SEKT - researching **SE**mantic **K**nowledge **T**echnologies.

Knowledge discovery

Knowledge discovery is concerned with techniques for automatic knowledge extraction from data. It includes areas such as data and text mining, machine learning, statistics and parts of artificial intelligence. In SEKT these techniques will be mainly applied to textual data and in particular for (semi) automatic ontology generation, ontology evaluation, metadata extraction and ontology population.

Human language technology

An understanding of the structure of human language, and of its syntax and semantics, will be used to complement the statistically-based knowledge discovery techniques. Taken together, knowledge discovery and human language technologies make for a powerful approach to ontology learning and metadata extraction.

Ontology management and reasoning

Ontologies evolve over time, as the underlying knowledge evolves. This evolution needs to be managed. Moreover, not everyone will use the same ontology to describe a particular domain, so we need to be able to mediate between ontologies. For really intelligent search engines, we need to be able to reason using ontologies. In particular, robust reasoning will be needed to cope with the inconsistencies to be found on the global Web. SEKT will research state-of-the-art techniques to tackle all these requirements.

These technologies represent key components of what is increasingly being known as the *Semantic Web*.

minded

SEKT is a European research and development project launched in January 2004 with a lifetime of three years. The total budget is €12.5 M.

SEKT receives a contribution of €8.3 M from the European Commission as part of the Sixth Framework Programme (DG Information Society, strategic objective: semantic-based knowledge systems). SEKT also receives a contribution of €0.4 M from the Swiss government.

The SEKT consortium has 12 partners, led by BT. The consortium includes systems integrators and software developers, as well as leading European universities and research institutes.

THE SEKT PARTNERS

- British Telecommunications Plc
- AIFB, University of Karlsruhe
- Empolis GmbH
- Jozef Stefan Institute
- University of Sheffield
- iSOCO S.A.
- Ontoprise GmbH
- Kea-pro GmbH
- Sirma AI Ltd
- Universitat Autonoma de Barcelona
- Universität Innsbruck
- Vrije Universiteit Amsterdam

SEKT is a member of the SDK project cluster, which seeks to strengthen European research and industry capability in semantic web technologies.
www.sdk-cluster.org

FOR MORE INFORMATION

SEKT website
sekt.semanticweb.org

Dr. John Davies
SEKT Project Director
email: john.nj.davies@bt.com

Prof. Dr. Rudi Studer
SEKT Technical Director
email: studer@aifb.uni-karlsruhe.de

Paul Warren
SEKT Project Manager
email: paul.w.warren@bt.com

SEKT
The next generation web

sekt.semanticweb.org

THE RESEARCH CHALLENGES

SEKT is extending technical boundaries in a number of important ways:

• SEKT is paving the way for the introduction of semantic web technologies, in particular by researching the key technologies supporting the use of ontologies and ontologically-represented knowledge. Ontologies make possible an understanding of a domain which can be shared between people, or between computers to create a semantic web. SEKT is developing techniques for the (semi) automatic generation of ontologies and extraction of metadata. Once this has been done, reasoning can be undertaken to make deductions from the ontology and the associated metadata.

• The ontologies used by SEKT will provide much richer descriptive frameworks than the strictly hierarchical taxonomies which the most advanced current commercial knowledge systems use.

• The use of ontologies, and the ability to reason with ontologies, are the key components of the next generation of semantic knowledge technologies. SEKT is developing techniques for creating and managing ontologies, populating them with metadata, and reasoning with them.

• Although there will always be times when human interaction is needed, SEKT will enable a much greater degree of automation than current systems employ.

SEKT research is not just about advances in basic technology. We will be investigating how the user best interacts with knowledge, not just at a computer terminal but also via a PDA or mobile phone. SEKT will also be looking at how knowledge technology is best used in the organisation, and at how an ontological framework and the associated knowledge can best be presented to the user.

for the
technically –

knowledge
at your fingertips

SEKT's vision is to make knowledge truly accessible. The Internet constitutes the largest body of knowledge ever to have existed in the history of mankind, with corporate intranets adding even more. However, knowledge is only valuable if you can find it and make sense of it. That means it must be placed in a descriptive framework.

The Internet is full of unstructured text – fine for people, but useless for computers. As a result, when you search for information, you typically receive half a million hits, many irrelevant to your needs. SEKT is developing automated techniques for extracting meaning from the Web. By generating structured descriptions of Web pages, SEKT is making those pages machine-processable. So that in the future you will be able to get straight to what you want.

Searching intelligently

When you use a search engine on the Internet or a corporate intranet, it has no understanding of the context of our search. For example,

you might key in 'European growth rate', but the search engine has no idea what kind of 'growth' you are interested in: population growth, economic growth or some other measurement. With SEKT you can establish a context for your search.

