Towards the Ontology-based Classification of Lymphoma Patients using Semantic Image Annotations

Sonja Zillner

Corporate Technology, Siemens AG
Munich, Germany
sonja.zillner@siemens.com

Abstract. Today, clinicians rely on medical images for screening, diagnosis, treatment planning and follow up, but still a generic and flexible medical image understanding is missing. Although, there exist several approaches for semantic image annotation, those approaches do not make use of practical clinical knowledge, such as best practice solutions or clinical guidelines. Our final goal is to enhance medical image annotations by integrating clinical knowledge, such as lymphoma staging systems. The contribution of this paper is to introduce a formal approach to the classification of patients in well-defined categories. As first step, we have developed an OWL DL ontology representing the Ann-Arbor Lymphoma staging system that is suitable for performing automatic patient classification. Our aim for the ontology design was to establish means for automatic staging that maps each patient to one staging degree.

1 Introduction

The vision of MEDICO is to automatically extract the meaning from the medical images and to seamlessly integrate the extracted knowledge into medical processes, such as clinical decision making. In other words, the computer shall learn to find, catalogue and interpret medical images. This requires the semantic representation of medical images' content and the preprocessing of semantic image annotations for seamless integration into clinical applications.

There exist several approaches for semantic image annotation, such as automatic image parsing [1], manual image annotation [2][3], the extraction of information from DICOM headers and DICOM structured reports [4], or the automated extraction from radiology reports. Although those approaches provide the very important basis for semantic image annotation, they do yet not make use of practical clinical knowledge, such as best practice solutions or clinical guidelines for fine-tuning and customizing the established annotations to reflect the particular requirements of a clinical application or work flow. Our final goal is to enhance medical image annotations by integrating clinical knowledge, such as lymphoma staging systems. The contribution of this paper is to introduce a formal approach to the classification of patients in well-defined categories. Our approach is based on external medical taxonomies and ontologies which promote re-usability and interoperability. We have focused our study on the Ann-Arbor staging classification of Hodgkin-lymphoma as defined in [5] as external clinical knowledge resource. Our aim for the ontology design was to establish means for an automatic staging system that maps each patient to uniquely one staging degree. We, thus, can automatically generate additional patient annotation data, that can be used for supporting the clinicians in their daily tasks. By integrating the patient's staging information, into other clinical applications, for instance, the clinical work-flow can be optimized or the search and comparison of patients be improved. We have developed an OWL ontology that represents the Ann Arbor staging system together with lymphoma patient records and that is suitable for performing automatic lymphoma patient classification.

The remainder of the paper is organized as follows. In Section 2 we will introduce the knowledge resources our approach is based on. Section 3 details challenges we faced and design decisions we made when developing the ontology for lymphoma patient classification. In Section 4 we will discuss related approaches and Section 5 concludes this paper with an outlook on future work.

2 Knowledge Resources

Ann Arbor Hodgkin Lymphoma classification Ann Arbor staging (Fig. 1) is the staging system for lymphomas. It was initially developed for Hodgkin's Lymphoma, but has some use in Non-Hodgkin lymphomas. The stage depends on two criteria. The first criterion is the place where the malignant tissue is located. The location can be identified with located biopsy as well as with medical imaging methods, such as CT scanning and increasingly positron emission tomography. The second criterion are systemic symptoms, such as night sweats, weight loss of more than 10 percent or fevers, caused by the lymphoma. Those systemic systems are called "B symptoms".

The principal stage is determined by the location of the tumor and reflects the grade of expansion of lymphoma occurrences. Four different stages are recognized: Stage I indicates that the cancer is located in a single region, either an affected lymph node or organ within the lymphatic system. In Stage II the cancer is located in two separated regions, an affected lymph node or an affected organ within the lymphatic system and a second affected lymph node area. Moreover, the affected areas are confined to one side of the diaphragm - that is, both are above the diaphragm or, both are below the diaphragm. Stage III indicates that the cancer has spread to both sides of the diaphragm, including one extra lymphatic organ or site. Stage IV shows diffuse or disseminated involvement of one or more extra lymphatic organs.

