Core Technologies for the Internet of Services

Tilman Becker, Catherina Burghart, Kawa Nazemi, Patrick Ndjiki-Nya, Thomas Riegel, Ralf Schäfer, Thomas Sporer, Volker Tresp, and Jens Wissmann

Abstract Information and knowledge are growing permanently and represent valuable resources for many enterprises. The efficient access to knowledge of an enterprise like expertise, contact persons, project and milestone plans, etc. may simplify business processes and lead to time and cost savings. Semantic technologies offer numerous possibilities to enrich data with background information about their meaning. Such semantic relations do not only lead to more efficient search in larger information repositories but they also assist the user in diverse processes like editing, annotation and processing of information. In addition they offer new means of access and transfer of knowledge. Each information unit is linked to other units in the same domain, which allows faster search and offers a way of information access that is close to the habits of humans, i.e. the creation of new

---

T. Becker (✉)
DFKI GmbH, German Research Center of Artificial Intelligence, Saarbrücken, Germany
e-mail: tilman.becker@dfki.de

C. Burghart • J. Wissmann
FZI, Forschungszentrum Informatik, Karlsruhe, Germany
e-mail: burghart@fzi.de; wissmann@fzi.de

K. Nazemi
Fraunhofer Institute for Computer Graphics Research (IGD), Darmstadt, Germany
e-mail: kawa.nazemi@igd.fraunhofer.de

P. Ndjiki-Nya • R. Schäfer
Fraunhofer Institute for Telecommunications (HHI), Berlin, Germany
e-mail: patrick.ndjiki-nya@hhi.fraunhofer.de; ralf.schaefer@hhi.fraunhofer.de

T. Riegel • V. Tresp
Siemens AG, Munich, Germany
e-mail: thomas.riegel@siemens.com; Volker.Tresp@siemens.com

T. Sporer
Fraunhofer Institute for Digital Media Technology (IDMT), Ilmenau, Germany
e-mail: spo@idmt.fraunhofer.de

knowledge and its association to already existing knowledge. This approach of the whole THESEUS research program was also the basis of the CORE TECHNOLOGY CLUSTER of THESEUS.

1 Introduction

The CORE TECHNOLOGY CLUSTER (CTC) had been established in parallel to the different use cases in order to research and develop innovative technologies, which are common and required by several of these use cases in order to avoid parallel and separate developments. The CTC was subdivided into *Work Packages*, five of which (WP2–WP6) were dedicated to specific technologies and one (WP8) to the evaluation of these technologies according to international benchmarks such as ImageClef (Müller and Tsikrika 2010) or TRECVID (Smeaton et al. 2006):

- WP1 – "Program Management"
- WP2 – "Video, Metadata, Platforms – Processing of Multimedia Data"
- WP3 – "Ontology Management"
- WP4 – "Situation aware Dialog Processing"
- WP5 – "Innovative User Interfaces and Visualization"
- WP6 – "Statistical Machine Learning"
- WP7 – (canceled in the planning phase)
- WP8 – "Evaluation"

Accordingly, this article is subdivided into six sections, each of them dealing with a technology mentioned above. The first section deals with new methods for image and video processing, and for semantic search in multimedia archives. 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. Different tools for ontology design, for the mapping of different, heterogeneous knowledge structures, for the tracking of modifications (evolution) and for reasoning have been developed and will be explained in detail. The next section describes a dialog shell, which is an architecture and development platform for new means of interaction. A platform for multimodal and situation aware dialog systems has been created, whose architecture and some applications are described. A further core technology is semantic visualization, which is a graphical presentation of information to enhance decision making through visual processing of complex interdependencies. Last but not least machine learning methods have been developed, which are intelligent data analysis processes that facilitate automatic recognition of data relationships and interconnections so that they can be modeled and structured. These methods are applied to texts, images and audio and video data, and they help identify relationships between different types of data. The last section is dedicated to the evolution of the above mentioned technologies. Experts were assessing the quality of the basic technologies developed within the framework of THESEUS. The developed technologies have been tested to determine their reliability, functionality and suitability, in an effort to ensure that

the research meets international quality standards. The results of this evaluation were also taken into account in the research and development process, helping to further optimize end results.

The CTC started its work in February 2008 and the last activities ended in March 2012. The following companies and organizations were involved in the CTC (lead marked in bold):

- DFKI GmbH, German Research Center of Artificial Intelligence, Saarbrücken and Kaiserslautern (WP3, **WP4**),
- Fraunhofer Institute for Computer Architecture and Software Technology (FIRST), Berlin (WP6, WP8),
- Fraunhofer Institute for Telecommunications (HHI), Berlin (**CTC**, **WP1**, **WP2**),
- Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS), Birlinghoven (WP6),
- Fraunhofer Institute for Computer Graphics Research (IGD), Darmstadt (WP2, **WP5**),
- Fraunhofer Institute for Digital Media Technology (IDMT), Ilmenau (**WP8**),
- Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), Karlsruhe (WP3),
- FZI, Forschungszentrum Informatik, Karlsruhe (**WP3**, WP4),
- Ludwig Maximilian University, Munich (WP6),
- Siemens AG, Munich (**WP6**),
- Machine Learning Group, TU Berlin (WP6).

2 WP2: Video, Metadata, Platforms: Processing of Multimedia Data

The amount of digital images and movies is ever increasing. To effectively manage this flood of content in archives, new methods for image and video processing and for semantic search in multimedia archives are required. The Work Package 2 has therefore developed innovative solutions for the processing of multimedia content within the THESEUS CORE TECHNOLOGY CLUSTER.

An overview of the WP2 tasks is given in Fig. 1. It is agreed that manual annotation of digital images and movies would yield tremendous costs. Therefore, only a very limited amount of metadata is available for a vast proportion of audio-visual media.

To effectively manage this flood of content in archives, new methods for image and video processing and for semantic search in multimedia archives were required. With respect to related key challenges, the work in WP2 was structured around four major axes:

- Image and Video Recognition
- Perceptual Hashing
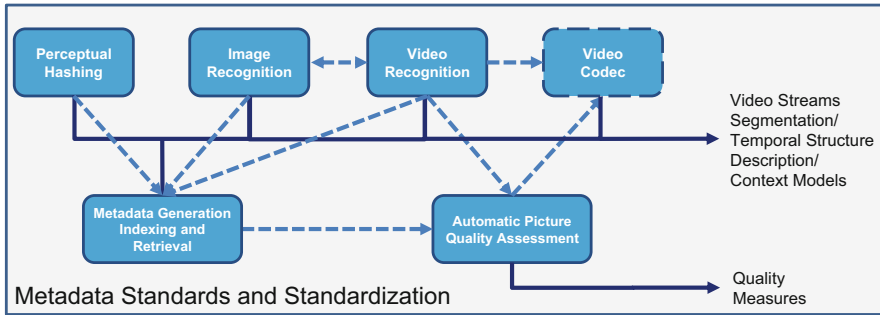
Fig. 1 Summary of the WP2 tasks

- Automatic Picture Quality Assessment
- Metadata Generation, Indexing and Retrieval.

Some peripheral activity was also conducted on video coding in order to generate transmittable video streams.

The *Image and Video Recognition* techniques aim at automatically extracting (vision-based) *metadata*, the fuel of any semantic search engine. Another kind of meta-information was explored in the *Perceptual Hashing* task, where the identification of several versions (e.g. image/video compression or changes in brightness) of the same content was the relevant use case. Automatic methods for the *Quality Assessment* of images/videos for their qualitative comparison were also addressed.

Searching videos and images in a semantic manner requires the consideration, utilization, and integration of the entire spectrum of the available information, from image analysis results to relevant context and background knowledge. To achieve this, a *modular framework for solving semantic queries* in videos and images has been developed.

Recognize and Classify Manual annotation of digital images and videos is extremely expensive. Often, only a very limited amount of metadata is available for audio visual media. Within WP2, one important goal was to develop image and video recognition techniques that are able to automatically extract important metadata. One of the fields focused on was semantic classification of images. The computer uses machine learning algorithms to learn which images belong to or contain given concepts, such as "outdoors", "vehicle", or "city." Afterwards, the trained system can be used to detect these concepts in new, previously unseen images. Besides these general purpose concept detectors, very specific model-based concept detectors were developed, e.g. for the detection of faces within images and videos. The proposed face detector is very robust with respect to lighting conditions. Besides the position of faces, it can also classify their pose, i.e. the direction of the face.

Different approaches can be used to transfer algorithms designed for images into the video domain. One possibility is to properly extract a set of (key-) frames

and use algorithms on these extracted frames (images). Within WP2, a temporal video segmentation algorithm was developed that is able to detect the temporal hierarchical structure of videos. It can extract scenes, shots, and sub-shots and is able to also extract key-frames from these temporal units. These key-frames are then utilized for further analysis or for the construction of a visual table-of-contents that assists the user in browsing and navigating through the video.

Enabling computer programs to automatically detect the content of a scene, the various settings and the genre of a film sequence was one further goal. In addition, mechanisms were investigated and implemented, which combine low-level features of the content (e.g. color, contour and motion) with information about the context and thereby derive complex relationships.

Effective Feature Storage and Indexing For search engines, two main development goals have to be reached: The results have to be of high quality and the amount of time needed to do the search and return results has to be low. The second goal can be addressed by the development of compression and indexing methods.

Compression methods are used to reduce the dimensionality of features, i.e. the amount of metadata per item to be compared during a search is reduced. This has to be done, however, without sacrificing too much in terms of search quality.

