

***THESEUS's way from semantic core technologies to
innovative services***

Statement

Ca. 15 Min.

Dr. Andreas Goerdeler

Federal Ministry of Economics and Technology

on the occasion of
the opening of the W3C office at the DFKI in Berlin

on 10 February 2011

at 9.00 a.m.

at the THESEUS Innovation Center
in Berlin, Salzufer 6

Checked against delivery!

Embargoed until delivery!

Ladies and gentlemen!

1. Introduction

It is a pleasure for me to have the opportunity to welcome the World Wide Web Consortium (W3C) here today to the THESEUS Innovation Centre for the Internet of Services!

And I am pleased that the German Research Center for Artificial Intelligence (DFKI) is opening the W3C office for Germany and Austria, thus making it a pioneer in shaping the future internet.

This step shows that the DFKI can already point to significant successes in this field and I will return to that later.

Owing to its widespread links, not only with the research but also with the business community, it has become an important driving force for future web technologies, as for example Web 3.0.

Decision-makers in research, industry, administration and government, and active players of the W3C Community have been invited to attend today's opening event.

I'm particularly glad to see Dr. Jeffrey Jaffe from W3C, Prof. Cosnard from INRIA, Prof. Wahlster and Prof. Sasaki from the German Center for Artificial Intelligence.

Together with partners from science and business, the German Economics Ministry opened the THESEUS Innovation Center for the Internet of Services here about six months ago. Following this, we planned to launch a dialogue on innovations emerging from research and development in the cutting-edge field of information and communication technologies or, in short, ICT.

Today's event marks an important milestone in this process.

Under the THESEUS research programme, new technologies for the future Internet of Services are being developed and tested.

W3C, which is represented here today, has paved the way for submitting important research findings to a worldwide standardization process for the internet.

W3C currently has 325 members, including a number of well-known companies and scientific institutions from Germany. W3C and the traditional standardization organisations, such as DIN, DKE, ISO, IEC etc. do not necessarily have to opposites in this process. On the contrary: Many W3C standards are also integrated into the traditional standardisation processes as Public Available Specification or, in short, PAS. W3C thus serves as a PAS submitter.

In the age of the internet, however, the pace of standardization is important. In this regard, W3C is often able to respond very quickly to current trends and to provide the right guidance. The benefit for industry is clear. The aim is

to accelerate the development of innovation through early-stage standardization activities.

2. From semantic core technologies to innovative services within THESEUS

Let me now comment on THESEUS and first results.

THESEUS is a research programme supported by the German Economics and Technology Ministry, which aims to facilitate access to information, to link data to create new knowledge and to lay the foundation for the development of new services in the internet. In addition, THESEUS paves the way for applications in numerous business sectors. Support is being provided for 6 application scenarios in powerful sectors of industry and 12 additional SME-led application projects.

THESEUS is a lighthouse project of Chancellor Merkel's IT summit and is now developing into a lighthouse for the future Internet of Services. THESEUS receives approximately 100 million euros in public R&D funding from the German Economics Ministry. Additional 100 million euros are contributed by the participating partners from industry and research. At present, THESEUS is thus the largest ICT project of the German government.

The initial reason for launching THESEUS was the information overload on the internet. In recent years, this flood of information on the internet has not decreased. On the contrary, the number of websites has risen to about 200 billion. The storage space for digitally generated and processed new information in 2009 reached a record high of 800 billion gigabytes. And this was already a 62 per cent increase on the previous year. In 2010, this number went up to 1203 billion gigabytes. A 44-fold rise is expected by the year 2020. And according to current forecasts of the Euro-

pean Commission (EU), internet data traffic by 2020 will double every 11 seconds. In 2010, it doubled already every 11 hours.

We therefore need to find new ways to support people in analyzing information, which is initially available in an unstructured manner.

In this context, high hopes are pinned on the use of semantics. For example, the semantic description of data - while also taking into account the respective context - allows a better analysis of data. This way of linking data creates novel knowledge networks and more specific search paths. Upon this basis, entirely new and innovative services, products and business models can emerge.

Basic semantic technologies developed under THESEUS facilitate the analysis, classification and linking of written texts, images, speech, sounds, videos and web services

based on content. This enables computer programmes to find information not only with the help of key words or content fragments as they do today. Instead, they are able to independently identify the meaning of information, draw connections to other information, model it as a classification system and employ certain rules to draw logical conclusions.

THESEUS does not only use methods and standards as depicted in OWL ontologies and RDF. It also creates new classification systems, such as for example an ontology for mechanical engineering. This ontology is currently being tested by various companies with the support from the German Engineering Federation (VDMA).

Existing medical ontologies used in radiology, anatomy or to classify diseases (RADLEX, FMA, IDC) are employed and partly developed further. Another new innovation serves to describe web services. Not only an ontology has

emerged from this innovation but also a universal description language for web services - I am speaking of USDL.

