currently provided for handling multiple ontologies and for integrating heterogeneous semantic web data.

At the time of writing the NeOn project is just about to complete its second year and I am happy to say that the collaboration between NeOn's 14 European partners has already led to significant advances in ontology engineering and Semantic Web technologies. This newsletter reports on some of these exciting developments. I hope you will enjoy exploring it.

Fighting hunger with ontologies

The Food and Agriculture Organization of the United Nations (FAO) leads international efforts to defeat hunger. Four million times a month, someone visits the FAO website to read a document or consult one of the many available databases. The constant growth in resources on the FAO website requires that the search facilities and the mechanisms to aggregate and correlate information need to be constantly improved. For this reason, FAO has been an early adopter of semantic technologies and continues to explore ways in which these technologies can help providing better services to FAO's users.

An important area for FAO is Fisheries Resource Management. Fish currently provide at least 20 percent of the animal protein needs of over 2.5 billion people in the world. Unfortunately, as indicated in a recent FAO report, 25 percent of marine fish stocks are either overexploited, depleted, or recovering from depletion. Hence, the effective management of shared fish stocks is one of the great challenges facing us.

FAO has several information and knowledge organization systems related to Fisheries, which manage and disseminate fishing statistical data, GIS data, information on aquaculture, description of fish stocks, etc. Such systems are developed and maintained by different people, who interact with experts and users worldwide.

Although much of the fisheries data are 'structured', they are very heterogeneous and not necessarily interoperable. Hence, we are working together with our partners in the NeOn project to develop solutions which will enable us to manage effectively these large, heterogeneous and distributed sets of data. Our ultimate aim is to develop a radically new ontology-based decision support system, which will make it possible for us to assist member countries in monitoring trends in the level of fishery stocks, thus ensuring that potential crises can be identified as early as possible.

NeOn Toolkit

The NeOn Toolkit is a new generation ontology engineering environment, combining industrial-strength robustness with a comprehensive support for the ontology engineering life-cycle. The target community for the platform includes both researchers from the Semantic Web, knowledge management and related areas, as well as practitioners from the world of industry and commerce. The core of the Toolkit is freely available as open source software and provides the reference implementation of the NeOn architecture. Building on the Eclipse platform, the NeOn Toolkit is designed around an open and modular architecture, which includes infrastructure services, such as a registry and a repository, and distributed components for ontology management, reasoning and collaboration in networked environments. The Toolkit defines an open framework for plugin developers and already within the NeOn community a number of plugins addressing different ontology-lifecycle activities have been released or are under development. These include support for visual modelling (OntoModel), ontology reuse (Watson), ontology learning (Text2Onto), ontology matching (R2O, FOAM), ontology diagnosis and repair (RaDON), as well as an ontology registry (Oyster). Other plug-ins will be produced at later stages in the project.

www.neon-toolkit.org

In addition to supporting an open, collaborative approach to developing functionalities, the extensible architecture of the NeOn Toolkit is also crucial to enable the realization of certain key advanced functionalities, such as integrating external infrastructures like DBMS, supporting a dual language approach – including both OWL and rules, and ensuring compliance with Service-Oriented Architectures (SOA).

The combination of a comprehensive set of state-of-the-art functionalities for ontology engineering with solutions which facilitate the integration of conventional software technologies in semantic applications, within an open source framework, distinguishes the NeOn Toolkit from other ontology engineering environments in what is rapidly becoming a strategic area for many companies. The first open source version of the Toolkit has been unveiled at the 2007 International Semantic Web Conference and, as befits an open, collaborative enterprise, we welcome contributions from both users and developers, who are keen to help us test and develop the Toolkit.

NeOn Partners

NeOn has brought together a critical mass of scientific and technological leaders in the fields of ontology engineering, Semantic Web, context awareness, collaborative technologies, software engineering and human-computer interaction from several European countries. The consortium is formed by fourteen institutions with extensive experience in research and development from United Kingdom, Germany, Spain, Italy, Slovenia and France.

The Open University

AIFB - Universität Karlsruhe

Software AG

Laboratorios Kin S.A.

Intelligent Software Components S.A.