Indexing methods are the second way of limiting the amount of time needed for a search. A linear search for example has a time duration or hardware complexity that grows linearly with the number of items in the database. Some archives have grown to a point where this is infeasible. Therefore, algorithms are needed that perform faster than linear variants. In WP2, a tree-based indexing method was developed that has a logarithmic complexity during search, which means doubling the archive size does not double the time needed for a search. Metadata nodes are chosen that are representative of other metadata nodes. In the search, nodes can be skipped if their representative nodes are not similar enough to what is searched for. This way, only a small set of the metadata in the database actually has to be compared for a search. Most of it can be skipped which considerably reduces search complexity (Ciaccia et al. 1997; Skopal et al. 2003).

Framework for Semantic Search in Media Data Searching videos and images in a semantic manner requires the consideration, utilization, and integration of the entire spectrum of the available information, from image analysis results to relevant context and background knowledge. To achieve this, a modular framework for solving semantic queries in videos and images has been developed. It allows for exploiting the (inter-) relations between objects/events across multiple images, the situational context, and the application context, e.g. for identifying persons in video sequences by considering contextual information or detecting events in a video sequence without direct observation. This evidence-based detection/recognition is carried out using a novel calculus approach ("Subjective Logic"), which allows us to aggregate different evidence for and against a hypothesis, hence deriving a degree of belief for it. This is achieved by taking into account all the available evidence,

including the inherent uncertainties (accounting for the vagueness of image/video analysis results).

For a unified access to distributed and heterogeneous (multimedia) data resources, a middleware component has been developed. This "Query Broker" acts as a mediator between a client and multiple heterogeneous retrieval engines or database systems, encapsulating all retrieval functionalities in one component and exposing a standardized query format. The retrieval process is further supported by modules for intelligent query segmentation and distribution, as well as federated result set aggregation.

The Query Broker was designed to support the integration of heterogeneous knowledge bases into the global query evaluation process, enabling semantically enriched retrieval. In addition to features available in meta-search engines, the Query Broker also supports specific multimedia retrieval paradigms (e.g., query by example), as well as cross system multimedia retrieval (i.e. both cross metadata and cross query language). In this context, WP2 additionally promoted the international standardization of MPQF (at ISO/MPEG) on search support for semantic concepts and on JPSearch (at ISO/JPEG) through successful contributions.

Quality Assessment Quality assessment plays an important role in various image and video processing applications. Although human observers can effectively and easily judge the quality of images and videos, subjective measurements cannot be integrated into real-time image/video processing systems because of high runtimes and processing costs. Therefore, a lot of research has focused on objective metrics with the intention of addressing the drawbacks of the subjective metrics. Nevertheless, it is difficult for objective metrics to assess image quality in a manner similar to human perception.

Three measurement classes, i.e. full-reference (FR), reduced-reference (RR) and no-reference (NR), are proposed in the area of objective image quality assessment. The FR metrics predict the quality of an image based on differences to a reference image. Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structure Similarity (SSIM) are widely used FR metrics. On the contrary to RR metrics, which use indirect information as a reference, NR metrics predict the quality of images by extracting and modeling prior knowledge on distortions (Wang and Bovik 2005).

The quality of images and videos is evaluated and assessed by humans in a highly complex process within a split second. Mimicking this natural process using technology is one of the most difficult tasks in image and video processing. WP2 had the objective to develop automatic methods for the quality assessment of images, thereby facilitating, inter alia, the qualitative comparison of different images. The automatic metrics incorporate perceptual quality measures by modeling human visual system characteristics.

Copy Detection In order to identify several versions of the same content in archives quickly, perceptual hashing algorithms can be applied. These algorithms extract characteristic features from multimedia content and combine them to form a small

"digital fingerprint", the so-called perceptual hash. This hash value is robust against several modifications of the content, such as image/video compression or changes in brightness. This property is crucial, as it enables the detection of content after it has undergone processing steps which are usual in image or video processing. Various application scenarios exist for this method. In addition to the identification of content in archives, it is also possible to detect copyright infringements on the Internet, to monitor the broadcast of commercials, or to link the content with additional information (such as content owner or copyrights) stored in a database. Furthermore, overlapping sections of video snippets can be detected, which allows for a reconstitution of the full video.

The robustness against distortions distinguishes these methods from cryptographic hash values, like SHA-1 (National Institute of Standards and Technology (NIST) 2002). The objective of those methods is to detect if the content has been modified. Therefore, they are bit sensitive, meaning that the hash value changes completely when a single bit of the input data changes. In contrast, perceptual hashing algorithms should be robust against small modifications. Content originating from the same source should get the same or a similar hash value. This enables identifying copies of the same content based on the distance between the hash values. The advantage compared to watermarking techniques, which have similar application areas, is that the content does not have to be labeled before it can be detected. Therefore, it is possible to detect already distributed content. WP2 has developed robust perceptual hashing methods which enable the identification of images and videos.

3   WP3: Ontology Management

Research and design of semantic technologies based on ontologies is one of the most prominent research topics within the THESEUS research program. These technologies enable computers to "understand" the meaning of an object or symbols with the help of ontologies, which are knowledge bases of formal models conceptually describing domain knowledge. Automatic processing of ontologies allows semantic interpretations – a typical characteristic of a human being. Using methodologies for structural processing and logical deductions permits ontology-based information systems to achieve intelligent behavior: i.e. within THESEUS anatomical relations between organic structures are modeled in order to enhance search in medical information systems. Another use case supports knowledge management within production and distribution of goods by describing functions and characteristics of industrial automation.

Ontology management actually denotes the handling of knowledge bases. Knowledge itself is formally represented and modelled by using either first order logic or an ontology language: within THESEUS we have agreed to support ontologies modelled by one of the W3C standards OWL or RDF(S).

The goal of WP3 is to provide an *ontology management* (OM) toolbox to be used by use cases and other CTC work packages for the handling of ontologies and semantic metadata. Management of ontologies comprises four different tasks within the THESEUS research program: ontology design, mapping of heterogeneous knowledge structures to one another, tracking of changes to ontologies (evolution), automatic deductions (reasoning) and semantic search. Within these four tasks several components were developed which can be used by other partners of THESEUS to manage ontologies; some of the tools are available in different platforms like NEON, webGenesis, information workbench or SMILA. Several of these tools can also be found as open-source plug-ins on sourceforge.net.

We partition the field of OM into several subjects concerned with the maintenance and the processing of ontologies:

- *Persistence* In addition to programmatic in-memory manipulation, ontologies need to be stored in databases for their persistence and efficient access. In particular for large ontologies a database persistence layer is required to ensure their scalable handling.
- *Change Management* The changes made to ontologies over time need to be maintained explicitly in order to allow for features like versioning and intelligent update management. This also comprises ontology evolution techniques that take into account the semantics of changes, such as quality management and semantic consolidation.
- *Alignment* For bridging the heterogeneity between several ontology models, techniques for ontology alignment need to be provided. They comprise ontology mapping, for finding corresponding entities across ontologies, as well as ontology merging, for integrating ontologies based on such correspondences.
- *Reasoning* The process of reasoning checks an ontology for its consistency or derives new knowledge that logically follows from what was explicitly stated in the ontology. *Reasoning brokerage* allows for the dynamic combined use of different features and advantages of several reasoners, while *approximate reasoning* addresses scalability issues by trading correctness for speed.
- *Querying* Mostly related to reasoning, the process of querying retrieves relevant entities from ontologies. Typically, derived knowledge is taken into account for the generation of answers.
- *Entity Disambiguation* Apart from direct manipulation of ontologies, the process of ontology engineering needs to be supported in various ways. One such way of support is to use a tool that disambiguates the meaning of entities that occur in ontologies (or related text documents) based on their context.

In the following, the main components[1] designed and implemented within the THESEUS research program are sketched:

---

[1]The OM components are named after gods or heroes of Greek mythology in accordance with naming in the THESEUS research program, and with some meaning related to the component's functionality.

- *MNEMOSYNE* augments the OWL API by a persistence layer for the (native) storage of OWL ontologies in a relational database. This aims towards the scalable handling of large ontology models, for which the current in-memory implementation in the OWL API is insufficient.
- *ARETE* performs an automated recognition of references between natural language text elements and entities in an ontology.
- *ACHILLES* provides an adapter from the OWL API to RDF stores. It enables the programmatic handling of RDFS ontologies from within the OWL API, and thus bridges between the ontology languages OWL and RDF(S).
- *HARMONIA* addresses the problem of semi-automated ontology mapping for identifying correspondences between entities in different ontologies that cover similar or overlapping domains of interest.
- *CHRONOS* aims at providing a change management system for OWL ontologies that covers aspects of versioning and quality management. In its current state it offers tools for taxonomy clean-up and redundancy consolidation.
- *HERAKLES* provides a reasoning broker system that allows the use of various OWL reasoners in a combined way. It supports scenarios where several reasoners run in parallel on remote machines and where appropriate reasoners are selected at runtime.
- *DELPHI* encompasses systems for approximate reasoning with OWL ontologies that trade a certain degree of correctness of reasoning results for a speed-up in runtime performance. This component has been integrated into the reasoning broker HERAKLES.
- *ATLAS* is a reasoning framework for OWL 2 Full ontologies and RDFS. It is based on the working principle of theorem provers and thus offers a reasoning methodology different from known OWL-DL reasoners.
- *PYTHIA* provides the possibility of accessing OWL ontologies through structured queries stated in the query language SPARQL.
- *KOIOS* is a semantic search engine that enables keyword search on graph-shaped (RDF) Data. For a given keyword query, KOIOS computes k relevant SPARQL queries, answers to which are retrieved from an inverted keyword index and a special summary graph of the origin data. Additionally, KOIOS combines search within ontologies, databases and texts. Search results are returned according to relevance.
- *HEPHAISTOS* supports a rule-based transformation of RDF(S) data in the context of generating explanations for reasoning results.

Figure 2 details the subjects of ontology management, arranged on the two layers of maintenance activities and processing activities for ontologies, respectively.

