Intelligent Interaction and Incremental Knowledge Acquisition for Radiology Images

Daniel Sonntag

German Research Center for AI (DFKI)
Stuhlsatzenhausweg 3, 66123 Saarbruecken, Germany

Abstract. Today, the major challenge in medical imaging is the so-called knowledge acquisition bottleneck. We cannot acquire the necessary medical image knowledge that ought to be used in the software application easily as it is hidden in the heads of medical experts. In this paper, we provide an example of how an incremental knowledge acquisition process for radiology images can be implemented to solve this problem. Thereby, we integrated Semantic Web technologies with a variety of automatic and manual annotation tools for radiology images. According to the complex medical finding process, the different annotation tools should be used for very specific purposes. This divide-and-conquer strategy turns out to be very effective in the radiology domain, but produces many infrastructure requirements and relies on high-end intelligent user interfaces such as dialogue systems which are not always available.

1 Introduction

A prior usability analysis to identify the requirements for industrial applications, where image semantics play a role, is very useful. In many circumstances, different requirements have to be met during knowledge acquisition, refinement, and retrieval. In addition, work from the area of the Semantic Web should be integrated in such a way that the process of using image semantics, and relying on it, does not produce too much knowledge engineering overhead. Unfortunately, in many industrial domains such as medical radiology, a vast amount of images is produced and manual annotations are not feasible. In addition, these medical image annotations must be refined and augmented during a complex medical workflow.

Our clinical partner,the University Hospital Erlangen in Germany, has a total of about 50 TB of medical images. They are currently doing about 150,000 medical examinations producing 13 TB of data per year. Many 2D and 3D image series in radiology, and individual images in particular, require specific semantic annotations of the image contents which cannot be automatically provided (figure 1). These annotations are extremely helpful and increase the quality of patient treatment processes; in addition to satisfying the trend to store and organise all patient data, including health records, laboratory reports, and medical images in digital libraries, effective retrieval of images builds on the semantic annotation of image contents. In the medical domain, the proper selection of specific image contents can improve the treatment process to a large degree since the doctor can consult similar cases and other doctors' treatment plans. This case-based reasoning is very effective in the medical domain. At the same time it is crucial that clinicians have access to a coherent view of image data within their particular diagnosis or treatment context. Semantic annotations should provide the necessary image (region) information.

In order to address these issues, namely the knowledge acquisition bottleneck and different user interface requirements at different medical workflow stages, we designed and implemented an incremental knowledge acquisition process for radiology images. This process relies on an integrated ontology-based approach of structured knowledge for medical images (section 2) and takes the special requirements of the radiology department into account. Based on these requirements, automatic and manual annotation frameworks can be constructed (sections 3 and 4) and combined, thereby implementing an incremental knowledge acquisition process (section 5). Section 6 provides a conclusion.