Ontological Knowledge Resources For capturing the semantics to the classification of lymphoma patients and for achieving re-usability and interoperability, we required third party taxonomies or ontologies that cover all possible regions of lymphatic occurrences. Two ontologies, the Foundational Model of Anatomy

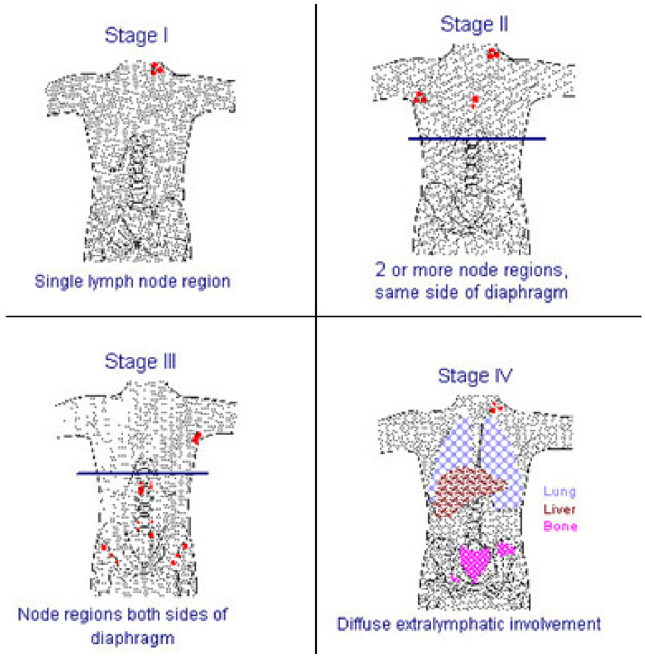

Fig. 1. Ann Arbor Staging System (Source: http://training.seer.cancer.gov)

(FMA)[6] and the Radiology Lexicon (RadLex)[1], provide the required coverage of anatomical concepts for the staging scenario.

The FMA is developed and maintained by the School of Medicine of the University of Washington and the US National Library of Medicine. Besides the specification of an anatomy taxonomy, i.e. an inheritance hierarchy of anatomical entities, the FMA provides definitions for conceptual attributes, part-whole, location, and other spatial associations of anatomical entities. FMA covers approximately 70,000 distinct anatomical concepts and more than 1.5 million relations instances from 170 relation types. It provides concepts that describe single lymph nodes, such as 'axilliary_lymph_node', and concepts that describe multiple lymph nodes, such as 'set_of_axilliary_lymph_node'. It contains 425 concepts representing singular lymph nodes and 404 concepts describing sets of lymph nodes. The FMA is freely available as a Protégé 3.0 project or can be accessed via the Foundational Model Explorer. There also exist conversions of the frame-based Protégé version of FMA to the OWL DL format [7][8].

RadLex is a terminology developed and maintained by the Radiological Society of North America (RSNA) for the purpose of uniform indexing and retrieval of radiology information, including medical images. RadLex contains over 8,000 anatomic and pathologic terms, also those about imaging techniques, difficulties and diagnostic image qualities. Its purpose is to provide a standardized ter-

[1] www.radlex.org

minology for radiological practice. RadLex is available in English and German language and covers 144 concepts describing lymph node concepts.

As the use case scenario relies only on high-level concepts, both ontologies - FMA and RadLex - are suitable in terms of coverage and both ontologies are used within the MEDICO project for annotating medical images and radiology reports. For this use case, we decided to use RadLex as primary third party resource for labeling lymphatic occurrences. As we are working with radiology reports in German language, the availability of a German translation was an crucial argument for this decision. Due to the large size of Radlex, we needed to establish an ontology fragment that is scalable and efficient for reasoning application. Moreover, we transformed the ontology fragment into an appropriate OWL DL format. The manually created ontology fragment covers all RadLex concepts describing lymph nodes as well as a list of likely regions for extra nodal lymphatic occurrences covering concepts such as liver, skin, or bone.

Semantic Image Annotation The MEDICO project is based on multiple ways of generating semantic image annotations. For instance methods for automated image parsing, such as [1], allow hierarchical - in terms of starting with the head and subsequently moving down the body - parsing whole body CT images and efficiently segment multiple organs taking contextual information into account. While automated image parsing remains incomplete, manual image annotation remains an important complement. In addition, ongoing work in MEDICO project is focusing on the semi-automatically identification of terms and relations in radiology reports that are generated by clinicians in the process of analyzing the patient's disease patterns by investigating medical imaging data. The extracted knowledge always relates to a particular medical image and can be used for the generation of further image annotation data.

