



Federal Ministry
of Economics
and Technology



THESEUS

New Technologies
for the Internet of Services



Innovation Policy, Information Society, Telecommunications

The THESEUS Research Program

New Technologies for the Internet of Services

www.theseus-programm.de

Credits

Compilation and design

THESEUS Pressebüro, c/o LoeschHundLiepold Communication

Print

Richard Thierbach, Buch- und Offset-Druckerei

Photo credits

BMWi, Corbis, Getty, Fotolia, Masterfile

A.Weber/PT/DLR, THESEUS Innovation Centre, mediaglobe

Publisher

Federal Ministry of Economics and Technology
Public Relations
10115 Berlin

In collaboration with:

THESEUS Program Office
c/o Attensity Europe GmbH
Meisenstraße 90
33607 Bielefeld

Last revised

September 2011



The Federal Ministry of Economics and Technology has been awarded the berufundfamilie® audit certificate for its family-friendly HR policy. The certificate is granted by berufundfamilie GmbH, an initiative of the Hertie Foundation.



Federal Ministry
of Economics
and Technology



THESEUS

New Technologies
for the Internet of Services

Innovation Policy, Information Society, Telecommunications

The THESEUS Research Program

New Technologies for the Internet of Services

Table of Contents

Preface	05
1 THESEUS – New technologies for the Internet of Services	06
2 THESEUS Joint Research	09
3 THESEUS Innovation Center – Bringing the Internet of Services to life	10
4 ALEXANDRIA – A knowledge platform on the Internet	12
5 CONTENTUS – Technologies for the media library of the future	14
6 MEDICO – Intelligent searches of medical databases	16
7 ORDO – Order in a digital world	18
8 PROCESSUS – Making better use of corporate knowledge	20
9 TEXO – Infrastructure for Web-based services	22
10 THESEUS Basic technologies	24
11 THESEUS SME 2009 – Application projects for small and medium-sized enterprises	26
12 Partners in the THESEUS program	33

Preface

Our economy derives its strength from its capacity for innovation. Information and communication technology (ICT) plays a central role as a driving force behind innovation for the entire German economy. Today, the availability of cutting-edge ICT decides which locations will prosper and which will fail in the global competitive arena. Therefore, our aim is to actively encourage the development of these key technologies and occupy leading positions worldwide in the field of ICT.

In the THESEUS technology program, some 60 partners are developing and testing new technologies for a future Internet of Services. For example, on online marketplaces it will be possible to find suitable service offerings quickly, and combine these with proprietary or other services as needed. This will give rise to completely new business models. The opportunities for growth of this forward-looking market are immense, as demonstrated by the rapid growth of cloud computing we are currently witnessing. It is forecast that revenue alone from software as a service from the cloud will be in the region of €11 billion in Germany by 2025. Small and medium-sized enterprises, in particular, are set to benefit from this new approach.

The THESEUS program was launched in 2007 and has now entered the finishing straight. With a program budget of roughly €200 million, and some €100 million in funding from the Federal Ministry of Economics and Technology, THESEUS is the German government's biggest ICT research project. Declared a flagship project by industry and policy-makers at the National IT Summit, THESEUS is a priority area in the government's "Digital Germany 2015" ICT strategy.

THESEUS combines scientific excellence and business potential. Leading research institutes are collaborating with universities and enterprises within the THESEUS project. With the special THESEUS SME program, THESEUS also focuses specifically on small and medium-sized enterprises (SME) and offers SMEs in Germany the opportunity to test THESEUS technologies and use them to create new products and processes. Early integration of scientific results into their business operations makes these companies more competitive, while also encouraging emulation and multiplier effects. This promotes innovation,



growth and employment in Germany as a business location. THESEUS SME extends the research program, bringing in 30 new partners, most of them SMEs.

The THESEUS Innovation Center for the Internet of Services brings the results of THESEUS to life. Visitors can take a closer look at the research results and test the new services and tools for themselves.

This brochure presents the latest results from THESEUS SME and provides an overview of the entire THESEUS program. I can but urge you to learn about THESEUS, take advantage of its many excellent results and become involved in the program.

Sincerely yours,

Federal Minister of Economics and Technology

1

THESEUS – New technologies for the Internet of Services



Every day the Internet provides new information. As the amount of data grows, however, the challenge is to find ways of providing users with simple and efficient access to this enormous store of knowledge. Partners from the academic sphere and the business community are working together within the framework of THESEUS, currently Germany's largest IT research program, to meet this challenge.

Just as the legendary figure of Theseus in Greek mythology succeeded in escaping from the Minotaur's labyrinth, the research program of the same name has developed ways to navigate through the increasing quantities of data found on the Internet. Under the THESEUS umbrella, some 60 research partners from academia and the business world have come together to develop new technologies and applications. Their goal is to facilitate access to information, combine data to form new kinds of knowledge and lay the groundwork for new services on the Internet.

The technologies being developed within the THESEUS program are preparing the way for a future Internet of Services. This will make it possible for services that are now available on the Web only separately, such as online shopping, flight bookings and research support, to be combined and linked with one another. If a user were then to tell his computer: "I want to move from Berlin to Hamburg," the program

would identify and coordinate the appropriate resources for finding housing, organizing the move and registering with the authorities.

The first step: Semantic technologies

To make this vision a reality, THESEUS is developing new semantic technologies that will be able to analyze, classify and link information based on content. Not only will computers find information with the help of key words or content fragments, as they do today, but they will also determine its significance, draw connections to other information, model it as a classification system and employ certain rules to draw logical conclusions.

THESEUS researchers are working on innovative procedures for automatically producing metadata. Research and development are also under way on other basic technologies for processing multimedia documents more quickly and designing new graphic user interfaces more efficiently.

Trials in application scenarios

To ensure that new products and services are created as quickly as possible with the help of the basic technologies, the partners participating in the project are putting them to the test in the following six application scenarios. These scenarios show how the technologies

can be used for innovative tools, services and business models.

In the ALEXANDRIA application scenario, the Web 2.0 approach – which involves linking users in virtual communities and attaching greater importance to user-generated content – is expanded to include innovative approaches to knowledge management. For example, a search can be easily run on historical events or current affairs based on famous people, specific developments or interesting places. With ALEXANDRIA, users can more easily access content on the Internet to suit their individual needs.



CONTENTUS uses semantic technologies to digitalize cultural assets in the form of texts, images and audio and video recordings. The goal is to better preserve cultural assets for future generations and to facilitate swift and comprehensive access to this knowledge through a variety of media. For example, CONTENTUS provides important technologies that are being used in establishing a German Digital Library (Deutsche Digitale Bibliothek, DDB), a joint initiative of the Federal Government, Länder and local authorities. The goal of the DDB is to provide universal access to the resources of over 30,000 cultural and scientific institutions through a national online access portal based on the most current semantic methods.

Within PROCESSUS, researchers are exploring how the knowledge available in a company can be used to better advantage. To that end, a platform is being developed that integrates resource planning and business process management while also making it easier to compare products, solutions and business partners. As a result, companies can reduce costs and optimize all of their business processes.

MEDICO improves the quality of medical diagnoses with the help of imaging procedures. Semantic technologies help physicians recognize abnormalities in medical images quickly and confidently, and improve their diagnoses and therapeutic decisions by rapidly analyzing image databases and specialized literature.

ORDO concentrates on developing innovative approaches and technologies for organizing digital information, on both the Internet and company

intranets. Its research is geared to developing tools that organize information automatically, making it easier to handle and analyze large quantities of data. Anyone who works with large, difficult-to-manage datasets will benefit from such tools.

To ensure that these resources on the future Internet of Services can be used in an open, yet secure environment, researchers involved in the TEXO application scenario are developing an integrated platform for providing, managing and combining Internet-based services.