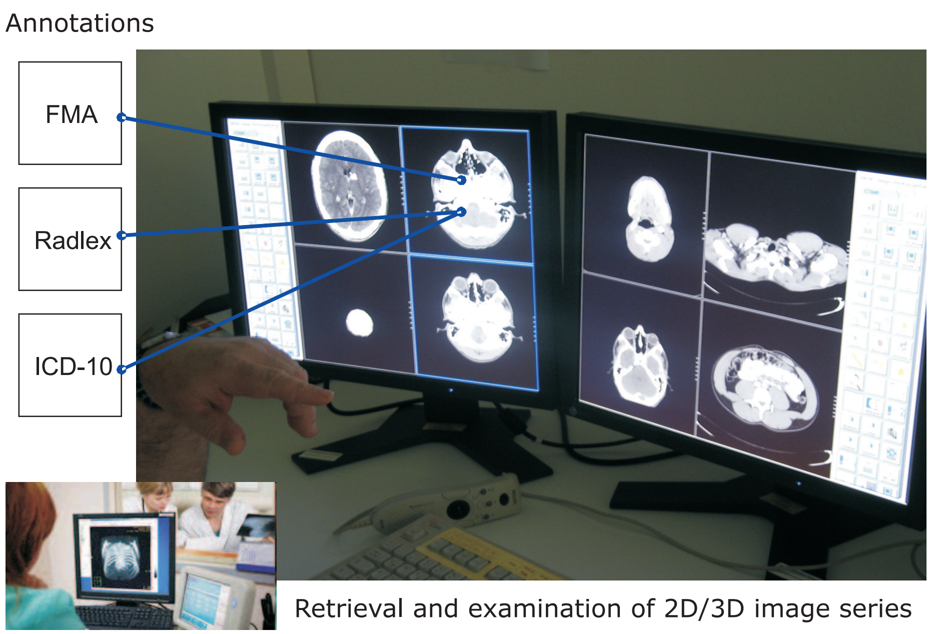

Fig. 1. Image series and semantic annotation requirements

2 Structured Knowledge for Medical Images

Structured knowledge in radiology has a multitude of different aspects, which can be divided into different representational ontologies in RDFS and OWL. The annotations for medical images are based on the assumption that those elements at higher levels are more stable, shared among more people, and thus change less often than those at lower levels. For example, the *Upper Ontology* describes very general concepts like *time*, *space*, *organization*, *person*, and *event*, which are the same across all domains. The *Information Element Ontology* represents the information elements of the incremental knowledge acquisition process (figure 2). For the *Medical Ontologies*, a separation into mid- and low-level ontologies is not so clear since they usually cover a broad spectrum of concepts ranging from very abstract ones like "heart" (which are not very likely to change) to macromolecules (which are updated and added frequently). However, the medical ontologies are

– the Foundational Model of Anatomy (FMA) ontology [7] for anatomical annotations;
– the International Classification of Diseases (ICD-10)[1] for disease annotations; and
– Radlex to express visual features of the visual manifestation of a particular anatomical entity or disease [6].

On the images, any combination of anatomical, disease, and visual annotations is allowed and multiple annotations of the same image region are possible. As a result, all messages transferred between internal and external components which deal with image contents are then based on RDF data structures which are modelled in the respective ontology instances (also cf. [1, 4, 14]). This is only possible when all the annotation ontologies are available in the respective format. Especially for the most critical disease part, the ICD-10 was not available in OWL, although the biomedical ontology community has focussed on establishing interoperability and data integration. Several country- and language-specific adaptations of ICD-10 exist which share the general structure of the WHO version but differ in certain details. We presented an approach for modelling the hierarchy of the ICD-10 using OWL so that we can easily use it in our ontology framework and have enough expressivity to convey special data relations (figure 2(1)). For example, specialties such as "Exclusion" statements, which make statements about the disjointness of certain ICD-10 categories, are modelled in a formal way. The important thing is that we crawled the necessary data from the language-specific ICD-10 web pages, and this procedure can be transferred to other image semantic domains, as far as the necessary image terminology is available online. Noy and Rubin have also presented an approach for translating the Foundational Model of Anatomy ontology (FMA) to OWL [10]. (From their approach we adopted the idea to split the generated ontology into an OWL-DL

[1] http://www.who.int/classifications/apps/icd/icd10online

and an OWL-Full component.) The resulting integrated data model is a prerequisite to get accurate annotations on decision-relevant image contents in medical imaging.