3 Ontology-based Classification of Lymphoma Patients

Our aim is to establish an ontology that enables the automatic classification of lymphoma patients by means of ontology-based reasoning services. For doing so, we transformed the Ann Arbor Hodgkin Lymphoma classification [5] into a formal and ontology-based representation suitable for reasoning and classification tasks. For capturing the requirements of the ontology design, we followed the formal approach of [9]. We used OWL DL – MEDICO's agreed semantic representation language – for representing the knowledge model. Description Logics, a family of formal representation languages for ontologies, are designed for classification-based reasoning. We developed two different version of ontologies; the first ontology model represents all patients as classes and the second all patients as instances.

In a first step, the Ann Arbor Hodgkin Lymphoma Classification was translated into complex questions for the subsequent decomposition into more simple and manageable queries. Both, the complex questions as well as the simple questions, provided us valuable guidance for evaluating the ontological commitment that has been made. This was possible by systematically establishing test patient classes and respectively patient instances according to the simple and the complex questions.

3.1 Decomposition of Queries

As already mentioned, the stages of the Ann Harbor Staging System classify lymphoma patients in accordance to their number and location of lymph node occurrences and extra lymphatic organ or site involvement. For accessing the number of lymph node occurrences and respectively extra lymphatic organ or site involvement, the following queries, i.e. defined OWL classes, were defined:

- The OWL defined Classes `N0`, `N1` and `N2` identify all patients with zero, one or two and more involved lymph node regions.
- The OWL defined Classes `E0`, `E1` and `E2` identify all patients with zero, one or two and more involved extra lymphatic organ or site involvement.
- The OWL defined Classes `N_AllAboveD` and `N_AllBelowD` identify patients with the location of the lymph node regions either only above or only below the diaphragm. For accessing patients with occurrence of lymph node on both sides of the diaphragm, one has to make sure that the occurrences neither are located all above or located all below the diaphragm. This can be formulated by the following axiom:

$$\neg\ \texttt{N\_AllAboveD} \sqcup \neg\ \texttt{N\_AllBelowD}$$

- The location of extra nodal occurrences are identified in an analogous manner, i.e. by establishing the defined classes `E_AllAboveD` and `E_AllBelowD`, as well as the corresponding complex axiom for accessing patients with extra nodal occurrences on both sides of the diaphragm.

In a further step, the above established simple or auxiliary queries could be used for specifying defined classes deducing the Ann Harbor Stages. Before detailing the staging classification, we will discuss the particular requirements towards the ontology design for realizing the auxiliary queries.

3.2 Implications and Requirements for the Ontological Model

For establishing the ontological model basically three different challenges needed to be addressed:

1. How to ensure that the ontological model provides the basis for inferring the number of lymph node occurrences and the number of extra lymphatic involvements?
2. OWL DL is based on the Open World Assumption (OWA), thus, it is only applicable to a limited extent for checking the validity of closed sets of information. Moreover, the patient diagnosis deals with increasing set of information. How to represent the patient staging classification dealing with an open set of information in combination with the OWA paradigm?
3. How to identify the relative location – above, below or on both sides of the diaphragm – of the lymphatic occurrences?

Counting Lymphatic Occurrences Although in most programming languages counting items is quite straight forward, this is not the case for OWL ontologies. Counting lymphatic occurrences requires some preparation to achieve the intended classification. A reasoner is only able to count concepts that are different. To provide the basis for counting the number of involved regions, the concepts for lymph nodes and for extra nodal lymphatic regions in the established RadLex fragment need to be labeled as disjoint. Subsumption in OWL means necessary implications, thus, classes and their subclasses can not be labeled as disjoint. Therefore, we decided to describe the parent-child relationship of RadLex by introducing a transitive is_a relationship and by using the following type of axioms

$$\begin{aligned} & \texttt{Tr(is\_a)} \qquad (\mathit{Transitive}) \\ \texttt{Right\_Hilar\_Lymph\_Node} \sqsubseteq & (\texttt{Anatomical\_Structure} \sqcap \exists \texttt{is\_a.Hilar\_Lymph\_Node} \\ & \sqcap \exists \texttt{hasNLocation.aboveDiaphragm}) \\ & \vdots \end{aligned}$$

Additional, we entered disjointness axioms indicating that the lymph nodes as well as the extra nodal lymphatic regions are not overlapping.