Promoting Germany as a business location

The research conducted as part of the THESEUS program is creating new technologies for the Internet of Services, and particularly for mapping complex knowledge. The participating researchers are making it easier to use the knowledge available on the Internet and are laying the groundwork for the development of top-quality services, ranging from innovative semantic searches for the future digital library to new service offerings such as cloud computing. In this way, Germany is not only strengthening its position in the area of Internet research, but is also ensuring that German companies will help to shape the Internet of Services by providing innovative products and services, and will benefit from this market of the future.

The first success stories of this strategy can already be reported. Although not even two years old, the research program has already spun off four startup companies: The software enterprise SemVox develops



user interfaces to simplify human-computer communication. TIQQER offers a Web service for analyzing Internet forums in order to rate products, for instance. Innoraise provides a flexible search application for social networks, enabling the automatic identification of experts on specific subjects, while Original 1 develops methods to make it easier to recognize counterfeit products.

When utilizing the results of the THESEUS program, a primary goal is to provide open interfaces and to develop standards. To this end, ORDO has initiated the “Semantic Information Logistic Architecture” (SMILA) open source project that provides the developer community with possible solutions – which can be tested and refined – for designing information logistics in service-oriented architectures.

With the “Extensible Multimodal Annotation” (EMMA) project, THESEUS has helped develop a new standard for multimodal input and output. EMMA enables the development of Web applications which users can operate by text, speech or gestures, and control with a wide range of computer equipment.

The “Integrated Service Engineering-Workbench” (ISE) open source project began as part of the TEXO application scenario for the simple and efficient development of Internet-based services. Furthermore, with the “Universal Service Description Language” (USDL), TEXO has also proposed a standard language to describe services, allowing them to be clearly identified on the Internet, and easily linked and combined.

The results of THESEUS research have already won awards at high-profile international competitions, such as ICDAR and Image Cleef.

A program that integrates different stakeholders and fosters international collaboration

The THESEUS program includes talented young scientists, outside experts, entrepreneurs and, in particular, small and medium-sized businesses in the research activities so they can take advantage of the research results at an early stage.

THESEUS is also involved in European and international initiatives geared towards shaping the Internet of the future. For example, THESEUS played a key role in the European Commission’s CHORUS program, which aimed to coordinate and combine European projects and initiatives focusing on multimedia search technologies. The international THESEUS symposium – “Internet of Services” – received global attention and underlined Europe’s strong position in the field of semantic technologies. In this way, THESEUS has had a significant bearing on the European Commission’s “Future Internet Public Private Partnership” initiative.

THESEUS was launched at the end of 2007 for a term of five years. It is receiving some €100 million in funding from the German Federal Ministry of Economics and Technology. Another €100 million is being contributed by partners from the spheres of industry and research, making a total of about €200 million that is being invested in this groundbreaking research work, which will produce technologies to benefit the Internet of Services and Germany as a business location.

2 THESEUS Joint Research



To a greater extent than other branches of technology, Internet technologies are developed and applied within global infrastructures. The task of THESEUS Joint Research is to put the research and development work that is taking place within the THESEUS program into an international context, to encourage a flow of information between the THESEUS consortium and external partners, and thus to promote the dissemination of the technologies and applications being developed by THESEUS.

The German Federal Ministry of Economics and Technology has charged the Joint Research program with focusing on the following three areas: profiling, networking and disseminating results.

Profiling involves the systematic collection of information about research areas and the development of technology, application and business models for a variety of international research and development activities. Such activities include, for example, the European technology platform “NESSI,” the large-scale Japanese project “Information Grand Voyage,” as well as industry-run initiatives such as service platforms.

The resulting profiles make it possible to link partners in the THESEUS consortium with scientific institutions and companies, by providing valuable information about shared development goals and

about complementary or competing interests. This allows them to engage in results-oriented planning, identifying areas in which cross-project and cross-company cooperation is feasible and mutually advantageous. In addition to linking individual partners, Joint Research also organizes events such as cross-sectoral workshops and international conferences.

Success in transferring results to industry, as well as to external research projects, depends on reconciling the needs of potential users with what the THESEUS consortium is able to provide. Joint Research supports this process, particularly by engaging in an ongoing exchange of information with user groups, regional service networks and industry-run standards groups. This effort, too, is based primarily on the results of profiling as well as on numerous closed and open events and a variety of publications.

3 THESEUS Innovation Center – Bringing the Internet of Services to life



The THESEUS Innovation Center for the Internet of Services in Berlin seeks to bring THESEUS research to life. Located in the former Siemens building, the center boasts 500 square meters of exhibition space where visitors can get a hands-on look at various interactive applications and prototypes, and experience the results of THESEUS research for themselves. Furthermore, the center also has facilities for presenting, discussing and testing the results of research into the “Internet of Services”.

Research results at your fingertips

What are semantic technologies? What solutions are the THESEUS researchers working on in the different application scenarios? How are research results quickly translated to marketable products? Answers to these questions and many more are given in the exhibition area of the innovation center, which provides insight into the THESEUS research program, explains how the individual projects are interconnected, and demonstrates the specific benefits of the research for ordinary citizens.

Various sample applications and demonstrations show how semantic technologies and innovative services will shape our everyday life in the future. The exhibition focuses on the areas of multimedia, knowledge, services, business processes and medicine, and presents groundbreaking solutions from THESEUS research for each of these fields.

The exhibits represent the entire spectrum of the THESEUS program and clearly illustrate the opportunities the research project gives Germany as a center for science and business. New research results will be constantly added to the exhibition and presented to the public in this way.

Collaboration and dialog

In addition to the exhibition area, the laboratory is one of the central elements of the THESEUS Innovation Center for the Internet of Services, and presents different examples of the specific work performed by the numerous researchers involved in the THESEUS project. The laboratory area is designed as an open



workplace to foster the exchange of knowledge between the THESEUS researchers and to support collaboration with outside experts. At the same time, visitors can also watch the scientists at work and get an insight into the practical side of the Internet of Services.

Furthermore, the Innovation Center also functions as a venue for events centered on THESEUS research and the Internet of Services. Topics such as the semantic Web, knowledge management and cloud computing, as well as new business models for small and medium-sized enterprises, are proposed and discussed in presentations, panel discussions and conferences. In addition, special topic weeks, workshops and guided tours are also on offer for various groups, including school-goers, junior scientists, entrepreneurs and the media.

The THESEUS Innovation Center thus promotes knowledge and the exchange of information about the Internet of Services as a key emerging area. Moreover, the center actively supports the integration of

science, business and politics with the aim of strengthening Germany's position in the field of information and communication technology.

The THESEUS Innovation Center is a "Selected Landmark 2011" in the Land of Ideas, and therefore one of the winners of the "365 Landmarks in the Land of Ideas" competition held as part of the "Germany – Land of Ideas" nation-branding initiative.

THESEUS – Innovation Center for the Internet of Services

Adress: Salzufer 6 | 10587 Berlin

Telephone: 030 31002-400
E-Mail: innovationszentrum
@theseus-programm.de

www.theseus-programm.de/en/innovation-centre.php



4

ALEXANDRIA – A knowledge platform on the Internet



The city of the same name on Egypt's Mediterranean coast was home to the ancient world's most famous library. Even then, its precise cataloguing system facilitated access to knowledge. Inspired by this model, ALEXANDRIA researchers are working to develop a knowledge platform on the Internet that will use entirely new ways of interacting to facilitate access to knowledge.

The starting point for the ALEXANDRIA application scenario is the growing importance of Internet users in what is known as Web 2.0. In addition to using search machines, e-mail and chat services, users are actively making their own knowledge available on the Internet. Common examples of this new kind of interaction are blogs, the online encyclopedia Wikipedia and social networks like Facebook and Xing. These platforms would be inconceivable without users who provide and exchange knowledge and information.