In the field of ontology design three components have been developed: two for persistent storage of large ontologies, another for disambiguation of entities in texts using associated ontologies. Up to now, it is rather tiresome and problematic to load large scale ontologies into an ontology editor and manipulate them. MNEMOSYNE solves this issue by using an object relational representation based on the open-source framework Hibernate and translating accesses to the ontology directly into
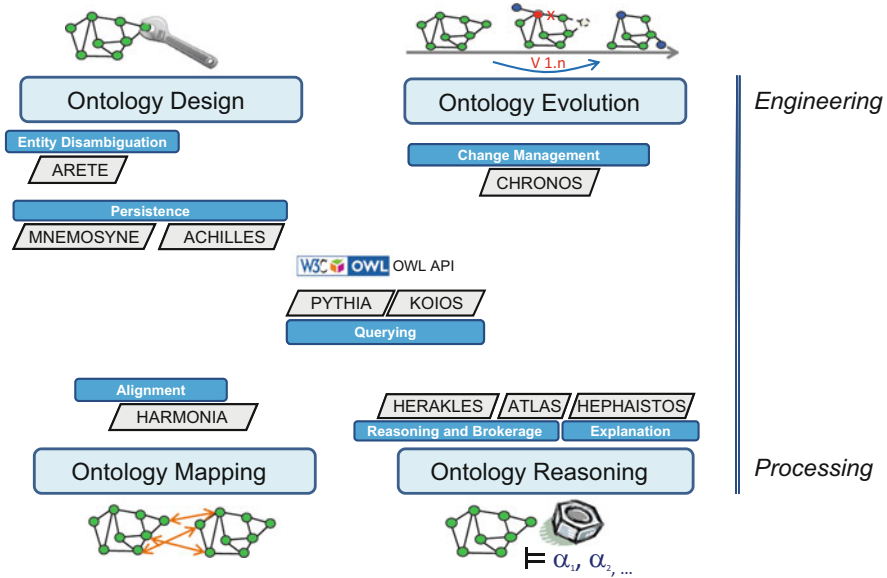
Fig. 2 Subjects of ontology management and their relation to WP3 components and Semantic Web technology

database queries via the OWL-API. The result of the query is transformed into axiom objects and entity objects. Therefore, only presently used objects of the ontology have to be kept in the memory when manipulating the ontologies. Graph-based queries in SPARQL can be made using the component PYTHIA, which directly interacts with MNEMOSYNE. The third component, called ARETE, is used for disambiguation of occurring ambiguities within texts, provided found text snippets refer to entities of an ontology at the same time, i.e. an ontology for geolocations can be used to distinguish between names referring to totally different locations, towns or cities. Disambiguation is achieved by using the spreading activation technique, which operates on partial graphs of the associated ontology with the help of statistical methods (Kleb and Abecker 2010).

Ontology mapping plays an important role in order to align different knowledge bases describing the same application domain but representing heterogeneous foci. Within this task a new component was designed and implemented using two different technologies: on the one hand a particle swarm algorithm is used to obtain alignments between two different ontologies (Bock 2010; Bock and Hettenhausen 2012); on the other hand correspondences between knowledge bases can be found with the help of evolutionary techniques. Both methodologies are incorporated within the component HARMONIA. Ontologies usually are built by different designers or ontology experts and thus have a subset of common representations or overlaps. These correspondences have to be automatically identified using HARMONIA. The results of this alignment process can then be used to form

new, bigger ontologies by adding the alignments to the two analyzed ontologies and obtaining a new, composite knowledge base. In this way the search space for queries can be enlarged to an enormous extent. This can be also illustrated by an example within the SUI (Search in Environmental Information Systems) research program using the components developed within THESEUS. Here, three different ontologies (GEMIT, catalogue of object types and ontology of life situations) are combined in order to search a much larger set of data when looking for building space within a given region. Without mapping these ontologies a person wishing to build a house in a specific region would only receive a subset or no results when entering the request in the environmental portal of the SUI server. For 3 years HARMONIA has also participated in the annual OAEI (ontology alignment) contest in the directory track and has there proven its capability to hold its own with the best mapping software programs. Additionally, the designed mapping component can exploit parallel computing infrastructures due to the techniques applied. Therefore, large ontologies can be aligned by exploiting parallelization and thus allowing good scalability. The algorithms can also be used in a cloud and have already been tested there (Bock et al. 2010).

There are many scenarios within which ontologies are developed on a collaborative basis, but if many different persons change, add or delete entities and data within a knowledge base, inconsistencies, mistakes and important differences can be the result. This issue is addressed by the third task of ontology management: ontology evolution. Here, a newly developed component called CHRONOS keeps track of different versions of an ontology, computes differences between different versions, performs repairs, executes unit tests, and corrects staging (Grimm and Wissmann 2011). CHRONOS is responsible for the grooming of ontologies and thus for quality assurance and quality control. With the help of a specially designed graphical interface users can design individual grooming processes for their ontologies. Additionally, the pursuit of development within a specific branch of an ontology is supported as well as the division of an ontology into different modules and the integration of new knowledge domains.

The fourth task within the ontology management work package focuses on ontology reasoning and semantic search. A specially designed reasoning broker called HERAKLES enables a processing of an ontology by several reasoning tools in parallel as well as the selection of the most suitable reasoning tool, if it is connected to the reasoning broker. Here, HERAKLES performs the communication between broker kernel and external reasoners via OWLLink. In this way, external reasoners can be addressed without explicit integration into the reasoning broker (Bock et al. 2009). Depending on the structure of the ontology some reasoners are more applicable than others. If there is a complex structure of several concepts and relations (called T-Box) another reasoner is chosen than if there were large scale data and a rather simple structure of the ontology (this is called A-Box). By analyzing the structure HERAKLES selects the appropriate reasoning tool and reasoning strategy. Two new reasoning techniques were also added to the broker system: approximate reasoning and anytime reasoning. In cases of very large knowledge bases it can be useful to start the processing of reasoning results before the actual reasoning process

has ended. In this case all results are continuously published while the reasoning tool is still working. Approximate reasoning produces either results which are all correct but do not cover the complete result set or all correct results plus some additional incorrect results. As this reasoning tool runs faster than others a slower reasoner running in parallel can be used to produce missing results or verify the output.

Finally, semantic search is the last semantic tool category designed and implemented within the ontology management work package. The component KOIOS performs keyword-based search on so-called linked data sets; these are data related to each other via a net of RDF triples. The keyword query is expanded by additional relations or details which the user can select before it is translated into a SPARQL query. Additionally, KOIOS searches for further existing links in the RDF graph, thus providing data not explicitly specified in the expanded query but relevant due to existing relations between data. The relations are computed based on feature vectors and weighted by probabilities (Bicer et al. 2011). For example, a movie can be classified as a German movie, although it is not labeled as such, solely by comparing producers, actors, directors and original language with those of other German movies.

All the above described components were developed within the THESEUS research program; some of them have been implemented in to several integration platforms, some of them are used in use cases; others have even been integrated by industry partners. Designing and implementing these components is an essential building block in getting enterprises to use semantic tools for knowledge management as future projects will surely show.

4   WP4: Situation Aware Dialog Processing

The Internet of Services poses new challenges, including the development of new interaction paradigms for the man-machine interface. As part of the THESEUS CORE TECHNOLOGY CLUSTER, we have developed a dialog shell, investigating new interaction styles and providing software tools for the rapid and efficient development of multimodal and situation-aware dialog systems.

Given semantically annotated content and access to semantic services, it becomes possible to extend existing interaction approaches that are menu-like and system-driven to a more natural interaction that is aware of the current situation and context. The dialog shell developed in WP4 enables flexible dialogs that integrate the user into complex business processes; see, e.g., the demonstrator for the Uses Case TEXO as described in Porta et al. (2014). Another interaction style, explorative browsing, enabled by semantically annotated content and services in conjunction with the dialog shell is demonstrated by the CALISTO system, described in Löckelt et al. (2014).

We followed an integrated approach, providing a comprehensive toolkit, the Dialog Shell, for the THESEUS research program and beyond (Sonntag et al. 2010). It addresses two basic research questions: how can semantic services in the Internet
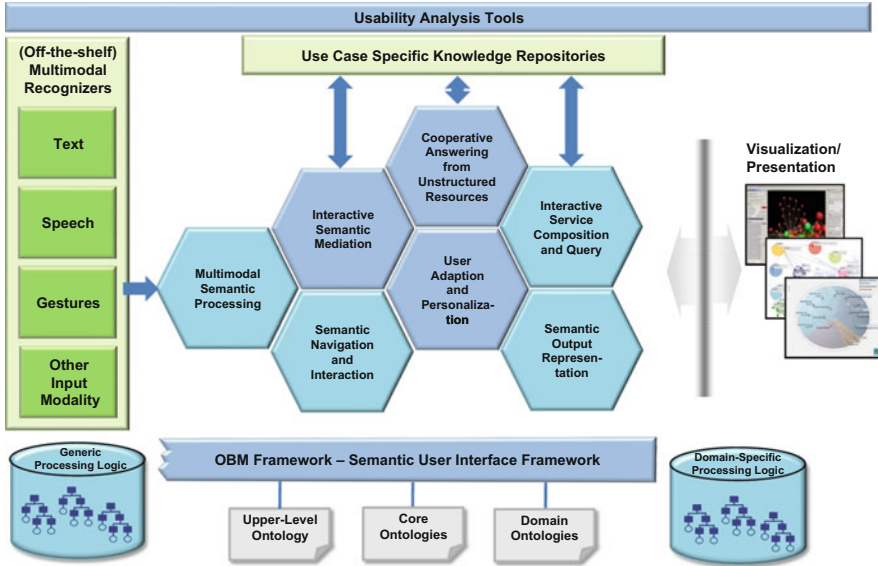
Fig. 3 Architecture of WP4

be found, addressed, and combined and what new forms of interaction are possible with them.