Open World Reasoning with Increasing Patient Information Our aim for the ontology design was to establish means for automatic staging that maps each patient to uniquely one staging degree. Due to open world reasoning, this could not be achieved directly. For instance, by establishing a defined class with necessary and sufficient conditions stating the existence of exactly one lymph node occurrence, e.g.

$$\texttt{ExampleStage} \equiv \texttt{Patient}(x) \wedge \exists^{=1} y \ \texttt{hasLymphaticOccurrence}(x, y)$$

the inferred ontological model will never classify patient classes as subclasses of the `ExampleStage`. In open world reasoning, anything might be true unless it can be proven as false. In the context of counting lymphatic occurrences, this translates to: Any patient might have more than the indicated lymphatic occurrences unless it can be proven as false. For instance, a Patient A with one explicit indicated lymph node occurrence can possibly have two or more lymphatic occurrences unless explicitly stated differently. Inferring Patient A as sub-class of a staging class consisting of patients with exactly one lymphatic occurrence would bear the risk of future inconsistencies. However, staging classes specified by a minimum number of occurrences, such as $\exists^{>=1} y$ $\texttt{hasLymphaticOccurrence}(x, y)$ are suitable for OWL DL reasoning. But this again will cause, for instance, that a patient with three indicated lymphatic occurrences on both side of the diaphragm is mapped to more than one staging class, i.e. stage I and stage III. Our goal is to achieve the unique mapping between patients and staging classes. This becomes possible by interpreting the highest staging class as the unique mapping result. Yet, this interpretation is not explicitly stated in the derived

OWL DL model and has to be reflected when realizing clinical applications that make use of the automatically inferred classification results.

The Location of Lymphatic Occurrences The patient staging results distinguish between patient that have lymphatic occurrences only above, only below or on both sides of the diaphragm. Thus, the relative position of lymphatic occurrences to the diaphragm has to be expressed in the knowledge model. For achieving this, we see two alternatives. The information about the relative position of lymphatic occurrences can either be derived by the image segmentation algorithm or be directly derived from the underlying ontological model. In other words, this information either comes with the patient record information or with the integrated anatomy and radiology knowledge model. The computation of relative spatial positions of lymphatic occurrences by segmentation algorithms will be addressed within our future work. In the meantime, we have to rely on the anatomy and radiology ontology for deriving this knowledge. RadLex and FMA do not explicitly capture information about relative positions between lymphatic regions or organs and the diaphragm. Thus, we required to enhance the RadLex fragment accordingly. This could be achieved by extending each lymphatic region by an axiom indicating that the region is above, or respectively, below the diaphragm. The classification of lymphatic regions - and of patients with lymphatic occurrences - above, below or on both sides of the diaphragm was modeled as value partition.

3.3 Representation of the Ann-Arbor Staging Classes

The Ann-Arbor Staging classes are represented as defined OWL DL classes. Each staging class is capturing the semantics as detailed in Subsection 2. Their formal representation makes use of auxiliary classes introduced in Subsection 3.1. Figure 2 summarizes the Ann-Arbor Staging Formalization. Thus, for instance, `Stage-II-N` gathers all patients with more that two lymph node region that are all on one side of the diaphragm and `Stage-II-mixed` all patients with one involved lymph node regions and one involved extra lymphatic organ or site on the same side of the diaphragm, and so on.

3.4 Evaluation

Evaluation is an ongoing topic. For evaluating the simple queries as well as the complex staging queries, by hand we systematically generated the corresponding test patient classes and respectively test patient instances. Each test patient is equipped with different numbers and kinds of lymphatic occurrences at different locations in the body. In this way, the evaluation (as well as the development) of the classification axioms was straight forward.

Moreover - in our ongoing work - we are using real radiology reports for challenging the capabilities of our ontology model. Our goal is to semantically annotate more than 250 radiology reports that provide the basis for translating

Stages	Description
Stage-0	N0 *or* E0
Stage-I	N1 *or* E1
Stage-II	Stage-II-N *or* Stage-II-mixed
Stage-II-N	N2 *and* (N_AllAboveD *or* N_AllBelowD)
Stage-II-mixed	N1 *and* E1 *and* ((N_AllAboveD and E_AllAboveD) *or* (N_AllBelowD *and* E_AllBelowD))
Stage-III	Stage-III-N *or* StageIII-mixed
StageIII-N	N2 *and* *not* N_AllAboveD *and* *not* N_AllBelowD
StageIII-mixed	E1 *and* N1 *and* (*not* (N_AllAboveD and E_AllAboveD) *and* *not* (N_AllBelowD and E_AllBelowD))
StageIV	E2 *and* (*not* N_AllAboveD *and* *not* N_AllBelowD)

Fig. 2. Formal Representation of Ann-Arbor Staging

the clinical findings - in terms of numbers and types of lymphatic occurrences - into the OWL representation. As input we are using the semantic annotations (see left side of Fig. 3) which is translated into OWL format (see right side of Fig. 3), i.e. the class based OWL DL representation of a real patient with a large number of lymph nodes who was classified as Stage-III.