Universidad Politécnica de Madrid

Institut National de Recherche en Informatique et Automatique

Institut Jozef Stefan

Universität Koblenz-Landau

University of Sheffield

Ontoprise

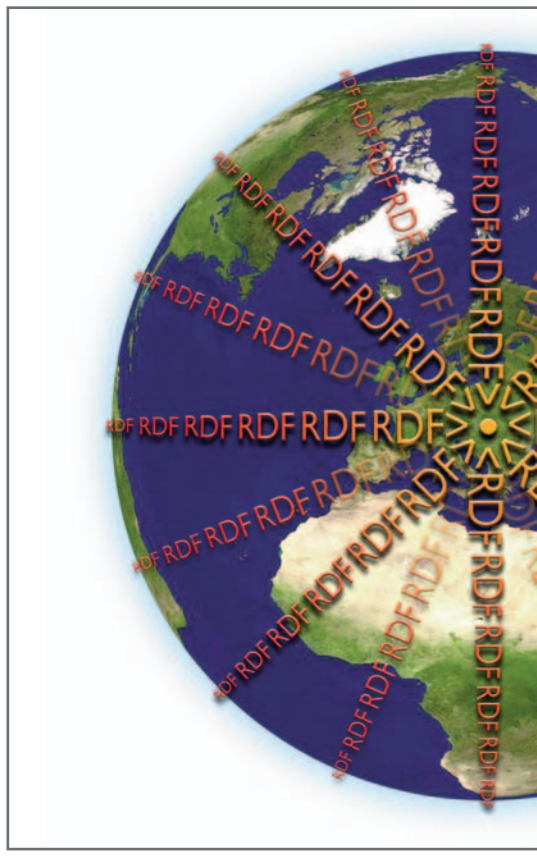
Consiglio Nazionale delle Ricerche

Food and Agriculture Organization of the United Nations

Atos Origin

Semantic Technologies: T

Imagine a world where cars autonomously attend to drivers' requests according to the current traffic conditions. A world where citizens and public administrations understand each other despite their use of different jargons. A world where people can collaborate at work or at home supported by virtual desktops, which intelligently anticipate the services required by the user.



In such a world, pharmacies will be able to access meaningful information on chemicals from a virtual vademecum regardless of the way in which such information is encoded – e.g., from distributed repositories in different languages. Furthermore, organisations involved in economic transactions will be able to exchange arbitrary electronic business documents by automatically extracting the information contained in them, even though these documents may use different structures and terminologies, thus saving large amounts of money in the process. And finally, in such a world we will be able to intelligently monitor our natural resources worldwide, preventing hunger for millions of people.

All these scenarios are part of the world in which we live, and semantic technologies can help us in realising them. In fact, providers of semantic technologies are experiencing an unprecedented pull from the market, which demands solutions to these problems. At the same time, the technology push related to automatic annotation, natural language processing, context awareness, customisation, and service and data integration, is now reaching the required level of maturity to satisfy this demand.

Ontologies are used to characterise the do
applications, and therefore are at the hear
Additionally, in a strongly interconnected w
knowledge-intensive systems include colla
dynamics, and human factors. NeOn, as par

Tackling diversity in the pharmaceutical sector

NeOn's activities in the pharmaceutical domain focus on two areas of the market. First, we aim to develop methods and tools to support and maintain homogeneous access to information about pharmaceutical products. Specifically, information about chemicals and their regulations for human consumption is highly distributed across a number of heterogeneous repositories, to which pharmacies do not have real-time access. As a solution to this problem, NeOn will produce both a global modelling schema and a concrete system, which will support the integration and management of distributed databases and regulations. Our target users include both pharmacies in Spain as well as European associations of pharmacy professionals.