Ontology-Based Dialog Platform The core result of this work is the ontology-based dialog platform (ODP), which is used in a number of use cases of THESEUS and has been successfully marketed by the spin-off company SemVox since 2008.

The starting point for the development of ODP was the standard architecture for multimodal dialog systems as described in Wahlster (2006) and Becker (2010). It comprises a number of components that are organized along a core processing pipeline. To use this architecture for the Internet of Services, it has been extended and adapted as shown in Fig. 3. The main steps are sketched here and elaborated in Porta et al. (2014).

User input can come from various modalities in paralle, e.g., spoken language and pointing gestures as in "Show me more videos about this!". After separate mono-modal interpretations of the input, the fusion module computes the combined intention of the user, taking context and a user model into account. Once the intention has been recognized, dialog and interaction management computes the system reaction. Typically, this is a call to a service, e.g. to provide requested information. Which service is applicable and if the request can only be fulfilled by the combination of multiple services is decided in the modules for semantic mediation and service composition, which call the corresponding services. The results are processed and presented to the user by the presentation planning module, which also allocates the information to the available media channels. It could be

the spoken output "17 videos were found" combined with a list of the videos on the screen.

The ontology-based dialog platform (ODP) provides modules for all of these processing steps. These modules are connected through a common, semantic data representation formalism, the so-called extended typed feature structures (eTFSs) (Pfleger 2007). This eliminates the frequent transformation steps between modules that are typical for earlier dialog systems. Not only does this speed up processing, it also supports error-free internal communication. To work with this common semantic representation formalism, the modules are provided with a common API that implements some powerful computations on eTFSs (Alexandersson et al. 2006). Additionally, all modules have access to PATE, a rule-based abstract machine based on cognitive models (Pfleger 2007) that allows writing processing rules at a high level of abstraction and yet features efficient processing. Through common tools in all modules, the dialog shell provides – for the first time – a developer with an integrated development environment (IDE) for the various, diverse tasks in a dialog system.

This specialized IDE is an important part of the ODP platform and it is based on freely available standards and systems. All development tools are integrated into the Eclipse platform and support the development of code and knowledge sources, e.g., the rules for speech interpretation or the PATE system. Wizards and specialized editors support the efficient setup of new ODP projects and maintenance of knowledge bases and processing rules for the modules. The workbench also provides specialized debugging and testing tools.

For the connection between front end user interface and back end semantic services, WP4 has developed the so-called "Joint Service Engine" (JSE) that provides service grounding and supports the semantic annotation process for services with a declarative data type mapping mechanism. It integrates a range of different representation formalisms, e.g., RDF, XML(Atom), or JSON.

Multimodal Interaction As an example for multimodal interaction beyond the boundaries of devices, we have developed an information kiosk for the THESEUS Innovation Center in Berlin that allows search for and interaction with media content (images, videos and texts) on the exhibits of the center and on the early history of computing and computer pioneer Konrad Zuse. For example, the starting point of a search can be an image that is stored on a mobile device and sent to the information terminal with a "frisbee gesture", i.e., a simulated throwing of the images with the mobile device. On the console of the information kiosk, the multitouch screen can then be used to activate various interaction windows (spotlets); see Fig. 4. The central spotlet is used for semantic search, triggered by simple drag and drop of semantic concepts or simply media objects. Search results are shown according to media type, and the related semantic tags can be inspected and used for continuing the search and browsing activity. Besides watching the media objects, e.g. with the video player, other Internet-based services are connected. All objects that have geospatial metadata can be shown on a map and the Internet message service Twitter is connected by a spotlet. The functionalities can also be addressed through spoken

Fig. 4 Touchscreen of the information kiosk CALISTO on "The progress from the mechanical calculator to the Internet of Services"

interaction, e.g. by saying "Please show the THESEUS Innovation Center on the map."

The components in WP4 provide functionality for spoken and typed input, for gestures such as pointing, selection, and drag and drop and the symbolic throwing gesture (the so-called "frisbee gesture" where a media object is selected on a mobile device and "thrown"). Modality-specific interpretation components map the input into a common semantic representation and a fusion component derives an integrated interpretation of the user input. System output includes written and spoken text, various graphical presentation forms, including outputs from WP5 and a desktop-like environment, as well as haptic feedback, e.g. the vibration in a mobile device. Rendering components exist for a multitude of operating systems and devices. Except for the latter, all components are device independent and the dialog shell has been employed on desktop PCs, Android, and iPhone devices as well as tablets and large multi-touch kiosks.

This new mode of interaction, being a variant of browsing in semantic datasets, in combination with the Internet of Services allows for a faster navigation through the exhibits in the THESEUS Innovation Center and guarantees a high relevance of the search results through the semantic annotation.

The integrated dialog shell is an innovative basis for new user interfaces for the numerous applications of the Internet of Services. It is already successfully marketed by SemVox GmbH, a spin-off company of DFKI GmbH, demonstrating the market readiness of the technology. The results of this work have been published in books, journals and at conferences (a total of 36 publications) and led to two Ph.D. and four Master's theses, eight lectures and seminars, four contributions to standards and a patent. Based on the dialog shell, three industrial projects, six EU projects

and three national projects have been started and contribute to DFKI's participation
in the EIT ICT Lab. Current work further extends the approach of the dialog
shell: a common semantic representation, strong modularization, efficient semantic
processing throughout the system, adaptation to the situation and personalization
and a strongly coupled development of mediation and composition of semantic
services.

5   WP5: Innovative User Interfaces and Visualizations

Work Package 5 investigated the graphical representations of semantic and massive
data in two main areas of research: Semantics Visualization and Visual Analytics.
Semantics Visualization focuses on human-centered graphical representations of
complex semantic structures or formal notations of ontologies. Research on this
topic incorporates the investigation of data manipulation and enrichment of graph-
ical algorithms for representing the structure of the semantic conceptualization.
Furthermore the question of how to present the graphical information, e.g. by visual
attributes or data filtering and recommendation, were topics of research.

The second main topic of WP5 was the research and interaction with huge data
amounts. Visual Analytics in WP5 investigated different methods for analyzing,
structuring and visualizing to support decision-making and analyzing tasks.

CTC work packages aim at providing core technologies for an instantiation in the
THESEUS use cases. In WP5 the human factor played a key role during the process
of research and development. One of the main challenges was to develop systems
that provide generic characteristics to be adaptable to various usage scenarios and
to consider the variety of human perception, preferences and knowledge. Thus
the THESEUS use cases did not only investigate different domains of knowledge,
e.g. medical computing, Internet of Services or social and semantic web, their
scenarios involved heterogeneous users. The users within just one use case differed
significantly, e.g. starting from service engineers to service providers and ending in
service consumers. Facing the heterogeneity of users in visualization, systems by
adaptable and adaptive visualizations was the main challenge of this work package.

Semantics Visualization   WP5 has investigated techniques and concepts to visual-
ize semantic data for heterogeneous users, usage scenarios and semantic character-
istics. The challenge of the research was the conceptualization and development of
a technology that is user-oriented on the one hand and features the characteristics of
a core technology on the other hand. Therefore the adaptable and adaptive SemaVis
technology was developed, which covers both the features of a user, use case
and context adaptable core technology and a user-centered and adaptive Semantics
Visualization.

SemaVis integrates a method for orchestrating visualizations in a multiple
visualization user interface as a visualization cockpit (Nazemi et al. 2010a). A set
of about 15 different visualization techniques can be positioned in a user interface

for an enhanced juxtaposed aspect-oriented visualization. Each of the visualizations focuses on certain characteristics of semantic data. The SemaVis cockpit metaphor is more than the combination of multiple visualizations, as known from Brushing & Linking. It provides the separation of the semantic information into several view-modes for supporting heterogeneous tasks, e.g. analytical comparison, knowledge exploration or information acquisition.

SemaVis integrates the visualizations and the general user interface using a layer-based approach for a fine-granular separation of visual information. The entire information visualization process is subdivided into different layers of abstraction. Visualizations can be described according to what is displayed, where with it is displayed, and how it is displayed. Accordingly, each visualization in SemaVis implements the layers semantics, layout and presentation (Nazemi et al. 2011).

Semantics defines which data is visualized. It contains information about the data, its structure and its implicit and explicit properties. Layout defines the graphical layout algorithm to be used. And presentation defines more precisely how the data will be presented. It is the visual layer of the visualizations and parametrizes the visual look. The input of the presentation is the geometry as calculated by layout.

The layer-based adaptability of SemaVis provides an adequate approach for generating heterogeneous user interfaces by manipulating the entire spectrum of the visualization characteristics. This ability is used to further provide an automatic adaptation to several impact factors, e.g. user interaction or data structure. The goal of the automatic adaptation of the visualization is to support the user in his navigation process, during the exploration of information. Therefore a new approach was conceptualized, where the impact factors are captured and enriched with a semantic structure provided by the data. A user interaction, for instance, is determined as a three-dimensional conceptualization of the interaction taxonomy. Thereby the semantic taxonomy of data is used and enriched with a semantic hierarchy (Nazemi et al. 2010b). Further the characteristics of the data, especially the semantic structure and the explicit and implicit properties, are considered in a similar way. Therefore a data semantics is generated that provides information about the current visible data.

Visual Analytics An ideal search and retrieval engine structures and organizes information in the way needed by a user in his current situation and task. It shows the right information at the right time – in the right way. While representing, refining and using such a structure was the main goal of the aforementioned achievements, the goal of Visual Analytics in THESEUS was the creation of this structure. For unstructured or weakly structured data – most raw texts fall in this category – this is a difficult feat. The ideal structure does not only depend on properties of text, like vocabulary, topic and context; in addition it should reflect how the user thinks about the data.