Given semantically-annotated knowledge, SEKT software will exploit class hierarchies and attribute information to give a more intelligent response. For example, if you were looking for information about cars with certain properties (e.g. having the colour red), but nothing was available, SEKT could report the fact but also return information about red lorries, since it knows

that cars and lorries are both types of vehicle. Similarly, SEKT could search for employees with specific attributes (e.g. over the age of 45) by exploiting extracted metadata.

Avoiding the information deluge

How about if you wanted to search the Web for the speeches and writings of a famous politicians? You would need to use a text-based search, which will inevitably deluge you with redundant information about the politician. With SEKT you can specify explicitly that you want material written by the politician, not information about him or her.

You will be able to specify that you want information written by a particular individual, about a particular topic, and containing a particular phrase or set of phrases.

Finding and sharing

SEKT isn't just about finding knowledge, it is also about sharing knowledge. Setting knowledge within a descriptive framework means it can be targeted at those who really need it. As a result, knowledge sharing between colleagues becomes a really valuable activity.

Knowledge from different sources can also be integrated, and used to provide intelligent advice to support business and professional goals.

case

SEKT has three case studies which will test the technology and enable feedback into research. These case studies will see SEKT applied to three major languages – English, Spanish and German.

Part of the user's environment

To be really useful, SEKT needs to be an integral part of your working environment. SEKT is being designed to be seamlessly integrated into proprietary business software, and to support a variety of devices, such as PDAs and mobile phones.

The SEKT philosophy

SEKT works with, not in place of, the user. No really satisfactory techniques exist for fully automating the process of describing documents. However, SEKT makes suggestions which can be accepted or overruled, giving the users the ability to specify how automated the process should be. The SEKT goal is to put the users in control, whilst freeing them from routine tasks to concentrate on value-creation.

The SEKT architecture

Development within SEKT will conform to a SEKT architecture.

So, all SEKT components will inter-operate, whether produced by SEKT's industrial or academic partners. This architecture will be published, so that other developers will also be able to develop inter-operable SEKT-compliant software. To encourage take-up from vendors and users, some key components of the SEKT infrastructure will be open source. This will encourage other developers to use the SEKT framework as a basis for their own software.

The SEKT architecture will build on agreed international standards, and in turn the SEKT consortium will seek to influence emerging standards for the Semantic Web. By acting together in the SEKT consortium, and also by working with other European 6th Framework projects in this area, we can ensure that the European voice is heard as the next generation of the Internet evolves.
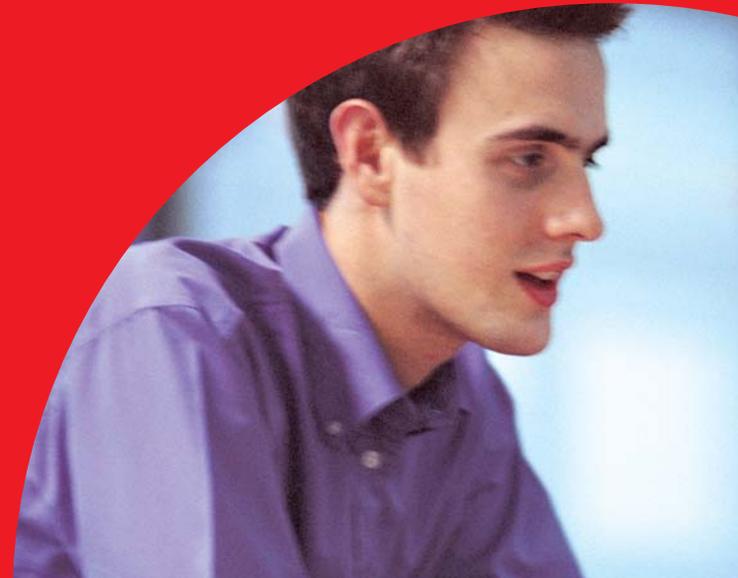
studies

Helping newly-trained judges in Spain

In Spain, as the world over, newly appointed judges are faced with complex decisions, and for assistance usually fall back on their former training tutor, a more experienced judge. This inevitably involves delay. SEKT seeks to help by providing precisely the information the newly appointed judges need.

Helping IT consultants serve their customers

Like all customer-facing personnel, IT consultants have a lot to learn from each other. Knowledge gained on one project needs to be shared with colleagues working on other projects. Often this doesn't happen and instead the knowledge is lost.

SEKT technology will help bridge the gap between the consultants' personal knowledge space and the organisational knowledge space. So that they can better serve their customers.

The digital library of the future

BT makes extensive use of its digital library, to provide electronic access to valuable information right across the company. This uses conventional textual searching to provide access to its 5 million documents. SEKT will give users a more powerful window into the library, and will also encourage digital library users to add knowledge to the library, and to share knowledge, all within a common framework.