One particular challenge we are facing is how to handle multiple lymph node occurrences at the same location. For deciding how to tackle this problem, a detailed analysis of the available patient record will be required.

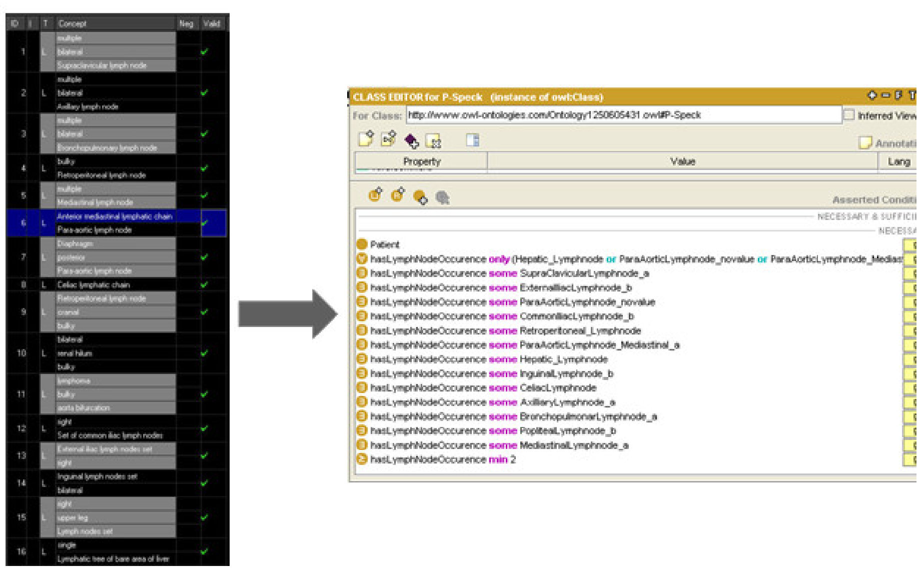

Fig. 3. The semantic annotation of patient records by tool support (left) and the corresponding OWL-based formal representation (right)

4 Related Work

4.1 Semantic Image Understanding

There exist a wide range of different imaging technologies and modalities, such as 4D 64-slice Computer Tomography (CT), whole-body Magnet Resonance Imaging (MRI), 4D Ultrasound, and the fusion of Positron Emission Tomography and CT (PET/CT) providing detailed insight into human anatomy, function and disease associations. Moreover, advanced techniques for analyzing imaging data generating additional quantitative parameters paving the way for improved clinical practice and diagnosis. However, for advanced applications in Clinical Decision Support and Computer Aided Diagnoses the comparative exploration of similar patient information is required. The missing link is a flexible and generic image understanding. Currently, the large amounts of heterogeneous image data are stored in distributed and autonomous image databases being indexed by keywords without capturing any semantics. The vision of the MEDICO project is to automatically extract the meaning from the medical images and to seamlessly integrate the extracted knowledge into medical processes, such as clinical decision making, and to improve clinical workflows. Within the MEDICO project, one of the selected use case scenarios aims for improved image search in the context of patients suffering of lymphoma in the neck area. Lymphoma, a type of cancer originating in lymphocytes, is a systematic disease with manifestations in multiple organs.

Generic medical image understanding is still a long-term agenda due to the high complexity of the problem. Several challenging research questions need to be addressed for tackling this vision. For determining the scope and level of detail of the semantics of the domain, i.e. the relevant metadata for annotating medical images, one needs to find out what kind of knowledge the clinicians are interested in. The scope of the constraint domain can be determined by the set of derived query patterns [10][11] that provide guidance in identifying the relevant (fragments of) ontologies [12]. Moreover, the low level features, segmentations and quantitative measures derived from image processing need to be associated with ontologies.