Purely automated methods for data analysis and machine learning reflect how the developer thinks about the data, but there is never a guarantee for a match between the user's and the developer's perspective. Visual Analysis combines techniques

for visual interactive and automated analysis. The goal of this combination is to compensate for the weakness of one technology with the strength of the other. Analysis becomes flexible by introducing the user into the process via visualization: visual representations and interaction.

Visual Analytics research in THESEUS explored how one of the most valuable assets could be introduced to the analysis – the knowledge and abilities of the human user. For example, visualization enables the user to detect patterns and structures in noisy and ambiguous data. It enables the user to evaluate and refine analytical results and to dynamically control the process depending on the users' knowledge. These leverage points have been investigated and structured within the Visual Analytics Framework. In addition, WP5 investigated strategies for how the leverage points were made accessible even to non-experts to analysis, allowing the users to interact with the data on their own terms.

Visualization Use Cases WP5 technologies were implemented in several usage scenarios and applications within and beyond the THESEUS research program. For instance SemaVis was integrated into different applications of the Internet of Services. The architectural design of the technology allows the instantiation of the visualizations as services. WP5 forced the idea to use Visualizations as a Service (VaaS) for providing rich and interactive graphical representations of highly complex and semantics data. The modular architecture allows a differentiated implementation for various data and for different users. Existing applications of SemaVis provide exploratory searching, learning and analysis tasks, e.g. in joint application with the THESEUS use cases ORDO and PROCESSUS (Stab et al. 2012). Further the exploration of services as a service browser provides the opportunity to examine virtual services for different roles, e.g. service consumer or service engineer. An open-access version of SemaVis opens the technology for everyone and provides a visual interactive interface to the most common semantic data repositories and search environments. In future applications SemaVis will be enhanced to do more than visualize semantically annotated information and especially to provide support for visual analysis and decision making tasks. Therefore existing open social media repositories will be used to visualize sentiments and opinions about topics of interest.

The visual interfaces of both researched areas, Semantics Visualization and Visual Analytics, will be more and more applied to the policy modeling life cycle. The visualization of the lacks in existing policies and the visualization of the impacts of new policies will be a main area of interest in the early future. Further related areas, e.g. the visualization of law and law consequences, will be very important to the related areas.

Semantics Visualization and Visual Analytics provide interaction abilities with graphical representations of information. With the changes in mobile computing and alternative interaction techniques, the graphical interaction will have a sustainable role in our way of interacting with information.

6   WP6: Statistical Machine Learning

Currently the Internet is dominated by unstructured content such as texts, images and videos, all of which are representations of information that are quite suitable to a human user but not to a machine. In the evolving Internet of Services, machines need to make decisions, which requires that machines have some form of understanding of unstructured content. A classical example, relevant even before the golden age of the Internet, but still of high importance nowadays, is postal automation, where a machine needs to understand the handwritten postal code and address on the envelope of a letter in order to properly sort the letter. Content extraction evolved into a major task on the Internet, where a text, image or video needs to be annotated with the right labels such that it can be found by a person with the corresponding information need. Statistical machine learning has become instrumental to solve these tasks and in WP6 a number of novel approaches have been developed.

Textual information is one of them. A first step to understanding a text is to extract the major named entities that are mentioned, such as persons, locations, companies, diseases and genes. Based on sophisticated statistical methods a number of approaches were developed for this task that are even able to deal with multilingual documents. More useful but also more challenging is the extraction of textual statements in the form of relation extraction and semantic role labeling. Thus, whereas named entity recognition identifies Kohl as a person's name and China as a country (and not as a ceramic), relation extraction would be able to extract from the text the statement (Kohl, visits, China) as the content of a news story. Finally, statistical topic modeling describes a document by extracting its main themes, and in the work package existing approaches were refined to combine the strengths of topic modeling and text annotation; see also Bundschus et al. (2008, 2009).

The developed methods were applied to news texts in the CONTENTUS Use Case, for the extractions of the relationships between genes and diseases out of medical abstracts in the MEDICO use case, and for the detection of the relationship between companies (GE, competitor of Siemens) out of news texts for the ORDO use case. The approaches have also been extended to detect opinions or sentiments in texts. Thus a company's product problems can quickly be detected by analyzing blogs and news texts. Finally, one can get insight into a document by extracting the main topics covered. We consider two applications developed with the THESEUS technologies. The QUOTE application regularly analyzes fresh news texts by relation extraction technology and detects persons and quotes by those persons, which is much more relevant than reported speech. A second application "Eat & Drink" scans the comments in a large repository of restaurant reviews and extracts important phrases highlighting some aspect of a restaurant. Hence the often very large number of reviews is condensed to a few highlighting citations. Together with statistical information on the reviews this allows the users to see the strengths and

weaknesses of a restaurant at a glance and alleviates the decision for the user. Both applications are provided as apps for Android and the iPhone.

Based on the topic models and the other developed technologies on text analysis, a demonstrator for the handling of claims was developed and presented at CeBIT. The demonstrator was also using results of WP6 developed for the analysis of document structures and for the recognition of handwritten text. For the recognition of handwritten text, a new learning approach based on recurrent hidden Markov models has been developed and has already been successfully integrated in various commercial applications. For the analysis of document structure, a self-adaptive solution has been developed. Consider that existing approaches for automatic logical structure analysis in documents have in common that they were developed for a specific task and document type. Adapting such a method to a different task requires modifying the existing set of rules or grammars, which is a laborious manual task. The document structure analysis developed within WP6 uses machine learning techniques instead of manually created rules or grammars. The module learns a structure model that takes layout, formatting and content features into account. The learned structure model is then applied in order to analyze the structure of a given document. For more details, see Schambach (2009) and Stoffel et al. (2010).

Machine learning is also very effective for the extraction of information from images. Image annotation is the basis for content-oriented retrieval of images, and highly competitive approaches for generic images were developed in WP6. As a special case we considered medical tomographic images, the focus of the use case MEDICO. We developed solutions for semantic localization, the task of automatically finding body regions which contain certain anatomical concepts like organs, bone structures or body landmarks. This information is very useful when working with CT Scans and for connecting them to other medical information. Thus, semantic localization is a powerful tool to connect the world of medical 3D images to the world of ontologies and semantics. Another important functionality is to align two CT scans for differential diagnosis and to optimize retrieval times from the picture archiving and communication system to the workstation of a physician. Compared to other solutions to this problem our new approach suffices with very limited information, i.e. a single 2D slice from the CT scan to precisely predict positions in the human body. Whereas some of our solutions are specialized for detecting a certain type of anatomical concept, for example the location of the vertebrae, we also developed a very generic approach mapping the positions of single slices to a general model of the human body. The model consists of a standardized coordinate system of the human body and statistical information about the location of multiple anatomical concepts within this coordinate system. Employing a further method being developed within the work package, it is now possible to align a CT volume to the coordinate system and to predict the most likely location of the anatomical concepts within the CT volume. A developed use case permits a physician to precisely localize a desired tomographic slide in the human body without loading the whole set of all tomographic slices, thus reducing loading times and network load substantially. This solution became part of the MEDICO use

case and found great interest in the medical community in general. For more details, see Grafa et al. (2009) and Graf et al. (2011).

Another special case constitutes the annotation and ranking of images for generic visual concepts. Such annotations may be employed for the search of images containing certain visual concepts based on the image content rather than keywords linked to an image. Solutions developed within the work package have achieved top ranks in international benchmark competitions with undisclosed ground truth such as the Pascal VOC Classification Challenge and yielded first-ranked submissions in the ImageCLEF2011 Photo Annotation Challenge. The methodological challenge consisted in the ability to deal with a large set of differing visual concepts, ranging from localized objects to overall emotional impressions; the large variability of image appearance present within general visual concepts beyond simple objects like Partylife, BeachHoliday, Mountains, Indoor, Euphoric, and AestheticImpression; and the disagreement of human annotators in labeling images for such concepts. Statistical methods are robust against annotator disagreement, labeling errors, varying image qualities and scales of visible cues, differing lighting conditions, occlusion and clutter in images. They are able to extract relevant structures based on implicit and statistical definitions, labeling example images as belonging to the same concept, even when there is no way to define deterministic rules for what a concept should be. This makes statistical learning methods a valuable complement to explicit and deterministic knowledge modeling. The results of this work are described in detail in Binder et al. (2014).

Finally, WP6 considered video information. The main focus here was the extraction of textual information in videos. The developed solutions are applied to the detection of logos in videos, e.g., for detecting placed advertisement in sport reports and for the detection of licence plates and vehicle types in traffic surveillance systems. The logo recognition technology developed within this work package works on previously learned logos and allows fast recognition and high recognition performance, enabling various real-time recognition applications. For example, in postal automation the online detection of postal stickers and value symbols like stamps, service symbols, return address logos or similar logos is already successfully in operation.

In addition to unstructured data, machine learning can be applied to structured data as well. Consider the Linked Open Data (LOD) initiative, where data from various domains is made available in the format of the Semantic Web. In addition, information sources cross-reference one another such that entities become unambiguous: e.g., if two data sources talk about Paris, it is clear if they both mean the capital of France and not a person popular with the media. Structured data show regularities that can be exploited via machine learning. In WP6 we have developed various approaches for machine learning in semantic domains in the form of the SUNS framework (Huang et al. 2010; Tresp et al. 2009). One application in MEDICO permits the prediction of diseases that might be affected by a given gene mutation. Other applications are in the recommendation of web services and in the recommendation of procedures in medical domains, as decision support component for physicians. Details of our approach can be found in Tresp et al. (2014).

7   WP8: Evaluation

There are two key paradigms researchers can follow: research just to learn, and research to solve practical problems. The first one in general is called basic research, the latter applied research. The result of applied research is usually something like a prototype which might be rather close to a final product. Most of the work in THESEUS was focused on applied research, but the work flow from technology to prototype was split into two separate functions: The basic technologies necessary for at least two application areas have been developed in the CORE TECHNOLOGY CLUSTER; the integration and prototype development was concentrated in the use cases. That way synergy effects between application areas could be exploited.