Collecting knowledge and making it accessible

ALEXANDRIA is seeking a better way of linking the knowledge that is available in Web 2.0 and working

to make that knowledge more readily accessible. To that end, a knowledge platform is being developed that not only collects information, but also helps to manage steadily growing quantities of data. This knowledge platform will help users publish, process and search for content. In order to make the platform intuitive and easy to use, researchers are developing semantic technologies and classification systems that use so-called metadata to link even large quantities of data in a logical way.

Comprehensive access to knowledge

To demonstrate the possibilities offered by this type of knowledge platform, ALEXANDRIA is initially focusing on "history and current events" to illustrate how the application works. With the help of an easy-to-navigate user interface, a demonstration shows how easy it is to access knowledge using newly developed technologies.

Instead of using key words to search for information, as we do today, the user can ask the system a question in normal language. For example, if asked

which 20th century Olympic athletes were natives of Berlin, the system is able to provide the relevant names as well as additional information about their lives and athletic achievements. By relating various categories of results – such as persons, places or events – to one another, the system is able to understand human knowledge and reproduce it in an organized fashion.

ALEXANDRIA also helps to display knowledge so that it can be easily visualized. Carrying out a search for information on the Olympics, for example, the knowledge platform combines data on past and future summer and winter games with geographical and scheduling information. The user is then able to find the various Olympic sites on a world map, navigate on a time line through the milestones of Olympic history, from the ancient world up to the present day, and access other information such as the winners in specific disciplines.

Integrating users' knowledge

Users determine which categories and types of information should be highlighted, by evaluating the relevance of specific data on a regular basis. They can also provide a detailed assessment of other information that is made available to them.

During the entry process, ALEXANDRIA indicates which portions of a query are automatically recognized. This allows the user to make corrections and narrow down the query. For example, if the user wants to know where Neil Armstrong was born, ALEXANDRIA automatically asks whether the American astronaut or the Canadian athlete is meant.

Questions that the system is unable to answer automatically are passed on to other users who have the necessary information, and they are asked to contribute relevant facts and multimedia files to the platform, leading to an ongoing expansion of the knowledge base. Such contributions can also be made using natural language. For instance, the system automatically recognizes the sentence “Neil Armstrong was born in 1932” and incorporates that information into the knowledge base.

To make users an integral component of the ALEXANDRIA platform, the system learns which user has what information and to what extent each user is contributing to the knowledge base. This also makes it possible to bring users together, based on their shared interests, to form knowledge communities.

Mobile access to knowledge

ALEXANDRIA is to be made available through mobile devices such as mobile telephones, cameras or GPS equipment, so that users will have access to the platform at any time and place. When a user looks at an object through the camera of his mobile telephone, ALEXANDRIA automatically superimposes semantic content on the picture, taking into account the user's location and the direction in which he is looking. Historical photographs might be displayed on a current photo of a building, for example, so that the user has a visual experience of traveling through time. Additional information and linked content can be accessed simply by touching the screen. The user's experience is therefore no longer limited to “searching,” but also involves “browsing” in semantic contexts.

Objectives of ALEXANDRIA

- ▶ Creating an easy-to-use, intuitive Internet-based knowledge platform
- ▶ Actively involving users and creating a cross-linked knowledge community
- ▶ Combining manual and automatic metadata production
- ▶ Developing new approaches to navigation and interaction
- ▶ Making it possible to use mobile devices

5 CONTENTUS – Technologies for media archives of the future



Germany's libraries, museums and archives contain an incredible wealth of knowledge in the form of millions of books, images, tapes and films. Researchers involved in CONTENTUS are exploring how these cultural assets can be made available to as many people as possible and preserved for future generations.

One of the major challenges facing the knowledge society is to provide easy access to knowledge and cultural assets through multiple media channels. Within the CONTENTUS project, ideas and technologies are being developed for an infrastructure that allows cultural institutions and information providers to work toward that goal. With the help of these new technologies, large quantities of data can be automatically processed and semantically linked, whether in the form of texts, images or video and audio recordings.

The work of CONTENTUS is closely coordinated with the German government's "German Digital Library" initiative. A comprehensive description of content using meta data – which are comparable to an index or table of contents for finding digital media – plays an important role in this context. Other challenges are linking the various digital knowledge records on the Internet and integrating user-generated content. To that end, researchers are developing intelligent algorithms that automate the necessary processes and allow users to add their own information to existing

digital content. This results in interlinked, next-generation multimedia archives.

Next-generation multimedia archives

These new types of archives cross-link the collections of traditional libraries, media archives and broadcasting stations to create a new information structure, bringing together providers and users via the Internet. Users are being offered new opportunities to contribute their knowledge to existing multimedia resources. At the same time, the data are linked semantically – based on their meanings and relationships – to form a new kind of knowledge network. This allows providers to optimize the structure and quality of their multimedia collections.

One example demonstrates the benefits of such next-generation multimedia archives: A search for "Lucullus" yields results in a variety of media, including texts, images, films and audio recordings. Thanks to the use of semantic technologies, it also shows content that is logically associated with that topic. Thus the search might show scores or recordings of "The Condemnation of Lucullus" as well as other works by the composer Paul Dessau or the author Bertolt Brecht, for example "The Good Person of Szechwan." Upon request, it will display information about the great interpreters of Brecht's works, or provide links to music companies that sell Brechtian compositions online.

While it generally takes a long time to find specific works or collection pieces in traditional archives, the technologies employed allow for efficient searches and to take advantage of semantically related results across multiple media formats.

Advantages for users and providers

With CONTENTUS, users will find it easier to search and navigate through digitalized cultural assets. It will be a simple matter for them to share results and insights with other users and providers, as well as to add their own information. With the help of technologies for quality control and semantic linking of multimedia resources, smaller libraries and archives, too, can be part of an interlinked information structure and make their multimedia collections available to a wider audience.

Research and development in the service of knowledge

The preparation of cultural assets for semantic multimedia access and online presentation can be broken down into six processing steps of considerable technical complexity.

Digitization of media - especially material threatened by decay like old magnetic tapes - is an essential step that influences the outcome of all subsequent tasks. Careful handling of fragile media and familiarity with digitization technologies are essential to obtaining good results. CONTENTUS and its project partners have compiled best practice guidelines for the digitization of media assets.

After digitization, the quality of the digital copies has to be analyzed and optimized. A wide variety of flaws can be automatically detected and removed. This comprises both flaws that result from defects of the original medium (scratches, stains, etc.) and quality problems that are introduced during digitization (e.g., dropouts in videos).

A content analysis step generates additional metadata based on the actual media content. For example, individual articles on newspaper pages are separated, as well as shots, scenes and key frames in videos. In addition, both printed and spoken text is transcribed

and within it entities of interest (Persons, Places, and Organizations) are identified.

The information gained in the previous step is then used to semantically link digital media and their metadata together with information from external sources such as Wikipedia or Geonames. This leads to an enriched knowledge network.

This network of knowledge can be opened for users to contribute further information or even add new items, making it possible to involve users with different levels of expertise in the accumulation and organization of knowledge.

Finally, the resulting knowledge base can be accessed using a semantic multimedia search interface, which enables users to intuitively navigate the information. Semantic technologies are used not only to rank search results but also to allow refined searches based on context information. As illustrated by the “Lucullus” search example, users are given suggestions regarding additional information that is related to the search term – so that they can explore the topic further, if they choose. This final step completes the process, resulting in cross-linked, multimedia access to information – a milestone in the development of a Web-based knowledge infrastructure.

Objectives of CONTENTUS

- ▶ Allowing easy access to cultural assets
- ▶ Developing concepts and technologies for next-generation multimedia archives
- ▶ Developing a complete, modular workflow from the analog medium to a semantic multimedia search
- ▶ Automating key steps such as quality control and media content analysis
- ▶ Creating new semantic and multimedia search possibilities
- ▶ Integrating user-generated content

6 MEDICO – Intelligent searches of medical databases



Imaging procedures, such as ultrasound and computed tomography (CT), are a critical part of medical diagnosis and therapy, as they make it possible to detect illnesses at an early stage and initiate the right treatment. Up to now, however, no single instrument has been able to intelligently structure all relevant information – written materials, images, laboratory data – and make it accessible. Researchers involved in the MEDICO application are working to close this gap.