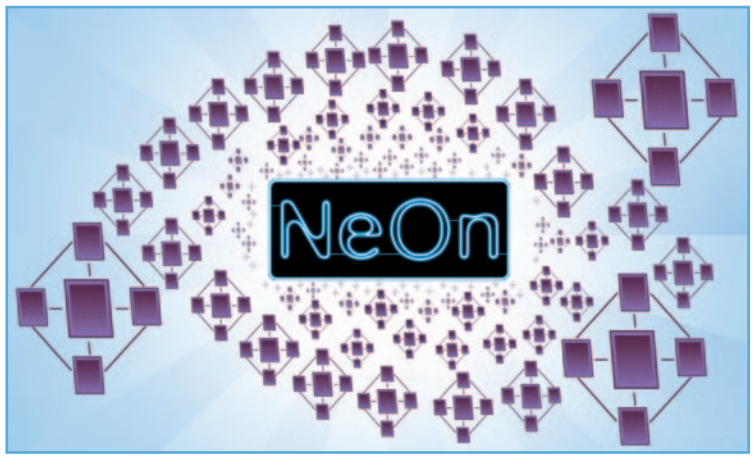
NeOn is also tackling the financial side of the pharmaceutical sector. Since a European directive back in 2002 authorised the use of digitally signed electronic invoices for commercial transactions, the use of electronic invoices has increased exponentially. However, the heterogeneity of the means to represent and exchange invoice information, as well as the lack of invoice standards adopted by the main players of the sector have been identified as key obstacles preventing the widespread use of digital invoices.

To make matters worse, the ERP systems managing this information and the different languages for exchanging electronic business documents that exist in the market are extremely diverse. It is the goal of NeOn to enable the different peers involved in a commercial transaction to automatically process arbitrary invoices by abstracting their content out of any particular representation format or technology.

Providing ontology-driven support for integrating and managing large amounts of heterogeneous data in the pharmaceutical sector makes a lot of sense from a business point of view: the savings from integrating business processes within this network are estimated at nearly 80% of the total current costs per year.

Managing Networked Ontologies

The Semantic Web already contains thousands of ontologies and is growing fast. As a result, more and more semantic applications tap into this very large resource, integrating information from a variety of sources. In such a scenario current ontology management models, which typically cater for a single ontology, become inappropriate, as future applications will require models to build and manage networks of ontologies, i.e., collections of ontologies that are interrelated in various ways and may not necessarily exhibit overall consistency.



Unfortunately, current ontology languages lack a number of features to explicitly express the relationships between ontologies. In particular mappings (also called alignments) are fundamental elements in a networked ontology model: they define relations between concepts of different ontologies, allowing applications to use these ontologies together, while keeping them distinct. Techniques for managing modular ontologies are another essential part of our model, which builds on the established notion of a module to support the design of ontologie characterized as the combination of components developed and managed independently.

For these reasons, NeOn has developed a networked ontology model, which extends standard representation languages for the Semantic Web, such as OWL, with primitives to express, amongst others, notions such as mappings and modules. Such model provides the formal basis for the NeOn infrastructure and ensures a formal common ground for the various functionalities we are currently developing in the project.

The European Dimension

mains targeted in semantic
t of this class of technologies.
world, key aspects for
boration, context, change
rt of Europe's strategy towards a

knowledge economy aiming to bridge the gap between technology and society, is dealing with these factors, working towards providing industrial-strength solutions to support the knowledge lifecycle.

Being technologically interesting is, however, not enough. Successful technologies need to be convincing in a variety of dimensions, including business context, potential applications, market readiness, economics, society aspects (what is hot and what is not), and legal issues. Semantic technologies are certainly in the position to satisfy such requirements and "drive the technology drivers". Indeed, success stories like the ones described earlier create a virtuous loop, influencing the strategic focus of many large organisations, and therefore opening up new business opportunities.

As in many other areas, there seems to be a race between US and Europe towards exploiting semantic technologies, with different strategies. In Europe, efforts driving the innovation process are still mainly funded with public money. However, as interest in semantic technologies continues to grow, we are witnessing a greater and greater involvement of private capital. While some of this is funding technology-centric Semantic Web start-ups, which are developing the core technologies required by semantic applications, we are also witnessing a number of enterprises adopting semantic technologies and developing their own Corporate Semantic Webs, to improve their ability to manage their core business processes. In other words, semantic technologies are slowly becoming mainstream. By bringing together large user organisations, major research centres and the key European developers of semantic technologies, the NeOn project is one of the key catalysts for this process.