Doing applied research is always an important topic for evaluation: Only if there are clear definitions of "good" and "better" is it possible to improve a technology. The evaluation tasks are different in the CTC and in the use cases: In a use case the complete system can be accessed concerning performance and usability. The weighting of different dimensions of quality often depends on subjective aspects and is application specific. In the CTC usually only technology components are available for testing. Each dimension is evaluated separately. An example can illustrate this: In the final application a question typed in by the user should be answered by the system. For the end user it is essential that he receives a useful answer in reasonable time. If the answer needs a lot of time or if there are too many results which do not make sense to the user he will not accept such a system. From a CTC point of view this system will consist of several parts: Algorithms for the analysis of the query, algorithms for the pre-processing of the database used in the query, algorithms to search in that database and tools to present the result of the search. For each of these components there are different performance parameters like computational time, storage requirements and correctness. Not all of the components are necessary for each query, but might influence the total performance. In the CTC it was essential to have the detailed performance parameters to monitor and control the development process. Note that the user interface, which is a key component in appearance for the end user, most of the time is application specific and is therefore part of the use case.

Due to the fact that each use case had different requirements on the overall performance, its evaluation was done in each use case separately. The evaluation of technology components developed in the CTC was concentrated in WP8. To fulfill the final requirements of the use case all evaluations had to consider the intended use. Special care had to be taken to use test data similar to the data in the use case. Due to the fact that the CTC developed a variety of different algorithms also a large number of test methods had to be implemented and performed. To exploit synergy effects expert groups performing the evaluations were organized in tasks which did not match to the structure of the CTC 1:1. To monitor progress and give hints about weak points in the algorithms all evaluations have been repeated annually during the existence of each development task. The following list gives some examples of

evaluations performed. Note, that due to the limitations of this article results cannot be detailed.

Databases In many evaluations audiovisual content was necessary in a raw impaired form. The database service task was responsible for the collection, creation and processing of data to be used in the evaluations. An important aspect for testing the performance of systems is that the databases used for the evaluation should not be used for the optimization of the algorithms. Therefore several new databases have been generated and kept secret from the partners. To enable the international comparison of results the task also organized challenges in international benchmarks under the umbrella theme of ImageCLEF (see also Liebetrau et al. 2014). Processing of data included the adaptation to different resolution for images and movies but also the controlled impairment like adding noise, coding, smearing and rotation of pictures.

Text Analysis Older data collections in libraries and medical archives are only available on file cards. They often are only available on paper and are even hand-written. Usually they have fixed layouts which simplify the automatic processing. Different font types can be used to recognize the function of recognized test. Within the CTC, algorithms for text recognition and text type recognition have been developed. To evaluate these algorithms, databases from medical records and libraries have been used. To simulate the effects of old paper and bad scanning, different impairments like adding of noise and rotating of picture have been used.

Media Data Analysis For the management of media databases it is essential to recognize duplicated pictures and videos. An important aspect is that copies which would be recognized as "identical scenes" by humans should be detected, while copies humans would annotate as "different takes by same actor" should be classified as different. The algorithms evaluated therefore created perceptual hashes for each item. For the evaluation modified copies of images and videos were used. Such modifications included rotation, changes of aspect ratio and blurring.

Picture Analysis In THESEUS, understanding the content of pictures is necessary in many use cases. Due to different demands a large number of CTC tasks performed work on different algorithms. The algorithms for face detection had to find one or several faces in pictures with different light and background. Evaluation criteria were that all faces present on the picture were found and marked and that no other structures were wrongly marked to be a face. The task of picture segmentation is to mark objects in pictures. Evaluation criteria were both the detection ratio and the precision for finding the form of the marked objects. In the context of analysis of films, algorithms for the detection of scenes are important. Detection might be slightly before or after the "switch" between scenes. The evaluation was counting the correct and false positive detections within a short time window around the correct scene borders. To find pictures with certain features, like "summer", "vehicles", "mountain" or "family&friends" in large databases of images, automatic algorithms for image annotation have been developed. Manually annotated databases have been

used for the evaluation. To enable international comparison, an evaluation task in the ImageCLEF workshop was organized. For further information about the test data see also Liebetrau et al. (2014). Algorithms for video genre classification (film, cartoon, news, commercial, music) were evaluated, based on video clips either containing only one genre ("pure") or an artificial combination of scenes of different genre ("mixed"). Each genre was represented by about 5 h of content. For "pure" data the analysis of whole clips had higher precision than the analysis of single video frames. The performance for the "mixed" content was worse than for the "pure" content. However in a real world usage scenario video genre classification would be combined with scene detection: As a result the algorithm would always work on "pure" data. Special algorithms in picture analysis are necessary for the localization of Computer Tomography slices. Such algorithms are useful in several medical applications (see Kuhlmann et al. 2014). In the evaluation 34 datasets from 24 patients and different regions of the body (pars cervicalis, thorax) were used. Depending on the region, different algorithms provided the best results.

Audio Quality Originally it was planned that the selection of algorithms for speech recognition and text2speech will be necessary for innovative dialog shells. In an early phase of THESEUS it was decided that such algorithms were application specific and are no longer on the list of research to be performed in the CTC.

Video Quality Algorithms for the efficient representation of video content at different resolutions are important aspects for many applications in THESEUS. The final instance for the assessment of video quality is the visual test with test subjects. Such tests are expensive and time consuming. Therefore WP2 conducted work towards the automatic measurement of the perceived video quality. Automatic measurement normally is based on the comparison of an unimpaired reference with the output of a technical system. Such algorithms can work based on either "full reference" or "reduced reference". Both configurations have the disadvantage that such measurement algorithms are not adequate in applications where the reference is not accessible or does even not exist at all (example: copy of an old film). The algorithm developed in WP2 therefore estimates the quality parameters "blur" and "blocky" by analyzing the output data only.

WP8 evaluated coding schemes and measurement schemes developed in WP2: One of the databases used to evaluate the performance of the video codec was used also to evaluate the performance of the video measurement algorithm. The two non-reference quality parameters "blur" and "blocky" in general showed a similar rank order correlation as the state of the state-of the-art full-reference algorithms PSNR and SSIM.

Iterative System Design For fulfilling the wide range of tasks, the CTC had to deal with ontologies, infrastructural measures, various visualization modes, and the annotation of various file types (images, documents), including machine learning aspects. WP8 thus aimed at identifying, preparing, performing, and analyzing appropriate evaluation means, methods, tools, and procedures for all the aforementioned tasks (components), frequently reporting the results back to the developers and into

their iterative development cycles. A total of more than 25 CTC tasks contributed to the list of components to evaluate.

The evaluation of the WP3 tasks was based on the *ontology management components* called DELPHI and HERAKLES (Ontology Reasoning), HARMONIA (Ontology Mapping) and MNEMOSYNE (Ontology Design and Ontology Evolution). The efficient manipulation and usage of large OWL ontologies is a major problem, because large ontologies often exceed the available main memory. This problem is caused by the fact that ontologies have to be parsed and loaded in the memory before they can be used. Performance and scalability are thus two major criteria for describing the quality of a persistence system. The evaluation measured the respective response times for several query tasks. Both the performance and the scalability of the systems have been tested against real-world ontologies of varying sizes and complexity.

The WP4 focused on a flexible *dialog shell* for the implementation of applications with multimodal user interaction. It consisted of several different integrated functional modules that can be connected in adaptable and flexible combinations to each other. The developed test framework also fulfilled basic requirements that pointed toward the possibility of measuring the time performance of the functional modules, forming the dialog shell, as well as debugging and tracking the event flow. The evaluation aimed at successfully running both single component tests as well as an end-to-end test, particularly for components relevant for the use case MEDICO.

The objective of WP5 has been the development of a reliable and *useful user interface* that is appropriate to the user. The user interface and the underlying *visualization* technologies developed aimed at being generic in order to be used in several Use Cases. The evaluation based on various cycles and a set of tasks to be fulfilled by human testers made sure that the modular approach was able to satisfy both functional and technical requirements of the different use cases. Apart from technical tests, a large usability evaluation completed the analysis of the WP5 components.

The evaluation of WP6 contained components concerning algorithms that extracted a latent semantic representation from *textual corpora* as well as *self-learning algorithms* for large-scale textual archives. The tests addressed mainly the correctness of the results, and made use of available tools and corpora as well as self-created tools and manually annotated corpora. The strengths and weaknesses of the components have been spotted and, where appropriate, the quality has been compared to state-of-the-art approaches. Further aspects under consideration have always been the scalability of the solutions and the required amount of annotated data for fulfilling the requirements from the use cases.

Privacy and Security Privacy and security aspects have to be considered in system designs from the beginning. In the THESEUS context this is especially important for the use case where core technologies are integrated into systems. Within the CTC only a few technologies have privacy aspects, and these tasks have been analyzed based on technical and legal requirements. In a final phase of THESEUS a workshop concerning privacy aspects for the whole THESEUS research program

was organized in cooperation with the ULD Schleswig-Holstein. An important result of this workshop was that in subsequent projects privacy has been considered as a separate key action across the whole program.

Field Tests The CTC presented its huge variety of developed technology components in the context of specific demonstrators at project meetings, trade shows, and various domain-related fairs. For figuring out both the functional and the non-functional capabilities of those components, a series of tests was necessary to perform aiming at much more than just pure functionality. The wide range of tests performed in that context was especially focused on, but not limited to, usability aspects. From that viewpoint, the standardized evaluation procedures along with selected methods and characteristics addressed preferably the entire chain of components (e.g. from input to analysis, to search, to processing, to visualization) integrated into a demonstrator application. Therefore non-functional aspects like usability, always based on standardized evaluation methodologies, characteristics, and sub-characteristics, have been evaluated.