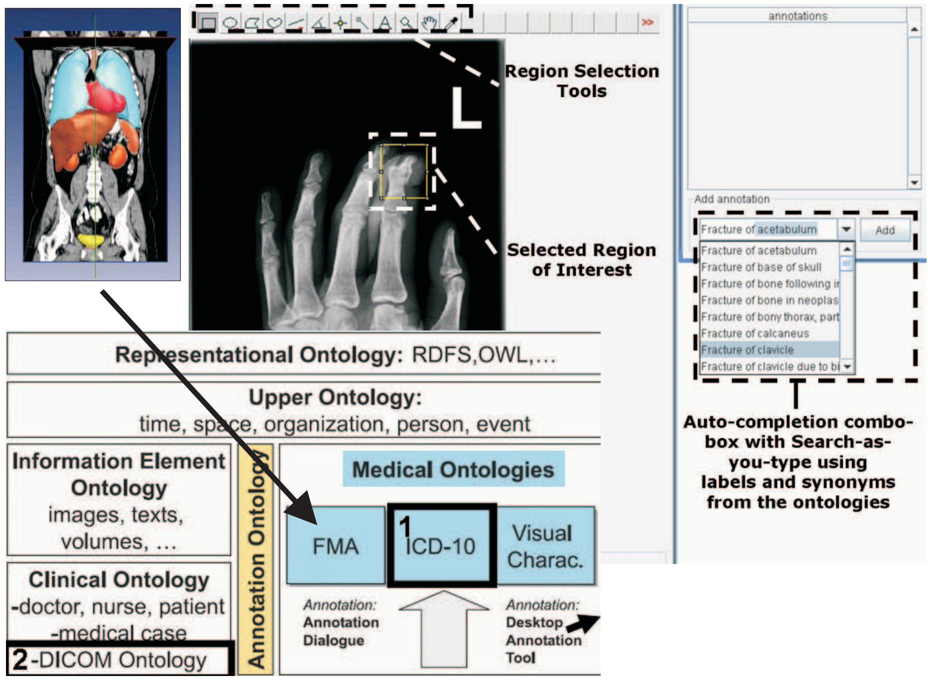

Fig. 2. Knowledge structure and automatic and manual annotation

3 Automatic Annotation

Automatic annotation of medical images has three basic components. First, you can extract knowledge from metadata that is produced during the image generation process. Second, you can use image recognition software to detect anatomical concepts and landmarks. Third, you can try to reason about the plausibility of special configurations being detected while using ontological background knowledge.

DICOM Standard The Digital Imaging and Communications in Medicine (DICOM) Standard (http://medical.nema.org/) ensures the interoperability of information on medical images. Manufacturers of imaging equipment and imaging information systems and manufacturers of peripheral equipment (e.g., computer monitors and image archives) conform to this standard. The Siemens computer tomography (CT) and magnetic resonance imaging (MRI), which we used to produce our image material, use this standard to encode a multitude of image metadata about the image generation process. Figure 3 shows the subset of these data we extract from the image headers in order to create ontology instances automatically (cf.the ontology model in figure 2(2)). As can be seen, we can extract about fifty image features in the context of the image study, the patient, and low level image characteristics. These metadata provide the necessary information to create the patient image instances to be augmented by the image content semantics of specific image regions.

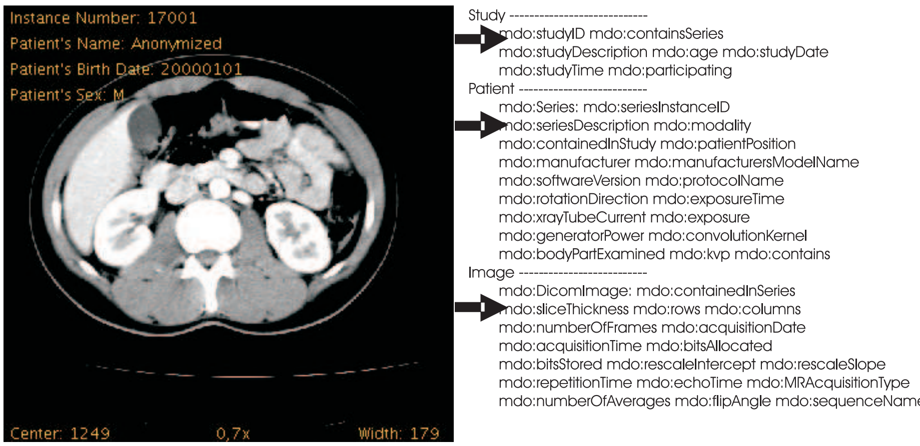

Fig. 3. DICOM data that can be extracted from the image header

Image Recognition The CT and MRI systems produce detailed pictures of organs, soft tissues, bone, and virtually all other internal body structures. Today, organs of the chest and abdomen—including the heart, liver, biliary tract, kidneys, spleen, bowel, pancreas—can be detected with great accuracy. But the automatic detection of image semantic of, e.g., malicious tissue in the context of cancer, is extremely difficult. Although we use state-of-the-art organ and landmark detection software [12] with a special focus on organs, landmarks (also cf. top left of figure 2), and lymph node segmentation [2], many further reasoning and manual annotation steps are necessary.