What do users of ontology engineering tools expect?

Ontology engineering is a complex task, which can be either facilitated or hampered by the tools in hand. To understand how existing ontology engineering tools fare in this respect, NeOn has carried out a user study to determine how IT practitioners (who are not necessarily ontology engineering experts) view their tools and what functionalities they are missing.

Specifically, we analysed 30 users undertaking three ontology engineering tasks with the help of two environments: Protégé and TopBraid Composer. Among our findings we uncovered issues concerning the level of support for representing complex logical statements and the lack of contextual factors in the amount of information shown by a tool during different activities. For instance, it is good to see a list of properties when defining logical restrictions, but this list is distracting if shown during the definition of new concepts.

Ontology engineers also explicitly identified a need for more expressive visualizations; better debugging support; the ability to work seamlessly with several ontologies; the ability to reuse subsets of concepts from other ontologies, rather than entire models; and the ability to express OWL statements using semi-natural (English) language. The findings were presented at the OWL: Experiences and Directions 2006 Workshop, attracting much interest from both academic and commercial developers - the workshop paper is available online at http://oro.open.ac.uk/6241/. The lessons learnt from this study form a key input to the design and implementation of the NeOn Toolkit.

Eloquent discussions with Cicero

Designing an ontology is a complex task that requires the collaboration of domain experts and ontology engineers. In order to come to a consensual model of the domain, the participants in the engineering process must share and discuss their different viewpoints in an efficient manner.

Thus, discussions between the participants are an important part of collaborative ontology engineering. For this purpose, the Cicero tool has been developed within the NeOn project to facilitate asynchronous discussions and decision making processes between the relevant stakeholders. These don't have to meet at one location or agree upon a common time for their discussions.

Cicero aims to improve both the efficiency and the robustness of the collaborative process. Efficiency is achieved by providing a methodology for structuring issues, their solutions and the corresponding arguments. By capturing the whole discussion, there remains a record of the pros and cons for each issue, which makes the decision making process more efficient. Indeed, concrete evidence has been found, that the methodology leads to a reduction in the time needed to find solutions to issues.

The process also becomes more robust, because the structured record of the discussions captures the rationale underlying the design of an ontology. By attaching a discussion to the elements in the ontology that are influenced by it, it is later possible to retrace the reasons for certain modelling decisions. Furthermore, if new requirements have to be taken into account, the availability of a detailed design rationale facilitates the resumption of the discussion.

Cicero is implemented as an extension of the Semantic MediaWiki and is freely available at **http://isweb.uni-koblenz.de/Research/Cicero**. Future work will produce a Cicero plugin for the NeOn toolkit, to ensure that the design rationale for the elements of an ontology can be accessed using the same environment used for developing and using the ontology.

Interview with Laboratorios KIN

The Spanish pharmaceutical company, Laboratorios KIN, S.A., is NeOn's newest partner. Part of the PharmaInnova initiative, KIN brings first hand domain expertise to the project case study on the automatic management of electronic invoices in the pharmaceutical domain. In addition, they play a key role at the interface between the technical output of the project and the end users: they will organize events for disseminating the relevant NeOn solutions, including the electronic invoicing prototype platform, within the PharmaInnova cluster, as well as co-ordinating the evaluation of the prototype by the PharmaInnova members.

Joan Candini works as Information Technologies Manager in Laboratorios KIN and is their team leader for the NeOn project. He explains why NeOn appeals to his organization.

"The interest of KIN in NeOn is twofold. On the one hand, KIN is interested in improving its R&D process, in particular in the context in which new materials are discovered that may give rise to new products. Currently, KIN researchers have to manage a huge amount of information and scientific literature and NeOn's advanced ontology technology is a promising candidate to improve this process. Hence by being part of NeOn we can improve our competence in this area.

On the other hand, KIN is interested in the e-Invoicing case study in the pharmaceutical industry. Becoming a full member of NeOn allows us to have a much closer involvement in the project, thus facilitating the communication and the transfer of results between NeOn and PharmaInnova.