8   Results

Altogether 45 different basic technologies have been developed within the THESEUS CORE TECHNOLOGY CLUSTER. A number of outstanding scientific results have been achieved, which is documented by a number of top rankings in international challenges. Furthermore the number of publications is remarkable. It includes 35 book chapters, 26 journal papers, 235 contributions to conference proceedings, 82 additional publications, 26 patents and 29 contributions to standardization bodies. In addition 20 doctoral theses have been completed and CTC results are used for teaching, including 25 diploma, bachelor and master theses and 3 lectures.

Besides the research and development described above, the CTC dedicated some effort towards the development of joint demonstrators. The reason for that was to better understand the potential of the technologies developed in the "other" work packages and the necessity to prove the interoperability of different technologies.

9   Conclusion

The exploitation of the developed CTC technologies is successful. They are used by all use cases and by many SME projects within THESEUS. Several CTC technologies found their way into international standards. A number of license agreements with third party partners have been concluded and five spin-off companies have been founded. In addition, a number of new R&D projects on national, European and industrial basis have been started to bring the CTC results closer to products.

References

J. Alexandersson, T. Becker, N. Pfleger, Overlay: the basic operation for discourse processing, in *SmartKom: Foundations of Multimodal Dialogue Systems*, ed. by W. Wahlster. Cognitive Technologies (Springer, Berlin/Heidelberg/New York, 7/2006), pp. 255–267

T. Becker, (Multimodale) Dialogsysteme, in *Computerlinguistik und Sprachtechnologie – Eine Einführung*, ed. by K.U. Carstensen, C. Ebert, C. Ebert, S. Jekat, H. Langer, R. Klabunde (Spektrum, Heidelberg, 2010), pp. 624–632

V. Bicer, T. Tran, A. Gossen, Relational Kernel machines for learning from graph-structured RDF data, in *The Semantic Web: Research and Applications*, Heraklion, 29 May–2 June 2011, ed. by G. Antoniou, M. Grobelnik, E. Simperl, B. Parsia, D. Plexousakis, P.D. Leenheer, J. Pan. Volume 6643 of Lecture Notes in Computer Science (Springer, Berlin/Heidelberg/New York, 2011), pp. 47–62

A. Binder, W. Samek, K.R. Müller, M. Kawanabe, Machine learning for visual concept recognition and ranking for images, in *Towards the Internet of Services: The THESEUS Research Program*, ed. by W. Wahlster, H.J. Grallert, S. Wess, H. Friedrich, T. Widenka (Springer, Berlin/Heidelberg/New York, 2014)

J. Bock, Map PSO results for OAEI 2010, in *5th International Workshop on Ontology Matching (OM-2010), Collocated with the 9th International Semantic Web Conference (ISWC-2010)*, Shanghai, vol. 689, Nov 2010, pp. 80–186

J. Bock, J. Hettenhausen, Discrete particle swarm optimisation for ontology alignment. Inf. Sci. **192**(1), 152–173 (2012). http://dblp.uni-trier.de/db/journals/isci/isci192.html#BockH12

J. Bock, A. Lenk, C. Dänschel, Ontology alignment in the cloud, in *Proceedings of the 5th International Workshop on Ontology Matching (OM '10), Collocated with the 9th International Semantic Web Conference (ISWC '10)*, Shanghai, vol. 689, ed. by P. Shvaiko, J. Euzenat, F. Giunchiglia, H. Stuckenschmidt, M. Mao, I. Cruz (CEUR Workshop Proceedings, 2010), pp. 73–84 http://ceur-ws.org

J. Bock, T. Tserendorj, Y. Xu, J. Wissmann, S. Grimm, A reasoning broker framework for Protégé, in *11th International Protégé Conference*, Amsterdam, June 2009

M. Bundschus, M. Dejori, M. Stetter, V. Tresp, H.P. Kriegel, Extraction of semantic biomedical relations from text using conditional random fields. BMC Bioinform. **9**(207), (2008) http://dblp.uni-trier.de/db/journals/bmcbi/bmcbi9.html#BundschusDSTK08

M. Bundschus, S. Yu, V. Tresp, A. Rettinger, M. Dejori, H.P. Kriegel, Hierarchical Bayesian models for collaborative tagging systems, in *Proceedings of the IEEE International Conference on Data Mining (ICDM '09)*, Miami, ed. by W. Wang, H. Kargupta, S. Ranka, P.S. Yu, X. Wu (IEEE Computer Society, 2009), pp. 728–733. http://dblp.uni-trier.de/db/conf/icdm/icdm2009.html#BundschusYTRDK09

P. Ciaccia, M. Patella, P. Zezula, M-tree: an efficient access method for similarity search in metric spaces, in *Proceedings of 23rd International Conference on Very Large Data Bases (VLDB '97)*, Athens, 25–29 Aug 1997, ed. by M. Jarke, M.J. Carey, K.R. Dittrich, F.H. Lochovsky, P. Loucopoulos, M.A. Jeusfeld (Morgan Kaufmann, 1997), pp. 426–435

F. Graf, H.P. Kriegel, M. Schubert, S. Pölsterl, A. Cavallaro, 2D image registration in CT images using radial image descriptors, in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Toronto, ed. by G. Fichtinger, A.L. Martel, T.M. Peters. Volume 6892 of Lecture Notes in Computer Science (Springer, Berlin/Heidelberg/New York, 2011), pp. 607–614. http://dblp.uni-trier.de/db/conf/miccai/miccai2011-2.html#GrafKSPC11

F. Grafa, R. Greila, H.P. Kriegela, M. Schuberta, A. Cavallarob, Enhanced detection of the vertebrae in 2D CT-images, in *Proceedings of the SPIE Medical Imaging 2010: Image Processing*, San Diego, vol. 8314, Feb 2009

S. Grimm, J. Wissmann, Elimination of redundancy in ontologies, in *Proceedings of the 8th Extended Semantic Web Conference (ESWC '11)*, ed. by G. Antoniou, M. Grobelnik, E.P.B. Simperl, B. Parsia, D. Plexousakis, P.D. Leenheer, J. Pan. Volume 6643 of Lecture Notes

in Computer Science (Springer, Berlin/Heidelberg/New York, 2011), pp. 260–274. http://dl.
acm.org/citation.cfm?id=2008892.2008916

Y. Huang, V. Tresp, M. Bundschus, A. Rettinger, H.P. Kriegel, Multivariate structured prediction
for learning on semantic web, in *Proceedings of the 20th International Conference on Inductive
Logic Programming (ILP '10)*, ed. by P. Frasconi, F.A. Lisi. Volume 6489 of Lecture Notes in
Computer Science (Springer, Berlin/Heidelberg/New York, 2010), pp. 92–104

J. Kleb, A. Abecker, Entity reference resolution via spreading activation on RDF-graphs, in *The
Semantic Web: Research and Applications: 7th Extended Semantic Web Conference*, ed. by
L. Aroyo, G. Antoniou, E. Hyvönen, A. ten Teije, H. Stuckenschmidt, L. Cabral, T. Tudorache.
Volume 6088 of Lecture Notes in Computer Science (Springer, Berlin/Heidelberg/New York,
2010), pp. 152–166. http://dblp.uni-trier.de/db/conf/esws/eswc2010-1.html#KlebA10

F. Kuhlmann, J. Hannemann, M. Traub, C. Böhme, S. Zillner, A. Cavallaro, S. Seifert, B. Decker,
R. Traphöner, S. Kayser, U. Lindemann, S. Prasse, G. Marczinski, R. Grützner, A. Fasse,
D. Oberle, The THESEUS use cases, in *Towards the Internet of Services: The THESEUS
Research Program*, ed. by W. Wahlster, H.J. Grallert, S. Wess, H. Friedrich, T. Widenka
(Springer, Berlin/Heidelberg/New York, 2014)

J. Liebetrau, S. Nowak, S. Schneider, Evaluation of image annotation using Amazon mechan-
ical Turk in ImageCLEF, in *Towards the Internet of Services: The THESEUS Research
Program*, ed. by W. Wahlster, H.J. Grallert, S. Wess, H. Friedrich, T. Widenka (Springer,
Berlin/Heidelberg/New York, 2014)

M. Löckelt, M. Deru, C.H. Schulz, S. Bergweiler, T. Becker, N. Reithinger, A unified approach for
semantic-based multimodal interaction, in *Towards the Internet of Services: The THESEUS
Research Program*, ed. by W. Wahlster, H.J. Grallert, S. Wess, H. Friedrich, T. Widenka
(Springer, Berlin/Heidelberg/New York, 2014)

H. Müller, T. Tsikrika, Global pattern recognition: the ImageCLEF benchmark. IAPR Newsl.
32(1), 3–6 (2010)

National Institute of Standards and Technology (NIST). Secure Hash Signature Standard (SHS)
(FIPS PUB 180-2), 2002

K. Nazemi, M. Breyer, D. Burkhardt, D.W. Fellner, Visualization cockpit: orchestration of multiple
visualizations for knowledge-exploration. Int. J. Adv. Corp. Learn. **3**(4), 26–34 (2010a). http://
dblp.uni-trier.de/db/journals/i-jac/i-jac3.html#NazemiBBF10

K. Nazemi, C. Stab, D.W. Fellner, Interaction analysis for adaptive user interfaces, in
Advanced Intelligent Computing Theories and Applications, ed. by D.S. Huang, Z. Zhao,
V. Bevilacqua, J.C. Figueroa. Volume 6215 of Lecture Notes in Computer Science (Springer,
Berlin/Heidelberg/New York, 2010b), pp. 362–371

K. Nazemi, C. Stab, A. Kuijper, A reference model for adaptive visualization systems, in *Human-
Computer Interaction: Design and Development Approaches*, ed. by J.A. Jacko. Volume 6761
of Lecture Notes in Computer Science (Springer, Berlin/Heidelberg/New York, 2011), pp. 480–
489