Reasoning with Special Configurations The ontology knowledge structures become effective when axiomatic relations apart from subsumptions can be exploited. Spatial relations are a promising area of research in this automatic reasoning area since they complement well-known linguistic phenomena being put into the ontology context (e.g., Wordnet relations) and at the same time allow both the automatic modelling of special configurations and the human judgement/evaluation for plausibility. One idea we evaluated is the incorporation of a spatio-anatomical ontology for automatic plausibility checks of the found configuration of automatically detected organs [8]. We first learned a model of plausible organ constellations inductively from an annotated corpus of 3D volume data sets. The model, an ontology-based canonical representation of the spatial relationships of organs in the human body, can be used to check the results of a state-of-the-art medical object recognition system for 3D CT volume data sets for spatial plausibility.

The interesting thing is that, on a dataset of 1118 instances, the model produces only 76 false positives and 213 false negatives. This means that while precision is relatively high, the recall is moderate with 65.5%. As a result, a lot of further manual control is needed to find the erroneous automatic recognition results. This is one of the reasons why manual annotations are needed not only for the disease, but also for the anatomical level on medical images in radiology.

4 Manual Annotation

Manual annotation means that the radiologist must use a special human-computer interaction system to perform the required image annotations. This process reveals many usability issues. We will first describe what the desktop workstation and the special multi-touch installation in combination with a dialogue system looks like, before we discuss the usability issues in the context of the combined incremental process.

Desktop Workstation For the manual semantic annotation on a regular desktop workstation we developed a new medical semantic annotation and retrieval tool RadSem [9]. It consists of a component that implements a method to annotate images and upload/maintain a remote RDF repository of the images and image semantics. In order to ease the task of finding appropriate annotations, we use *auto-completing* combo-boxes.

A screenshot of parts of the annotation tool is depicted in figure 2 (right) which shows a simple orthopaedic example. The broken bone of the index finger can easily be annotated while using the auto-completion combo-boxes with a search-as-you-type functionality. The resulting annotation is accurate but very time-consuming.

Radiology Dialogue System It is crucial that clinicians have access to a coherent view of image data within their particular diagnosis or treatment context (we experimented with a large touchscreen installation). These data include previous (rudimentary) annotations. A semantic dialogue shell should be used to ask questions about the image annotations and refine them while engaging the clinician in a natural speech dialogue at the same time. In the construction of a dialogue system for radiologists, we learned some lessons which we used as guidelines in the development of *semantic* dialogue systems [11, 15]; over the last years, we have adhered strictly to the developed rule “No presentation without representation.” All the items presented on the touchscreen are basically surface representations of more complex ontological entities according to the described knowledge structure. This knowledge structure (section 2) allows a specific user to ask questions about the displayed image content and other region-based image elements. The domain-specific dialogue application for the radiology department (also cf. [16]), which uses a touchscreen (figure 4, upper right) to display the medical image windows, is able to process the following dialogue:

- **U:** “Show me the CTs, last examination, patient XY.”
- **S:** Shows corresponding patient CT studies as DICOM picture series and MR videos.
- **U:** “This lymph node here (+ pointing gesture) is enlarged; so add the annotation: *lymphoblastic*.”
- **S:** Shows new annotation on the image and confirms database update.