How can a computer learn to interpret images and recognize similarities, as well as to relate them to text-based content? These are the questions with which the THESEUS application MEDICO is concerned. Using semantic technologies, the participating researchers are developing applications for simple, cross-linked searches of medical databases. The aim is to facilitate the work of physicians and other healthcare workers by intelligently compiling all of the relevant information on patients that is derived from image- and text-based findings.

New possibilities for diagnosis and treatment

MEDICO is developing procedures that not only recognize anatomical structures, such as bones, blood

vessels and organs, but also automatically classify data and combine reference images and treatment reports from various different databases, making it possible to identify pathological changes.

The benefits of MEDICO can be seen, for example, in the diagnosis and treatment of lymphomas: A lymphoma patient is being treated at a local hospital. The attending physician wants to know whether chemotherapy is working. The algorithms developed by MEDICO researchers automatically compare current CT images with those taken prior to treatment, yielding the relevant information: Chemotherapy appears to be working, since the lymph nodes have not increased in size. However, the program detects damage to the spleen, raising new questions. The attending physician has no comparable cases on site that might provide useful information. To clarify the patient's situation, he takes advantage of another feature of MEDICO that compares his patient's data with information on numerous lymphoma cases treated at a university medical center. This much larger set of data provides the attending oncologist with specific recommendations for further tests and treatment.

Using the THESEUS application ORDO, the physician is also able to conduct a quick search of current

scientific literature that is relevant to the illness of this particular lymphoma patient. The most current research results help to determine the optimal approach to treatment.

Support for physicians and patients

The intelligent interpretation and retrieval of a wide variety of images and texts offer a multitude of opportunities for the medical sector. Attending physicians are provided with an efficient and precise tool to help them reach decisions. Patients also benefit, since their physicians are able to base their diagnostic and therapeutic decisions on a wide variety of experience. Moreover, MEDICO will save hospitals time and money by avoiding tests that are unnecessary and often difficult for patients.

The researchers involved in MEDICO are initially concentrating on semantic searches and structuring the information contained in medical image databases. However, the potential applications of MEDICO go much further. As soon as ethical and data protection issues have been resolved, epidemiological researchers and the pharmaceutical industry alike will benefit from quick, efficient access to relevant studies, which can further improve the quality of diagnosis and treatment.

Structuring data appropriately

The purpose of the technologies developed through MEDICO is to enable computers to generate information on their own, using the contents of the images displayed; to combine that information with other medical data; and to structure the results to provide recommendations for the attending physician. This requires the development of numerous components, for example pattern recognition methods and clinical decision-making tools. The central task, however, is to model an appropriate ontology – a system for organizing and cross-linking copious amounts of information, based on medical terminology. Here MEDICO can build on ontologies that have already been developed by medical experts, such as “RadLex,” which characterizes medical images, or the very detailed “Foundational Model of Anatomy,” which – as the name indicates – describes human anatomy.

As methods of diagnosis and treatment improve, physicians will have access to more and more information about patients and their illnesses. MEDICO quickly and easily interprets, links and compares medical information in image and text form, making an important contribution to medical progress.

Objectives of MEDICO

- ▶ Intelligently researching and structuring medical information such as images, reports and texts, and providing information for use in a variety of different healthcare situations
- ▶ Providing help in making diagnostic and therapeutic decisions
- ▶ Providing more rapid and efficient access to medical knowledge in research and in day-to-day hospital routines

7

ORDO – Order in a digital world



Human knowledge is available in many forms, with information stored in a variety of data media and formats as texts, images or video and audio recordings. Researchers involved in the ORDO application scenario are developing a system that will make it easier to manage this diversity.

Modern information technology makes our lives easier, but sometimes more confusing as well. This can be seen, for example, in the development of new medications. Today, pharmacologists seeking information about active ingredients find that the search can be a complicated and drawn-out process. They may find chemical formulas in specialized databases and information on how medications work in pharmaceutical archives, while finding a complete list of medications with illustrations may require consulting a manufacturer's marketing materials. In addition he receives feedback on the effects of the relevant medication via discussion forums, weblogs and social networks.

A better way to organize knowledge

To make searching for information easier and more efficient, ORDO researchers are developing new services and software tools to help users organize and prioritize contents and search results. The idea is to provide users with all of the relevant data and help them distinguish quickly between important and unimportant information and recognize how one thing relates to another.

New technologies, for example, will allow pharmaceutical researchers to link searches in different databases and sources with a central query. Search processes are easily coordinated and simplified. The computer may search the database for a chemical formula and will find the corresponding molecular diagram in an entirely different source. Consumer opinion about a medication is determined by linguistic analysis of contributions in discussion forums or weblogs.

All of this information is clearly presented to researchers so that they can organize the search

results appropriately and select certain items about which they want to learn more. This overview of the available information is automatically generated. Procedures are also being developed to process large quantities of data and display them in graphic form. The result is comprehensive, individualized knowledge management that consistently allows the user to find the right information.

Practical testing in the context of research and development

ORDO's research results are being tested specifically to determine whether they meet the needs of research and development departments, such as those in the chemical and pharmaceutical industries, which need to handle large amounts of data. In addition to primary data, such as patents, a great deal of secondary information – research reports, market analyses and other academic publications – needs to be analyzed as well. E-mails, archives, in-house wikis and blogs also have to be processed. ORDO's researchers are developing methods and technologies that will make it possible to use all of these data resources more efficiently and link them with Internet content.

Managing information using semantic technologies

With ORDO, computers will soon be able to identify concepts and relationships in text documents, extract facts and generate summaries automatically. Relative to existing recognition methods, this represents a major step toward more efficient information management. In the future, these technologies will also be used for analyzing and organizing multimedia content.

This ambitious project is based on the enhancement of semantic technologies that enable a computer to automatically recognize data content and meaning, and to put data into context. This is particularly important when organizing and analyzing unstructured data. Another focus of ORDO is to develop efficient procedures for creating classification systems (such as ontologies) for modeling, managing and organizing data. With the help of innovative linguistic and statistical methods, these ontologies can be used still more effectively, and in some cases they may even be generated automatically.

An open, efficient platform as a basis

Different components to collect and process the data will be needed to develop the content. These can be generally used components such as a text analysis or bespoke components such as identification of the chemical structure. A powerful and efficient platform is needed to allow an optimum combination of these components. The open source project SMILA – Unified Information Access Architecture – is developing such a platform. ORDO is making a substantial contribution to the development of SMILA, which is placed at the disposal of interested users as a basis for their own solution free of charge in an eclipse Foundation project.

The key is personalization

To achieve truly individualized knowledge management, it is important to take into account personal information and perspectives – for example, specific work areas and locally used computers. The kinds of information that are relevant to a project manager, for example, will differ from the information a member of the project staff might need. Preparing and organizing data in a personalized way will make it possible to give due consideration to the individual needs and preferences of each user.

Objectives of ORDO

- ▶ Semantically organizing and prioritizing information
- ▶ Recognizing concepts and relationships in texts, extracting facts and generating summaries automatically
- ▶ Visualizing how different kinds of information relate to one another



8 PROCESSUS – Making better use of corporate knowledge



In an environment of global competition to produce innovative goods and services, companies are faced with the challenge of making the best possible use of their knowledge. Within the PROCESSUS framework, a solution is being developed, with the help of semantic technologies, to ensure that employees have access to the right information at the right time. This will lower costs and optimize operations.