Laboratorios KIN generate a high number of paper invoices each year. Hence, we are looking to implement technological solutions to manage these electronically and reduce the associated tasks that don't add value. For example, we would like to integrate the invoices received in our ERP and apply logical business processes to solve discrepancies, thus addressing the inevitable human errors associated with entering information in the invoice.

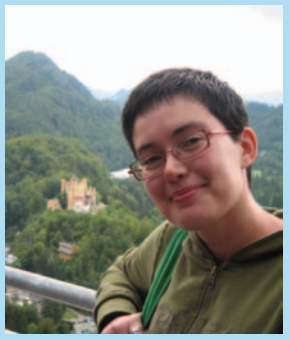
In addition to looking at new technological solutions for KIN, we aim to show these solutions to the companies in the pharmaceutical sector and by using their feedback, to make the system even more attractive for them. By doing this we can start to standardize electronic invoice processes in this sector.

Organizational change management is not an easy task. In order to deploy this technology successfully throughout the company and make a significant impact on work-processes, this activity will have to be managed carefully through effective dissemination.

Therefore, the dissemination task is essential, both within KIN and across the Spanish Pharmaceutical industry. Promoting this vision – the standardization of both the invoicing process and the e-invoice format – will facilitate communication between the different stakeholders and is a key success criterion for the NeOn project.

These factors make NeOn a key project for us. Because of the huge technological challenges and the financial cost, it is simply not viable for KIN to achieve these goals alone. NeOn provides us with the ideal environment in which we can bring our experience and know-how, and contribute to the development of a solution which will produce tangible benefits for the Spanish pharmaceutical industry as a whole".

NeOn People



What sort of people work in NeOn? We went to Spain to meet Mari Carmen Suárez-Figueroa, who is a PhD student in the Ontology Engineering Group at the Universidad Politécnica de Madrid.

• **What are you working on at the moment?**
My current work focuses on creating methodological guidelines for building ontologies embedded in ontology networks that are in continuous evolution. The goal underlying such methodology is to support the collaborative aspects of ontology development and the re-use of existing knowledge resources (e.g., ontologies, thesauri, databases, etc.).

• **Who are you collaborating with?**
I work very closely with the partners involved in the NeOn use cases (ATOS, FAO, and iSOCO), who provide me with real-world scenarios which I use as a key input for developing the methodology. I also collaborate with organisations such as CNR and the OU, since our research areas are related.

• **What are your future plans in NeOn?**
I expect to create a practical methodology for building ontology networks, easy to use and understand. Hopefully this methodology will prove helpful to ontology developers!

Additionally, as a member of the NeOn Gender Observatory, I am working to promote the active participation of women in scientific research.

• **What are your expectations for NeOn?**
The Semantic Web will be characterised by a very large number of ontologies embedded in ontology networks. Future Semantic Web applications will be based on such networks. Within NeOn, we have the mission of providing methodological and technical support to facilitate the development of applications which can successfully make use of this complex Semantic Web.

I'm confident that NeOn will provide a good solution for managing the ontology networks used in particular Semantic Web applications. There are many outstanding researchers working in NeOn, I am absolutely sure that together we will obtain great results!

• **Could you tell us what you enjoy most about working in NeOn?**
Working in a European project, such as NeOn, is very exciting. I have the opportunity to work with world-class researchers from different countries and, thus, I can learn from them and also practise different languages. I also have very good moments with fellow NeOn colleagues during our project dinners and when we go to pubs to discuss our problems ;-)

• **What do you do when you are not working in NeOn?**
I'm also working on my PhD thesis, which is closely related to my work in NeOn. Regarding my personal life, I love travelling by car around different countries. I like learning languages. I speak English and Italian, and I'm learning German.

The NeOn Glowfest

The NeOn Glowfest is a series of events, unique to the NeOn project, which bring together developers and users who wish to learn about the latest NeOn technologies, interact with the members of the NeOn project in an informal, social context, and get complimentary glowsticks.

The first Glowfest was held in Innsbruck in June 2007 and was co-located with the European Semantic Web Conference. The key focus of this event was the release of the initial version of the NeOn Toolkit. The second Glowfest took place in Korea in November 2007, co-located with the International Semantic Web Conference. During this second event, the first open source version of the NeOn Toolkit was unveiled and a number of plugins demonstrated to the audience. Both events have proved very successful, attracting a large crowd, despite competing for attention in very busy and high quality schedules. In both Glowfests, the participants clearly enjoyed the presentations, the drinks and the glowsticks.