The dialogue-based annotation can be done at a rate of approximately 6 annotations per minute (including the visual feedback phase) whereas the desktop-based annotation comes to a rate of approximately 3 annotations per minute. Most importantly, the prototype dialogue system delivers new semantic annotations instantly which are unavailable in the current clinical finding process so that the (senior) radiologist can directly detect errors visually.

5 Incremental Knowledge Acquisition Process

The incremental knowledge acquisition process (figure 4) relies on the structured ontological knowledge as introduced in the first section. Based on this prerequisite, we have been trying to formulate the process of automatic and manual image annotation. Hereby, two factors play a major role: the quality of automatic annotations and the usability of different intelligent user interfaces to control, correct, and add annotations. For us, usability means that people can use an Artificial Intelligence (AI) prototype easily and efficiently to accomplish their tasks. Prototypes that are usable enable clinicians to concentrate on their tasks rather than paying attention to the tools they use to perform their tasks. The prevalent interaction design issue that follows this definition is that the intelligent interfaces are

- efficient to use;
- quick to recover from errors; and
- visually pleasing.

To achieve all three of these a careful selection of involved components for manual annotation is vital. This can be substantiated by the current developments in clinical practice where *structured reporting* should be introduced. This means that the radiologists fill in special standardised forms. Radiologists feel restricted by these standardised forms and fear a decrease in focus and eye dwell time on the images [3, 17]. As a result, the acceptance for structured reporting is still low among radiologists while referring physicians and hospital administrative staff are generally supportive of structured standardised reporting since it can be used more easily for further processing. As a matter of course, the image semantics with RDF are a further step in this direction. These issues are explained in the context of industrial usability and our basic process steps for industrial dissemination.

5.1 Binocular View and Industrial Usability

As [5] point out, many research prototypes that use technically advanced but unimportant or unrealistic functionality for the specific domain or personal activities do not provide the AI support that users would appreciate most. This can, e.g., make a complex speech dialogue system languish as an infertile research prototype on demonstration computers which cannot be used in the context of industrial prototypes or real-world industrial dissemination. Accordingly, the binocular view of intelligent interfaces for industrial dissemination should study not only the suitability of a single algorithm and a component performance for a given user task, but also the industry user's interaction requirement in which the interaction will be used. In our specific radiology case, the feature that only a senior radiologist is responsible for the treatment plan, implicates that his or her interaction with the annotation system must be designed to be very effective. Although it is widely reductive to put it this way, a senior radiologist has three main goals: (1) access the images and image (region) annotations (a summary can also be synthesised), (2) complete them, and (3) refine existing annotations. These tasks can best be fulfilled while using a multimodal dialogue system. In contrast, less demanding manual annotation tasks, such as the correction of organ detection algorithms of image region selection can be done by, e.g., a first-year resident with the help of our desktop-based annotation tool. This tool can also easily be installed on virtually every computer in a hospital, whereas a speech dialogue system requires a specific hardware infrastructure.

5.2 Process Steps

The incremental knowledge acquisition process (figure 4) has four steps.

First, the automatic metadata are extracted from the DICOM images and instantiated according to the structured/structural knowledge model. After that, a direct access to the RDF statements is possible while using, e.g., the query language SPARQL.

Second, the automatic image recognition software runs over the images to produce anatomical annotations according to the structural knowledge model. According to the spatio-anatomical ontology, automatic spatial plausibility checks can be executed. Hereby, the spatial reasoning process runs completely automatically and only the outlier configurations are presented to the medical experts.

Third, the experts can then use the manual annotation tool to correct or extend these configurations. At this stage, a very comprehensive set of image semantics, namely the study, patient, and low-level image feature information in combination with the automatically detected anatomical concepts and manual annotations with the desktop tool are available. These image model instances are not accurate enough for a treatment plan, but accurate enough to be used in a semantic search and annotation system, the dialogue shell, which the senior radiologist can use. For this purpose, the image annotations are accurate enough. The point is that annotation is accurate enough, or timely enough at this stage. Then making it more accurate sooner is unnecessary and will increase costs without increasing benefits for the treatment process.