Knowledge is crucial to the competitiveness of German industry. In practice, however, the information that many companies possess is not adequately linked, making it difficult to access. One reason is that companies use a wide variety of databases and software systems (ERP, PLM, CMS systems, etc.) that are unable to communicate with one another. In addition, critical information is often hidden within internal documents – texts, contracts, analyses, plans, records and e-mails – so employees have only limited access to this valuable resource. PROCESSUS is developing a platform to combine the various types of information a company has in its systems, based on the significance of and relationships between data. This makes isolated bits of corporate knowledge more transparent.

Traditional software systems are capable of analyzing documents, but they cannot put them in a broader context. PROCESSUS, as a semantic service platform, will make it possible to use corporate information to support dynamic processes. It will link the contents of software systems and documents semantically. Where information is stored will no longer be a concern for users, who will be able to click on a link to find the right information at the right time.

Practical testing in the fields of mechanical engineering and information technology

The concepts and solutions developed in PROCESSUS are being tested in the fields of mechanical engineering and information technology. In mechanical engineering, the focus is on applications related to drive systems and automation technology. The mechanical engineering sector, in particular, involves highly complex processes and diverse user groups. A typical process chain consists of preliminary development, construction, production, sales and customer service. One and the same product is dealt with in very different contexts as these steps are carried out. Construction specialists are generally

concerned with technical information, while their colleagues in the sales department are more likely to focus on market and competitive analyses. PROCESSUS links the data contained in the various software systems, based on their significance. This gives employees exactly the information they need – in the context of the respective process step.

In such situations, PROCESSUS will make it possible to access information in an entirely new way: Users describe their problem, and while processing the query, the computer considers solutions that have proved successful in a similar context. This makes users aware of aspects and opportunities they may not have considered, such as the idea of using existing products in other markets. For example, a PROCESSUS-supported search might alert a member of the sales staff of a company that manufactures paint production systems to the fact that a machine component, such as a gripper arm or a labeling unit, might also be used in the production of bottling systems.

In the area of information technology, these insights can be applied to the linking of diverse computer applications, or so-called service-oriented architectures. Within the framework of the PROCESSUS use case, an infrastructure is being developed for managing content in the software industry. This classifies software components according to their ability to solve problems, and uses this information for drawing up technical documentation. The PROCESSUS platform combines the semantic information in the user's query with information and text phrases contained in the knowledge base, and dynamically produces a document with the necessary information.

These two examples demonstrate how the technologies developed through PROCESSUS make it possible to link industry and process knowledge more quickly and to develop new products, solutions and business models. This offers companies a competitive edge, since they can respond more quickly because their employees have access to knowledge wherever they need it, and companies can save money on research and development. This also gives companies an advantage in terms of quality, since the information they receive meets an especially high standard.

Consequently, PROCESSUS helps companies make the best possible use of their knowledge so that they can better meet the challenges of global competition.

Synergy with TEXO and ORDO

The results of PROCESSUS offer great potential, and this is particularly true when they are combined with the technologies developed through the TEXO and ORDO applications. Because of their modular structure, the components of PROCESSUS can be deployed as independent services on TEXO or ORDO platforms. The information provided by PROCESSUS is accessible individually and in a specific context, making it possible to take full advantage of a company's internal knowledge.

Objectives of PROCESSUS

- ▶ Expanding conventional applications such as document- and content-management systems to include process orientation
- ▶ Simplifying access to a company's process-relevant knowledge
- ▶ Reducing costs and increasing efficiency through the use of semantic technologies



9

TEXO – Infrastructure for Web-based services



While it has become commonplace to sell content such as music and videos on the Internet, Webbased services are not yet as widely used. The TEXO application is developing an infrastructure that will make it easier to combine and utilize Internet services, an important step toward creating an Internet of Services.

Many consumers and companies find searching for services on the Internet to be tedious and time-consuming. Exacerbating the problem is the fact that complex processes often require different service providers for different tasks, and there is no logical way of combining these tasks.

Foreigners seeking to open a business in Germany, for example, quickly discover how difficult the process can be – the choice of a location, arranging to be entered in the commercial register, a market analysis and the hiring of suitable personnel consume both time and resources. Smaller and medium-sized companies in particular would find it very helpful if all of these steps could be carried out through one central partner. But which providers can offer all of the necessary services? And how does one go about finding them?

A basis for an innovative service economy

The TEXO service marketplace is developing an integrated platform to solve these problems. Using the Internet, it is easier to advertise, combine, utilize and pay for services, and it is increasingly feasible to tailor services to the needs of the customer (whether an individual consumer or a company). This is an important step toward a new service economy on the Internet, in which services are traded just as goods are today.

A variety of benefits for providers

Particularly for small and medium-sized companies, offering services through the Internet brings with it substantial opportunities. They can advertise their services worldwide and attract new customers. They can also expand their range of services by linking them to the services of other providers, which allows them to open up new markets by offering innovative service packages. TEXO is creating the necessary technological conditions for accelerating the development of new services and facilitating the integration of services from different providers. This allows small and medium-sized companies to concentrate on their areas of expertise while saving time and

money – since they have the option of outsourcing tasks outside of their core business to external service providers.

New opportunities for customers

Both consumers and companies can benefit from the new TEXO service platform, as illustrated by the following example.

Using TEXO in connection with vehicle insurance claims

In 2007, €42 billion was spent to process nine million vehicle damage claims. Such claims involve a large number of entities – injured parties, insurance companies, repair shops, appraisers, damaged-vehicle exchanges, automobile manufacturers, attorneys and IT companies. By providing information about prices, availability and expertise and helping to link services appropriately, the TEXO service marketplace makes it possible for all of these parties to work together effectively. This expedites claims and makes it easier for customers to understand the process, as well as sparing both insurance companies and the insured unnecessary time and effort for processing and following up on claims.

The basis: SOA

Semantic technologies are being used to develop a standard for implementing the TEXO system, as a way of describing services and related information flows so that services can be more easily utilized and linked. This process employs service-oriented architecture (SOA) to group individual applications into a single user application. Building on standards to describe technical services in a service-oriented architecture, an integrated service description language is developed in TEXO which makes it possible to describe the technical, business-specific and operational aspects of services.

Objectives of TEXO

- ▶ Creating a service platform on the Internet
- ▶ Making it possible to trade services like goods and combining services to create new service packages
- ▶ Creating an open and secure environment for buying and selling services on the Internet
- ▶ New business models, particularly for SMEs



10 THESEUS Basic technologies



There is more knowledge on the Internet than in all of the archives and databases in human history. The THESEUS program is developing the basic technologies and standards necessary to make this knowledge more widely available in the future.

THESEUS focuses on the development of semantic technologies for capturing the meaning of information. This allows computer programs to automatically analyze the contents of texts, images and sound and video recordings, link them and draw logical conclusions.

In order to make this evolution of the Internet possible, it is essential to develop technical standards and innovative basic technologies, and these technologies are being created and tested by the partners in the THESEUS consortium in six application scenarios. The researchers are also exploring how these technologies might be used, as soon as possible, for creating new tools, services and business models on the Internet.

An overview of these basic technologies:

1. Generating metadata automatically

THESEUS researchers are developing new methods of comprehending media contents, based on metadata. Metadata contain information about an item included in a database, for example the name of an author or the period in which a film takes place. These new methods make it possible to create metadata for various types of media content, including texts, photographs and audio and video files. One focus is to find ways of generating metadata automatically. Another involves the use of semantic technologies for understanding the context of media content. The goal is to take similar contents from a variety of sources and put the information together in a grouping that provides the user with as much information as possible.

2. Rapid processing of multimedia documents

To avoid unnecessary delays as users wait for the results of a search of a complex multimedia database, THESEUS

is developing highly efficient algorithms for producing metadata. They make it possible to search images and videos quickly, even in databases containing several hundred thousand items. To enhance image searches, researchers are also developing image recognition systems that will allow computers to identify objects in a photograph or video. Another focus area is data compression, particularly for image files.