Future events will be announced on the NeOn project website, www.neon-project.org. So, please visit it to find out about the next GlowFest and to receive regular updates about the NeOn project.

Ontology Design Patterns

Producing ontologies, which are both practical and of high-quality, is notoriously difficult. An approach that attacks the problem at its root is the creation of a catalogue of reusable, modular *ontology design patterns*. These describe use cases, guidelines, examples, and ontology modules that can help with specific ontology design issues, such as, for instance, how to design ontologies that talk about:

Water regions within geographical regions within different countries;

Fishery techniques involving certain gear types according to fishing species;

Pharmacological products having certain components;

Drug evaluation guidelines for pharmaceutical labs;

The validity of this approach is well known in both software and knowledge engineering, and it was recently put to the test with the students of a PhD Course we have recently held at the Computer Science Department of the University of Bologna.

The response from the students was very positive. For example, this is how one of them assessed the approach:

"I've found the course about computational ontologies both interesting and helpful as it has encouraged us to think about ontology creation in new ways. In my past experiences I created a student ontology and a medical ontology related to neurological knowledge. I understood, attending the course and hands on sessions, that ontological patterns may improve the quality of my future work".

Our catalogue of ontology design patterns is directly related to the NeOn use cases and is rapidly growing. In addition, it is also open to the wider community, as it builds on Web2.0 and semantic technologies to support open, collaborative development and evaluation. Hence, it provides the first worldwide initiative of this kind. The catalogue, which can be found at http://www.ontologydesignpatterns.org, includes a variety of types of patterns, such as patterns derived from high-quality reference ontologies, patterns contributed from web users, patterns derived from *data models*, and lexical patterns from resources such as FrameNet and VerbNet.

A gateway to the Semantic Web

While the Web is essentially a network of documents, even though many of them are generated from underlying databases, the Semantic Web is envisioned as a network of data. These can be automatically handled by software programs to facilitate their use and exploitation. The technologies to represent and publish data and knowledge alongside classical web pages are already well established and the Semantic Web is now steadily growing. Providing a gateway to support the development of Semantic Web applications is the goal of Watson, http://watson.kmi.open.ac.uk. Roughly speaking, the ambition of Watson is to become a sort of "Google for the Semantic Web", paying particular attention to its fundamental aspect: the representation of meaning for machines.

watson
exploring the semantic web

Ontologies are the pillars of the Semantic Web: they provide schemas that capture a common understanding of a given domain and structures for semantically representing data on the Web. Collecting, analysing and indexing these ontologies and the corresponding semantic data are the core activities realised by Watson. The Web is crawled in order to discover semantic data, traversing common hyperlinks as well as the semantic links relating pieces of data together. Various elements of information are extracted from the collected files, paying particular attention to assessing the quality of the represented knowledge, to help applications in selecting the most appropriate and useful elements.

As a Web system, Watson provides a range of interfaces for users to query, access and retrieve the indexed data, but, more importantly, as a Semantic Web system, Watson provides interfaces for software programs, enabling the development of applications that automatically select, exploit and combine the knowledge published on the Semantic Web. Watson also exploits the latest standards for querying Semantic Web data (e.g., SPARQL), hence ensuring interoperability with the various applications and portals that have already adopted the Semantic Web as their basic infrastructure. Revyu.com, http://revyu.com, the review and rating site that won the 2007 Semantic Web Challenge, is an example of a platform that interacts with Watson through Semantic Web technologies.

In a nutshell, Watson supports the development of a new generation of intelligent applications that dynamically exploit the knowledge published on the Semantic Web as a smart way to solve real-life problems. The number of applications relying on Watson is increasing, ranging from semantic browsing to folksonomy enrichment, to question answering on the Sematic Web. This is indeed the most exciting aspect of this research: thanks to the features provided by Watson, entirely new applications – the ones that we have not yet envisaged – are now just waiting to be developed.