N. Pfleger, *Context-Based Multimodal Interpretation: An Integrated Approach to Multimodal
Fusion and Discourse Processing*. PhD thesis, Saarland University, 2007

D. Porta, M. Deru, S. Bergweiler, G. Herzog, P. Poller, Building multimodal dialog user interfaces
in the context of the Internet of services, in *Towards the Internet of Services: The THESEUS
Research Program*, ed. by W. Wahlster, H.J. Grallert, S. Wess, H. Friedrich, T. Widenka
(Springer, Berlin/Heidelberg/New York, 2014)

M.P. Schambach, Recurrent HMMs and cursive handwriting recognition graphs, in *10th
International Conference on Document Analysis and Recognition (ICDAR)*, Barcelona (IEEE
Computer Society, 2009), pp. 1146–1150. http://dblp.uni-trier.de/db/conf/icdar/icdar2009.
html#Schambach09

T. Skopal, J. Pokorný, M. Krátký, V. Snášel, Revisiting M-tree building principles, in
Advances in Databases and Information Systems, ed. by L. Kalinichenko, R. Manthey,
B. Thalheim, U. Wloka. Volume 2798 of Lecture Notes in Computer Science (Springer,
Berlin/Heidelberg/New York, 2003), pp. 148–162

A.F. Smeaton, P. Over, W. Kraaij, Evaluation campaigns and TRECVid, in *Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval*, Santa Barbara, 26–27 Oct 2006 (ACM, New York, 2006), pp. 321–330

D. Sonntag, N. Reithinger, G. Herzog, T. Becker, A discourse and dialogue infrastructure for industrial dissemination, in *Spoken Dialogue Systems for Ambient Environments: Second International Workshop (IWSDS 2010)*, Gotemba, 1–2 Oct 2010, ed. by G.G. Lee, J. Mariani, W. Minker, S. Nakamura. Volume 6392 of Lecture Notes in Artificial Intelligence (Springer, Berlin/Heidelberg/New York, 2010), pp. 132–143. http://www.springerlink.com/content/5149m52mt5378316/

C. Stab, K. Nazemi, M. Breyer, D. Burkhardt, J. Kohlhammer, Semantics visualization for fostering search result comprehension, in *The Semantic Web: Research and Applications*, Heraklion, ed. by E. Simperl, P. Cimiano, A. Polleres, O. Corcho, V. Presutti. Volume 7295 of Lecture Notes in Computer Science (Springer, Berlin/Heidelberg/New York, 2012), pp. 633–646

A. Stoffel, D. Spretke, H. Kinnemann, D.A. Keim, Enhancing document structure analysis using visual analytics, in *Proceedings of the 2010 ACM Symposium on Applied Computing (SAC '10)*, Sierre, Mar 2010, ed. by S.Y. Shin, S. Ossowski, M. Schumacher, M.J. Palakal, C.C. Hung (ACM), pp. 8–12. http://dblp.uni-trier.de/db/conf/sac/sac2010.html#StoffelSKK10

V. Tresp, Y. Huang, M. Bundschus, A. Rettinger, Materializing and querying learned knowledge, in *Proceedings of the First ESWC Workshop on Inductive Reasoning and Machine Learning on the Semantic Web (IRMLeS '09)*, Heraklion, vol. 474 (RWTH Aachen, 2009)

V. Tresp, Y. Huang, M. Nickel, Querying the web with statistical machine learning, in *Towards the Internet of Services: The THESEUS Research Program*, ed. by W. Wahlster, H.J. Grallert, S. Wess, H. Friedrich, T. Widenka (Springer, Berlin/Heidelberg/New York, 2014)

W. Wahlster (ed.), *SmartKom: Foundations of Multimodal Dialogue Systems*. Cognitive Technologies (Springer, Berlin/Heidelberg/New York, 2006)

W. Wahlster, H.J. Grallert, S. Wess, H. Friedrich, T. Widenka (eds.), *Towards the Internet of Services: The THESEUS Research Program* (Springer, Berlin/Heidelberg/New York, 2014)

Z. Wang, A.C. Bovik, *Modern Image Quality Assessment*. Synthesis Lectures on Image, Video, and Multimedia Processing (Morgan & Claypool Publishers, San Rafael, California, 2005)

Further Reading on WP2: "Video, Metadata, Platforms: Processing of Multimedia Data"

U. Engelke, H. Kaprykowsky, H.J. Zepernick, P. Ndjiki-Nya, Visual attention in quality assessment. IEEE Signal Process. Mag. **28**(6), 50–59 (2011). http://dblp.uni-trier.de/db/journals/spm/spm28.html#EngelkeKZN11

S. Han, A. Hutter, W. Stechele, A reasoning approach to enable abductive semantic explanation upon collected observations for forensic visual surveillance, in *International Conference on Multimedia and Expo (ICME)* (IEEE, 2011), pp. 1–7 http://dblp.uni-trier.de/db/conf/icmcs/icme2011.html#HanHS11

S. Han, B. Koo, A. Hutter, V. Shet, W. Stechele, Subjective logic based hybrid approach to conditional evidence fusion for forensic visual surveillance, in *Advanced Video and Signal Based Surveillance (AVSS)*, vol. 7 (IEEE Computer Society, 2010), pp. 337–344. http://dblp.uni-trier.de/db/conf/avss/avss2010.html#HanKHSS10

M. Liu, I.V. Konya, J. Nandzik, N. Flores-Herr, S. Eickeler, P. Ndjiki-Nya, A new quality assessment and improvement system for print media. EURASIP J. Adv. Signal Process. **109**, 1–17 (2012). http://dblp.uni-trier.de/db/journals/ejasp/ejasp2012.html#LiuKNFEN12

E. Mbanya, S. Gerke, P. Ndjiki-Nya, Spatial codebooks for image categorization, in *Proceedings of the 1st ACM International Conference on Multimedia Retrieval (ICRM '11)*, ed.

by F.G.B.D. Natale, A.D. Bimbo, A. Hanjalic, B.S. Manjunath, S. Satoh (ACM, New York, 2011). http://dblp.uni-trier.de/db/conf/mir/icmr2011.html#MbanyaGN11

P. Ndjiki-Nya, M. Köppel, D. Doshkov, H. Lakshman, P. Merkle, K. Müller, T. Wiegand, Depth image-based rendering with advanced texture synthesis for 3-D video. IEEE Trans. Multimed. **13**(3), 453–465 (2011). http://dblp.uni-trier.de/db/journals/tmm/tmm13.html#Ndjiki-NyaKDLMMW11

C. Petersohn, Temporal video structuring for preservation and annotation of video content, in *16th IEEE International Conference on Image Processing (ICIP)*, Cairo, Nov 2009 (Institute of Electrical and Electronics Engineers (IEEE), 2009), pp. 93–96. http://dblp.uni-trier.de/db/conf/icip/icip2009.html#Petersohn09

C. Petersohn, *Temporal Video Segmentation*. PhD thesis, Institute of Technology, Berlin, 2010

T. Schwarze, T. Riegel, S. Han, A. Hutter, S. Nowak, S. Ebel, C. Petersohn, P. Ndjiki-Nya, Role-based identity recognition for TV broadcasts. Multimed. Tools Appl. **63**, 501–520 (2013)

F. Stegmaier, M. Döller, H. Kosch, A. Hutter, T. Riegel, AIR: architecture for interoperable retrieval on distributed and heterogeneous multimedia repositories, in *11th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, Desenzano del Garda, Apr 2010 (IEEE, 2010), pp. 1–4 http://dblp.uni-trier.de/db/conf/wiamis/wiamis2010.html#StegmaierDKHR10

F. Stegmaier, M. Döller, K. Schlegel, H. Kosch, S. Seifert, M. Kramer, T. Riegal, A. Hutter, M. Thoma, H.P. Kriegel, A. Cavallaro, Generische Datenintegration zur semantischen Diagnoseunterstützung im Projekt THESEUS MEDICO, in *41. Jahrestagung der Gesellschaft für Informatik: Datenmanagement und Interoperabilität im Gesundheitswesen (DIG)*, Berlin, Oct 2011

Further Reading on WP5: "Innovative User Interfaces and Visualization"

M. Breyer, K. Nazemi, C. Stab, D. Burkhardt, A. Kuijper, A comprehensive reference model for personalized recommender systems, in *Human Interface and the Management of Information. Interacting with Information*, ed. by M.J. Smith, G. Salvendy. Volume 6771 of Lecture Notes in Computer Science (Springer, Berlin/Heidelberg/New York, 2011), pp. 528–537. http://dblp.uni-trier.de/db/conf/hci/hci2011-11.html#BreyerNSBK11

D. Burkhardt, C.E. Hofmann, K. Nazemi, C. Stab, M. Breyer, D.W. Fellner, Intuitive semantic-editing for regarding needs of domain-experts, in *Proceedings of ED-Media 2010, World Conference on Educational Multimedia, Hypermedia and Telecommunications* (AACE, 2010), pp. 860–869

T. May, J. Kohlhammer, Towards closing the analysis gap: visual generation of decision supporting schemes from raw data. Comput. Graph. Forum **27**(3), 911–918 (2008). http://dblp.uni-trier.de/db/journals/cgf/cgf27.html#MayK08

C. Stab, K. Nazemi, M. Breyer, D. Burkhardt, J. Kohlhammer, Semantics visualization for fostering search result comprehension, in *The Semantic Web: Research and Applications*, ed. by E. Simperl, P. Cimiano, A. Polleres, O. Corcho, V. Presutti. Volume 7295 of Lecture Notes in Computer Science (Springer, Berlin/Heidelberg/New York, 2012), pp. 633–646. http://dblp.uni-trier.de/db/conf/esws/eswc2012.html#StabNBBK12