3. Innovative ontology management

THESEUS is exploring semantic technologies based on so-called ontologies, which enable computers to “understand” the meaning of content. Ontologies are formal knowledge models that conceptually represent the knowledge within a given subject area and make it possible to process that knowledge automatically at the level of meaning – something that, thus far, only human beings have been able to do. The THESEUS working group on ontology management is developing methods for improving the design and development of ontologies and enhancing automatic reasoning through their use.

4. Machine learning

THESEUS is also working on intelligent data analysis processes that facilitate automatic recognition of data relationships and interconnections so that they can be modeled and structured, much as is done with the help of ontologies. These methods are being applied to texts, images and audio and video data, and they help identify relationships between different types of data.

5. Situation-sensitive dialogue processing

Before a computer can act on behalf of a user, it needs to understand what the user wants. To facilitate this dialogue between humans and machines, researchers in the THESEUS program are developing new functions that can be deployed in different applications. Innovative algorithms make it possible to create multimodal user interfaces that can be controlled using speech, gestures and other inputs. Such interfaces allow users to formulate their queries intuitively and refine them through spoken dialogue with the system. A special component within the computer serves as an interface between the multimodal user interface and the

various sources of metadata. It transforms a spoken query into the semantically appropriate data record required by the system for running a search.

6. Innovative user interfaces

THESEUS is also developing new graphic user interfaces to make it easier to identify the relationships between data, metadata and documents. For example, the results of a query can be presented in the form of a knowledge network, which shows how the search results are related to the search term as well as to one another. This provides users with a clear overview of the topic at hand and helps them to find the information they need more quickly.

7. Evaluating basic technologies

Experts are assessing the quality of the basic technologies developed within the framework of THESEUS. New technologies for speech and image recognition and for the automatic classification of metadata are being tested to determine their reliability, functionality and suitability, in an effort to ensure that the research meets quality standards. The results of this evaluation are also taken into account in the research and development process, helping to further optimize end results.

11

THESEUS SME 2009 – Application projects for small and medium-sized enterprises



Small and medium-sized enterprises (SMEs) are at the forefront of many fields of technology, since they are particularly quick and flexible when it comes to opening up new markets. However, SMEs often lack sufficient time and resources for research and development. This is where the “THESEUS SME 2009” competition to promote small and medium-sized enterprises comes in.

The THESEUS SME 2009 competition offers German small and medium-sized enterprises an opportunity for early testing of technologies developed through THESEUS, and allows them to use these technologies to generate new products and processes. The goal is to develop new services and business models. The competition seeks to encourage top-level research by small and medium-sized enterprises. Early integration of scientific results into their business operations makes these companies more competitive, while also encouraging emulation and multiplier effects, and allows other companies to take advantage of these results more quickly. This promotes innovation, growth and employment in Germany as a business location.

THESEUS SME 2009 expands the THESEUS research program by adding 12 application projects, selected by an independent jury, and bringing in 30 new partners, most of them SMEs. The SME initiative is receiving some €10 million in support from the German Federal Ministry of Economics and Technology, and the various projects have been assigned to work with existing application scenarios.

Objectives of THESEUS SME 2009:

- ▶ Greater involvement of SMEs in THESEUS research results
- ▶ Early transfer of scientific findings to medium-sized businesses
- ▶ Expanding the THESEUS program by developing and testing innovative technologies for new products and processes
- ▶ Enhancing the impact and sustainability of the THESEUS research program

The topics and participants involved in these twelve projects are as follows:

“B2B in the Cloud – Making SME business relationships easy”

Karlsruhe research centre, SEEBURGER AG

Large enterprises already rely heavily on the Internet to conduct business with customers and suppliers, and capitalize on the potential savings and increased output afforded by automated processes, efficient data transmission, short response times and low cost. Up to now, however, the costly initial investment, the expense of hardware procurement and maintenance, and the complex administration involved have prevented many small and medium-sized enterprises (SMEs) from benefitting from the multiple advantages business-to-business (B2B) integration has to offer.

B2B in the Cloud intends to make it significantly easier for SMEs to take advantage of B2B solutions, backed by the ability to provide standard descriptions for B2B services, trade the services on the Web, combine them to create value-added services and incorporate them into the customer’s individual application environment.

This will be centered on the development of an Internet-based platform that makes all the necessary steps involved in the exchange of business data as easy, standardized and inexpensive as possible. The vision involves an innovative evolution of the software-as-a-service (SaaS) principle toward the more demand-oriented use of external B2B services with pay-per-use pricing – similar to an electricity bill. This enables flexible cost models for SMEs, making B2B solutions significantly more attractive for the midmarket. The aim is to give SMEs greater access to B2B solutions and thereby make them more competitive on the global marketplace. As efficient IT helps drive down administrative costs, SMEs stand to benefit all the more: compared to large businesses, administration accounts for an above-average share of total costs in SMEs.

“ChemProspector – Intelligent searching for chemical structures in documents”

InfoChem GmbH

Researchers and developers in chemical and pharma-

ceutical companies need to retrieve all the information available on existing patents as quickly as possible. If patent information is available on time, such companies can avoid expensive investments in the wrong products and reduce financial and legal risks. This is why the ChemProspector project focuses on the automatic extraction of generic structural formulae from chemical and pharmaceutical patent documents.

The chemical structures hold the key information. When filing patents, companies use “Markush structures” – with a central core and variable groups – to increase the scope of their claims and cover many different potential compounds. Since this information can appear in the graphics, body of text and image captions of the patent documents, the process of clearly identifying already patented substances, active ingredients and substance classes is very complex and time-consuming.

The aim of ChemProspector is to make it possible to sift through and identify the generic chemical structures (Markush structures) in documents more quickly and comprehensively. This applies both to core structures from graphics and to generic elements from text passages or image captions. The system automatically exports the chemical structures, merges them into a logical data structure and saves this information. It is also possible to search through proprietary documents and external data sources via a Web-based research platform.

A demo system is already available: A Web-based research platform allows end users (researchers, patent examiners, patent agents etc.) to run fact, full-text and structure searches conveniently across different data sources, such as the patent data of the United States Patent and Trademark Office, and delivers clear results in the form of structural formulae.

“GoOn – Semantic search platform for life sciences”

antibodies-online GmbH, RESprotect GmbH, Technical University of Dresden, Transinsight GmbH

Antibodies play a central role in both biological research and medicine. Given the diverse range of products and antibodies available, however, the antibody market is broad and lacks transparency. For example, no standard naming and spelling conventions apply to the labeling of genes and proteins. Day after day, doctors and researchers therefore face the

problem of being unable to select suitable antibodies quickly and efficiently to treat patients or for their research work. GoOn provides a search platform that makes it possible to retrieve and systematically organize information on antibodies. It can extract related information from specialist literature and link data intelligently. The use of ontological background knowledge is a central goal behind the development of this new semantic search platform. This knowledge enhances the search quality as searches no longer focus solely on keywords. The search technology developed in GoAntibody is designed to retrieve and analyze information on specific antibodies and helps deliver transparency to the market. GoDisease, on the other hand, systematically analyzes information on the known interaction of genes and proteins to provide new insights into the treatment of certain diseases.

The technologies developed in the GoOn project have already been incorporated into real products. As part of the semantic search platform, www.GoPubMed.com, thousands of users benefit from the new services. Ten years online, GoPubMed is the world's first semantic search engine on the Internet and the leader in its field. Science and industry equally benefit from the integration of GoOn applications since researchers can search complex contextual relationships between proteins and genes, and products can be advertised correctly in connection with scientific texts.

“GREEN Mobility – Easy access to new location-based services”

B2M Software AG, HTTC e. V., raumobil GmbH, Yellowmap AG

Carpool arrangements organized through a Web-based lift-sharing agency can be quite inconvenient. Green Mobility delivers a solution for comprehensive and dynamic mobility services on mobile devices. To this end, Green Mobility links a range of external carpool services advertised on the Web with traffic jam updates, routing services and information from the yellow pages: Green Mobility suggests the optimum meeting point for the driver and passenger at the start of the trip and even identifies possible places the passenger can take shelter while waiting for the driver in bad weather. If the driver is delayed in traffic and one of the waiting passengers is in a



hurry, Green Mobility can suggest the best alternative by checking whether the passenger could consider another lift-sharing option or use public transport.