THESEUS has proposed a standard language under the name “Unified Service Description Language” (USDL) to harmonize the description of services. USDL allows that services can be clearly identified on the internet, easily linked and combined. The main basis for this are the research findings of the TEXO application scenario.

The World Wide Web Consortium has approved the request by the THESEUS partners SAP, Siemens, Attensity Europe and DFKI to establish an Incubator Group.

THESEUS considers the Incubator Group as a valuable instrument of W3C. With this instrument, W3C can provide support, for example, in the technical field. I have the strong hope that if its work proves to be a success, the In-

cubator Group will make recommendations for developing a standard for the description of services on the internet.

I am pleased to say that THESEUS has already proved to be such a success. With its project "Extensible Multimodal Annotation" (EMMA), THESEUS has made a major contribution to shaping a new standard for multimodal input and output. EMMA, which was created at the DFKI, allows the development of web applications, which users can operate by handwritten text, spoken language or gestures, and which can be controlled by a wide range of different terminals. It was possible to standardize EMMA with the help of W3C.

Standardization through W3C results in disclosure of the respective standard, making it available royalty-free. So, it is not about proprietary standards. This is precisely why standardization is an important strategy to make use of publicly funded research findings.

THESEUS has already established the fundamentals for standards in 4 areas.

In addition, it is promoting the distribution of research findings in 2 areas via open-source platforms. Within the scope of the ORDO application scenario, which is led by Attensity, the open-source project called “Semantic Information Logistic Architecture” (SMILA) was launched. This project provides assistance to the developer community in shaping information logistics in service-oriented architectures.

The scientific success of semantic work is reflected by very good positions that have achieved in international competitions. In the area of semantic analysis of newspaper articles, THESEUS results scored better than Google (ICDAR competition); in the area of semantic image annotation it achieved third place (Image Cleef Competition). Semantic technologies have also proved to be very attractive in our own contest - the THESEUS ideas competition for young

talents. A competition contribution to semantically supported indexing even won the first prize.

New fields of business have developed on the basis of the application scenarios in areas such as medicine, mechanical engineering, media and business software. As early as after half of the programme period, THESEUS brought about 4 spin-offs, two thereof originating from the DFKI, also in the field of semantic technologies and Web 3.0.

The THESEUS co-ordinator Attensity is a German medium-sized company that successfully develops products and services on the basis of semantic technologies. Attensity meanwhile is expanding not only in Germany, but also internationally.

THESEUS also affects Europe. THESEUS participates in European and international initiatives to shape the Internet's future. The international THESEUS symposium "In-

ternet of Services”, for instance, which took place in the Federal Ministry of Economics and Technology in 2009, attracted international attention and underlined Europe’s strengths in the field of semantic technologies. THESEUS thus had a great impact on the European Commission’s initiative regarding the Internet’s future that was launched in 2010 as public private partnership project.

3. Prospects

The THESEUS innovation centre established in 2010 provides a stimulus for the dissemination of developments regarding the Internet of services. It aims to enter into a dialogue with future users at an early developmental stage. Here, the centre’s activities focus on small and medium-sized enterprises and start-ups. Within the framework of the Federal Economics Ministry’s start-up competition, the THESEUS consortium supports start-ups with a special award in the field of the Internet of services.

THESEUS is further developing in the direction of cloud computing. In this context, a new initiative taken by business, academia and policy-makers was supported. The Federal Economics Ministry in 2010 launched the Cloud Computing Action Programme and the Trusted Cloud Technology Competition for this purpose. According to forecasts of a study commissioned by the Federal Economics Ministry, the German market alone will perceptibly grow in the coming years in the field of “software as a service”. Turnover in the field of cloud standard software, for instance, is expected to rise to about 11 billion euros by the year 2025 - this corresponds to a share of 90 percent in total expenditure for standard software in Germany.

I hope that synergies between the newly established W3C office and the THESEUS innovation centre will also trigger new activities. This event today provides a first signal.

Since co-operation with the W3C office has the potential to expand. The W3C office offers valuable support to secure sustainability regarding the USDL standardisation efforts. And it will no doubt be useful to attract further active participants for the Incubator Group. We must aim to make the Incubator Group a regular W3C working group. In case of success, W3C will also continue to be sought after as PAS Submitter for other standardisation organisations.

Today the Web is no longer only an HTML platform. And here, the W3C is well positioned. On the basis of the successful work in the HTML and XML environments, trends were further developed at an early stage. This was recently the case with XML3D and the new 3-D technologies, which are increasingly gaining significance. The DFKI in particular has made a considerable contribution in this field.

In concluding my speech let me refer to the former US president Kennedy: “Alone, there is very little we can do. Together, there is little we cannot do.”

I wish the new W3C office in Germany every success in its important work and I am convinced that it will considerably help to shape the Internet’s future.