Fourth, only when the images are retrieved and considered for a medical treatment plan, can accurate disease annotations be added by the senior radiologist while using the dialogue system which displays the image and patient data on a large touchscreen. It is even possible to search for similar disease annotations in other patients' contexts for a comparable study. Currently, we are trying to extend the high-level process of patient findings and image annotations to a mobile scenario, where we can use a special pen to recognise annotations on normal paper and/or used the iPad as a mobile dialogue system and touchscreen device for the senior radiologist (also cf. the project Radspeech, http://www.dfki.de/RadSpeech/).

Our hope is that the resulting process successfully supports the complex healthcare process in which radiology images are used. The development of automatic processing applications is as essential as the design and implementation of intelligent user interfaces for specific purposes. In our view, only this combination will produce successful decision support systems for industrial dissemination.

6 Conclusion

In discussions with radiologists we found out that three typical clinical scenarios are of interest for further analysis of clinical knowledge requirements and (incremental) knowledge acquisition: (1) the clinical reporting process; (2) the patient follow-up treatment (i.e., monitoring the patients health condition and the development of the disease); and (3) the clinical disease staging and patient management. In this paper we have explained a process that takes structured medical knowledge as input and provides an incremental process for the clinical disease staging process by addressing the bottleneck to annotate appropriate image semantics.

We explained exemplarily how to generate a multi-lingual OWL model of the ICD-10, and how this fits in the annotation framework as a prerequisite. In addition, we provided automatic and manual annotation scenarios and a MEDICO server architecture with several HCIs/dialogue systems to meet the requirements of a distributed software infrastructure and/or usability issues. These issues have been explained in the context of our basic architecture approach for industrial dissemination. An incremental knowledge acquisition process for radiology im-

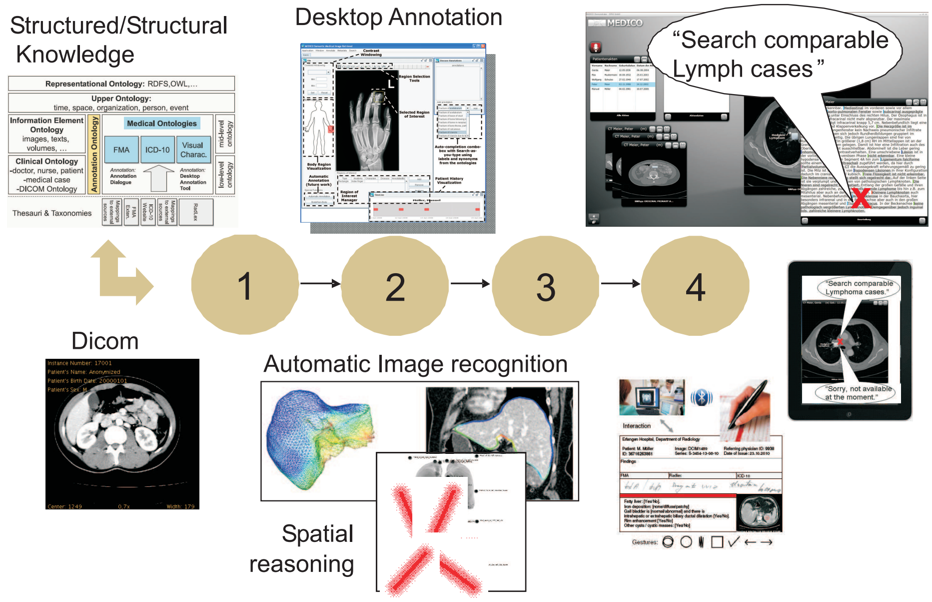

Fig. 4. Incremental Knowledge Acquisition Process

ages seems to be adequate. But we produced many infrastructure requirements and relied on high-end speech-based dialogue systems which are not available in the industrial sector today.