Green Mobility also demonstrates how services from the Internet of Services can complement one another and be combined to create a single innovative mobile application. The development of a central mobility platform – the Mobility Mediation Layer (MML) – enables optimized access to the Internet of Services via mobile terminals. The MML middleware makes specific functions for mobile usage available as ready-to-use modules, which significantly drives down the cost of developing an application. This access to the Internet of Services will play a key role in the flexibility of future business apps. Furthermore, the MML technology is also suitable for different scenarios on mobile terminals and offers key improvements, such as reduced data volume for transmission, semantic methods for automatic service replacement and a push service for a wide variety of mobile devices.

“HIPPOLYTOS – An easier way to use geographical and environmental data”

disy Informationssysteme GmbH

The vast amounts of geographical and environmental data available are often too confusing and complex for the layman. This is set to change with HIPPOLYTOS. By enabling user-friendly access to heterogeneous databases – as we know from Web searches – even non-professionals should be able to use geographical and environmental information quickly and effectively. Semantic technologies are the key, allowing users to ask the system simple questions via a search

field. For example users could ask “How has air pollution in Berlin changed in the past ten years?”. By linking relevant data intelligently and presenting the information in an organized manner, the system delivers clear search results that even laypersons can comprehend.

When used as a planning base for environmental protection, renewable energy or large-scale building projects, HIPPOLYTOS delivers real added value by intelligently linking geographical and environmental data. HIPPOLYTOS can also make a key contribution to integrating environment-related disciplines – such as ecology, environmental medicine, agricultural science or earth science – at the national and international level by making it easier to retrieve and use environmental data.

Interim results of the HIPPOLYTOS project include a term-based search function for a reporting and evaluation system for factual and spatial data, and a live preview function for the graphic presentation of results (such as thematic maps). Furthermore, a module has also been developed for the semi-automatic creation of a semantic index for environmental data analyses which will make it easier to find data.

“jCPEX! – A platform for cross-company business processes”

Metasonic AG

In today’s business environment, business processes go beyond company borders, with interconnected, cross-company processes now the norm. While a business collaborates with its suppliers, it also maintains close contact with its customers as a service provider. To enable the different parties to coordinate their business more effectively and easily, jCPEX! is developing a platform for planning and implementing business processes across companies. The creation of an RfQ, for example, involves many individual steps, as a business must make arrangements with its suppliers and its customer. With jCPEX!, this task is easier since the platform interconnects such cross-company processes. The individual companies can identify the process interfaces which they find interesting, adapt them individually, and link them to other interfaces. Specifically, the platform gives small and medium-sized enterprises (SMEs) the

opportunity to offer their services globally, systematically develop and expand their service portfolio, and present their business expertise to a wider potential audience.

This flexible system of loosely connecting business processes via the jCPEX! platform gives businesses added agility and the ability to implement changes within just a few hours at minimum expense.

“MachInNet – Better leveraging of technological knowledge in businesses”

CIM Aachen GmbH, EXAPT Systemtechnik GmbH, RWTH Aachen

Many small and medium-sized enterprises (SMEs) have in-depth specialized technical knowledge. At the same time, however, they often lack expertise in related technological fields which could be used to optimize processes further within the businesses. For example, a manufacturer of turned parts has expert knowledge of precision turning techniques. If, however, additional milling operations are required, the manufacturer cannot deliver optimum results as the company lacks the specialist knowledge. Even the repetition of production steps that were performed years beforehand with an older machine can suddenly pose a problem.

The knowledge needed can often only be found in the machine’s “cryptic” numerical control (NC) program. Data cannot, however, be recovered from this program.

The Machining Intelligence Network (MachInNet) project therefore focuses on developing a semantic Web-based knowledge network for SMEs. The goal is to help businesses to manage the wide range of technologies and significantly accelerate industrial engineering processes. Using the NC program and other tool process descriptions it should be possible for businesses to make explicit use of implicit technological knowledge via a reference database. All the SMEs in the MachInNet network feed information into the reference database, which then delivers possible solutions to specific problems using special algorithms and fuzzy searches.

MachInNet factors in both the needs of the manufacturing industry, which benefits from the broad knowledge base when optimizing its own production processes, and the needs of the tool production

companies, which can offer their customers significantly better data supply services. The MachInNet solutions have already been trialed by six pilot teams – two tool manufacturers and four tool users – with successful results. Interested companies can use the MachInNet solution either in part or as a complete service.

“mediaglobe – The digital archive”

Medien Bildungsgesellschaft Babelsberg gGmbH, Flow Works GmbH, defa-spektrum GmbH, Hasso-Plattner-Institut für Software-Systemtechnik GmbH

Contemporary history has been documented and recorded more and more for many years. It follows that Germany’s 3,000 to 4,000 media and broadcasting archives have also seen an enormous increase in the amount of audiovisual material stored. Managing this mostly analog information poses a serious challenge. Only a fraction of the material can currently be searched through with digital technologies. Consequently, the potential this material offers to our information and knowledge society remains largely untapped.

The “mediaglobe – the digital archive” project prepares the way to the digital age for both providers and users: cultural institutions and media archives can record, manage, digitalize and offer their audiovisual material without infringing copyright law, while individuals can access the data stored in multiple locations via the Web, search through the material and use it.

In order to retrieve the content, it must be possible to run a smart semantic search through the vast data repositories. For this reason, mediaglobe focuses on the optimization of the entire digitalization workflow and the automated analysis of the AV material as this generates the metadata required for the search.

As many archives have a tight budget, the technologies developed will be launched as a scalable system so that every archive only has to purchase the particular functions it needs. Smaller establishments can also opt for external hosting to avoid complex administration work. Furthermore, mediaglobe also offers consulting services for AV digitalization and rights management to ensure every establishment can reap the maximum benefit from the project. Mediaglobe has already



established a partnership with the German National Library of Science and Technology (TIB) in Hanover. In addition, a guide to the digitalization of audiovisual archives will be published in autumn 2011 and deal with the very complex topic of digitalizing analog AV material and the need for extensive advisory support on digitalization projects.

“openXchange – Easier damage claims processing using service networks”

Metris GmbH, University of Stuttgart

If property is damaged because a pipe has burst, for example, insurance companies need to perform a quick and in-depth assessment of the damage to ensure optimum claims settlement. Furthermore, any repair work must be assigned to damage remediators quickly and efficiently.

Large insurance companies recognized the potential of electronic claims processing a long time ago and have invested in the development of proprietary solutions. It is, however, only feasible for very large service providers to connect to the insurance companies’ proprietary networks. Smaller businesses are put off by the high investment costs and are therefore increasingly elbowed out of the market.

openXchange is specifically designed to tackle the competitive disadvantage facing small and medium-sized trade businesses. By running the service network on an open platform, all the parties concerned can take part in electronic settlement processes without the need for major investments in IT or specialized knowledge. The IT support makes a significant contribution to simplifying coordination processes

and driving efficiency. Loss adjusters have an insight into how insurance cases are assigned in practice, while the insurance companies can track and ensure the quality of loss adjuster services. The service network platform is specifically designed to help analyze and optimize existing settlement processes and develop quality models. Furthermore, the project will provide standards for data exchange formats and interfaces for the integration of openXchange into existing applications, and develop auxiliary reference processes and communication strategies. Large insurance companies can run openXchange solutions independently while smaller establishments can use openXchange as a SaaS solution.

“RADMINING – Better access to radiological images and texts”

Averbis GmbH, Freiburg University Hospital,
Leipzig University Hospital

The bureaucratic burden in the medical sector is growing. Given the plethora of reports and diagnostic records, the volume of medical texts and images is constantly increasing. To manage this unstructured data growth more effectively, experts from the fields of radiology, IT and computational linguistics have joined forces in the RADMINING project to develop a novel radiological knowledge system that optimizes processes and helps doctors compile clinical findings. Patients also benefit from RADMINING as diagnoses are faster and more reliable, and unnecessary radiation exposure can be avoided. This helps drive down health care expenses over the medium term by reducing the number of costly and unnecessary examinations.

Three initial RADMINING prototypes are currently being tested in participating hospitals in Freiburg, Leipzig and Erlangen. One such prototype makes it possible to link information in text and images to run a semantic search in the database. This enables doctors to verify their presumptive diagnoses and ensure faster and more reliable diagnoses. Another RADMINING prototype compares examination requirements against current guidelines to check whether a particular examination is the most appropriate for an individual patient. This helps accelerate the diagnostics process and avoid exposure to radiation. The third prototype involves comparing radiological presumptive

diagnoses with data from other departments to eliminate any uncertainties in the diagnosis. This gives radiologists prompt feedback on the proposed medical treatment and makes a key contribution to improving patient care and patient safety.

SABINE – Semantic support services for professional integration and personal skills development imc AG, University of Saarland

Job specifications and skills profiles in the business sector are no longer aligned with education and training qualifications. This means that businesses tend to spend a lot of time searching for suitable staff, while individuals looking for a new job face a complex task. SABINE delivers an efficient solution to address the needs of both parties. The project aims to implement a Web-enabled application for the use of semantic technologies that provides customized support for employers, employees and intermediaries in recruiting processes and human resources development.

Specially developed Web technologies systematically interlink job placement, recruiting, skills management, learning and social networks. SABINE searches for, filters, assesses and links possible jobs, appropriate education and further training courses, tailor-made learning material, occupational integration opportunities and individuals.

SABINE closely compares the personal skills and professional goals of the applicants against the employers' job requirements and finds the best possible match between open vacancies and candidate profiles. This makes it easier to find and compare job advertisements and candidate profiles on the Internet and systematically link the information to personnel development opportunities. The system also integrates any suitable further training courses from other databases, such as the KURSNET pages of the Federal Employment Agency and open opportunities for further training in Web 2.0, rendering the training and education market more transparent.

The efficiency of SABINE services can already be tested on a demo system for job-seekers available at www.learning-demo.eu.

SERAPHIM – Service platform for mechanical engineering and plant construction

Infoman AG, Festool GmbH,
Ligmatech Automationssysteme GmbH

A machine failure poses a serious problem to plant operators. Outages can restrict production and trade, and cause a system shutdown in the worst-case scenario. Such situations demand quick and effective coordination and information processes between all parties concerned. This is why the SERAPHIM project focuses on developing a service platform that uses semantic technologies to optimize services and sales processes in mechanical engineering and plant construction. In this way, everyone involved is fully informed and valuable time is not wasted in critical coordination processes.

The necessary digital services are made available to the maintenance, repair and overhaul (MRO) partners on the platform through the development of services for a paperless repair process, for example. This enables the swifter exchange of information between all parties concerned, and downed equipment is back in operation sooner. Furthermore, the solution also helps avoid the excessive use of paper, a bonus for both the environment and the business balance sheet.

Industry can also benefit from a service configurator which SERAPHIM provides as a digital service. This configurator compiles all the relevant and updated information on service offerings for specific machines and ensures that manufacturers, staff, customers and retailers worldwide are always on the same page.


Thanks to SERAPHIM, coordination difficulties and problems in the flow of information in global service and after-sales processes will be a thing of the past. The project results are available in the form of specific demo systems and have been incorporated into the development of a service portfolio for manufacturing companies.

12

Partner in the THESEUS research program

- ▶ antibodies-online GmbH
www.antibodies-online.com
- ▶ Attensity Europe GmbH
www.attensity.com
- ▶ Averbis GmbH
www.averbis.de
- ▶ B2M Software AG
www.b2m-software.de
- ▶ CIM Aachen GmbH
www.cim-aachen.de
- ▶ defa-spektrum GmbH
www.defa-spektrum.de
- ▶ Deutsche Nationalbibliothek
www.d-nb.de
- ▶ Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH
www.dfki.de
- ▶ Deutsche Thomson OHG
www.technicolor.com
- ▶ DISY Informationssysteme GmbH
www.disy.net
- ▶ EXAPT Systemtechnik GmbH
www.exapt.de
- ▶ Festo AG & Co. KG
www.festo.com
- ▶ Festool GmbH
www.festool.de
- ▶ Flow Works GmbH
www.flowworks.de
- ▶ Forschungszentrum Informatik (FZI) Karlsruhe
www.fzi.de
- ▶ Fraunhofer-Gesellschaft
www.fraunhofer.de
- ▶ Hasso-Plattner-Institut für Software-systemtechnik GmbH
www.hpi.uni-potsdam.de
- ▶ Hessisches Telemedia Technologie Kompetenz-Center (httc e. V.)
www.campus-innovation.de
- ▶ imc information multimedia communication AG
www.im-c.de
- ▶ InfoChem Gesellschaft für chemische Information mbH
www.infochem.de
- ▶ Infoman AG
www.infoman.de
- ▶ Institut für Rundfunktechnik GmbH (IRT)
www.irt.de
- ▶ intelligent views GmbH
www.i-views.de
- ▶ Karlsruher Institut für Technologie (KIT)
www.kit.edu
- ▶ LIGMATECH Automationssysteme GmbH
www.ligmatech.com
- ▶ Ludwig-Maximilians-Universität München (LMU)
www.uni-muenchen.de
- ▶ Medien Bildungsgesellschaft Babelsberg gGmbH
www.mb-babelsberg.de
- ▶ Metasonic AG
www.metasonic.de
- ▶ Metris GmbH
www.metris.de
- ▶ mufin GmbH
www.mufin.com
- ▶ Neofonie GmbH
www.neofonie.de

- ▶ ontoprise GmbH
www.ontoprise.de
- ▶ raumobil GmbH
www.raumobil.de
- ▶ RESprotect GmbH
www.resprotect.de
- ▶ Rheinisch-Westfälische Technische Hochschule Aachen (RWTH)
www.rwth-aachen.de
- ▶ SAP AG
www.sap.de
- ▶ SEEBURGER AG
www.seeburger.de
- ▶ Siemens AG
www.siemens.de
- ▶ Siemens IT Solutions and Services GmbH
www.it-solutions.siemens.com
- ▶ Technische Universität Darmstadt
www.tu-darmstadt.de
- ▶ Technische Universität Dresden
www.tu-dresden.de
- ▶ Technische Universität München
www.portal.mytum.de
- ▶ Transinsight GmbH
www.transinsight.com
- ▶ Universität des Saarlandes
www.uni-saarland.de
- ▶ Universität Stuttgart
www.uni-stuttgart.de
- ▶ Universitätsklinikum Erlangen
www.klinikum.uni-erlangen.de
- ▶ Universitätsklinikum Freiburg
www.uniklinik-freiburg.de
- ▶ Universitätsklinikum Leipzig
www.uniklinikum-leipzig.de
- ▶ VDMA Gesellschaft für Forschung und Innovation (VFI) mbH
www.vdma.org
- ▶ Yellowmap AG
www.yellowmap.de



This publication is available free of charge as part of the public relations work of the Federal Ministry of Economics and Technology, and may not be sold. It may not be used by political parties or campaigners, or by electoral assistants during an election for the purposes of campaigning. In particular, it is forbidden to distribute this publication at campaign events or at information stands run by political parties or to insert, overprint, or affix partisan information or advertising. It is also forbidden to pass it on to third parties for the purposes of electoral campaigning. Irrespective of when, in what way, and in what quantity this publication reached the recipient, it may not be used even when an election is not approaching in a way that might be understood as suggesting a bias in the federal government in favor of individual political groupings.