The question of how to integrate the acquired image knowledge with other types of data, such as patient data, is paramount. In a further step, individual textual findings should be organised according to a specific body region and the disease context both of which can be interlinked to several text passages. Currently, we are evaluating the proper usage of information extraction technology for this purpose. The main problem is that the text processing tools cannot be easily adapted to the medical domain. Finally, educators may find our process can help trainees learn the important elements of reports and will encourage the proper use of radiology terms (structured reporting). We hope that structured reporting will also help to ease the task of text mining.

Acknowledgements

I would like thank all colleagues at DFKI and all partners in the MEDICO and RadSpeech projects for their valuable contributions to the iterative process described here. This research has been supported by the THESEUS Programme funded by the German Federal Ministry of Economics and Technology (01MQ07016).

References

1. D. Fensel, J. A. Hendler, H. Lieberman, and W. Wahlster, editors. *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press, 2003.
2. J. Feulner, S. K. Zhou, M. Huber, J. Hornegger, D. Comaniciu, and A. Cavallaro. Lymph node detection in 3-d chest ct using a spatial prior probability. In *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010*, pages 2926–2932, 2010.
3. F. M. Hall. The radiology report of the future. *Radiology*, 251(2):313–316, 2009.
4. P. Hitzler, M. Krötzsch, and S. Rudolph. *Foundations of Semantic Web Technologies*. Chapman & Hall/CRC, August 2009.
5. A. D. Jameson, A. Spaulding, and N. Yorke-Smith. Introduction to the Special Issue on Usable AI. *AI Magazine*, 3(4):11–16, 2009.
6. C. P. Langlotz. Radlex: A new method for indexing online educational materials. *RadioGraphics*, 26:1595–1597, 2006.
7. J. L. Mejino, D. L. Rubin, and J. F. Brinkley. FMA-RadLex: An application ontology of radiological anatomy derived from the foundational model of anatomy reference ontology. In *Proc. of AMIA Symposium*, pages 465–469, 2008.
8. M. Möller, P. Ernst, D. Sonntag, and A. Dengel. Automatic spatial plausibility checks for medical object recognition results using a spatio-anatomical ontology. In *Proc. of the International Conference on Knowledge Discovery and Information Retrieval (KDIR 2010)*, Valencia, Spain, 25 - 28 October 2010.
9. M. Möller, S. Regel, and M. Sintek. Radsem: Semantic annotation and retrieval for medical images. In *Proc. of The 6th Annual European Semantic Web Conference (ESWC2009)*, June 2009.
10. N. F. Noy and D. L. Rubin. Translating the foundational model of anatomy into owl. *Web Semant.*, 6:133–136, April 2008.
11. S. Oviatt. Ten myths of multimodal interaction. *Communications of the ACM*, 42(11):74–81, 1999.
12. S. Seifert, M. Kelm, M. Möller, S. Mukherjee, A. Cavallaro, M. Huber, and D. Comaniciu. Semantic annotation of medical images. In *Proceedings of SPIE Medical Imaging*, San Diego, CA, USA, 2010.
13. D. Sonntag. Distributed NLP and Machine Learning for Question Answering Grid. In *Proceedings of the ECAI 2004 Workshop on Semantic Intelligent Middleware for the Web and the Grid*, Valencia, Spain, 2004.
14. D. Sonntag. *Ontologies and Adaptivity in Dialogue for Question Answering*. AKA and IOS Press, Heidelberg, 2010.
15. D. Sonntag, R. Engel, G. Herzog, A. Pfalzgraf, N. Pfleger, M. Romanelli, and N. Reithinger. *SmartWeb Handheld—Multimodal Interaction with Ontological Knowledge Bases and Semantic Web Services*, volume 4451 of *Lecture Notes in Computer Science*, pages 272–295. Springer, 2007.
16. D. Sonntag and M. Möller. Unifying semantic annotation and querying in biomedical image repositories. In *Proceedings of International Conference on Knowledge Management and Information Sharing (KMIS)*, 2009.
17. D. L. Weiss and C. Langlotz. Structured reporting: Patient care enhancement or productivity nightmare? *Radiology*, 249(3):739–747, 2008.