4.2 Formal approaches to the Classification of Patients

There exist several approaches that analyzed to what extent tumor grading and classification can be performed automatically using the OWL-DL description logic language, such as [13] aiming for the classification of lung tumors and [14] for the classification of glioma tumors. Different kinds of tumors rely on different kinds of staging systems. Whereas, lung and glioma tumors - similar to the most tumor kinds - can be classified by the TNM classification, lymphoma draws on a particular staging system. The so-called Ann-Arbor Classification System for lymphoma depends on the number, type and location of lymphatic occurrences. Thus, it raises different ontological design requirements, such as counting lymphatic occurrences and determining the relative location.

[15] is similar to our approach, inasmuch it introduces an application that provides support for the semantic annotation of medical images. Yet, the external knowledge used for enhancing the semantic annotation and the application focus are different. Our approach aims to integrate external clinical knowledge for enhancing existing image annotation to optimize and improve clinical applications.

OWL DL-based reasoning is also used in the context of other clinical use cases. [15], for instance, relies on the anatomy model and its regional relationships for assisting the labeling of the MRI image content. Due to the facts extracted from MRI images, rather topological relations are capturing the required knowledge. For representing and extracting the topological information, i.e. the interdependencies of properties, reasoning with rules in combination with ontologies is required. [16] aims for improved and concise patient data visualization by incorporation of medical ontological knowledge. The proposed solution uses an OWL DL view of the patient database with external semantics allowing for the patient record classification by a reasoner, from where the inferred hierarchy is directly fed into an appropriate visualization tool.

5 Conclusion

We are using the reasoning capabilities of OWL DL to provide means for the automatic classification of lymphoma patient. The enhanced patient annotation data can be used in a multitude of clinical applications, such as the recommendations for treatments, the search and visualization of similar patients, or clinical studies. In our future work, we are aiming to extend our knowledge model by the "B symptoms" of the Ann-Arbor classification system. Moreover, we will integrate our OWL DL based staging approach into the MEDICO system for enabling improved clinical applications.

Acknowledgements This research has been supported in part by the THESEUS Program in the MEDICO Project, which is funded by the German Federal Ministry of Economics and Technology under the grant number 01MQ07016. The responsibility for this publication lies with the authors.

References

1. Seifert, S., Barbu, A., Zhou, S., Liu, D., Feulner, J., Huber, M., Suehling, M., Cavallaro, A., Comaniciu, D.: Hierarchical parsing and semantic navigation of full body ct data. In: SPIE Medical Imaging. (2009)
2. D. Channin, P. Mongkolwat, V.K.K.S.D.R.: The cabig annotation and image markup project. In: Journal of Digital Imaging, online (2009)
3. Rubin, D., Mongkolwat, P., Kleper, V., Supekar, K., Channin, D.: Medical imaging on the semantic web: Annotation and image markup. In: AAAI Spring Symposium Series, Semantic Scientific Knowledge Integration, Stanford, USA (2008)

4. Möller, M., Regel, S., Sintek, M.: Radsem: Semantic annotation and retrieval for medical images. In: of The 6th Annual European Semantic Web Conference (ESWC). (2009)
5. Wittekind, C., Meyer, H., Bootz, F.: TNM Klassifikation maligner Tumoren. Springer Medizin Verlag, Heidelberg (2005)
6. Rosse, C., Mejino, J.J.: A reference ontology for biomedical informatics: the foundational model of anatomy. J. of Biomedical Informatics **36** (2003)
7. Golbreich, C., Zhang, S., Bodenreider, O.: The foundational model of anatomy in owl: Experience and perspectives. Web Semantics: Science, Services and Agents on the World Wide Web **4** (2006)
8. Noy, N., Rubin, D.: Translating the foundational model of anatomy into owl. Web Semantics: Science, Services and Agents on the World Wide Web **6** (2008)
9. Uschold, M., Gruninger, M.: Ontologies: Principles, methods and applications. Knowledge Engineering Review **11** (1996)
10. Wennerberg, P., Buitelaar, P., Zillner, S.: Towards a human anatomy data set for query pattern mining based on wikipedia and domain semantic resources. In: Proceedings of a Workshop on Building and Evaluating Resources for Biomedical Text Mining (LREC), Marrakech, Marocco (2008)
11. Buitelaar, P., Wennerberg, P., Zillner, S.: Statistical term profiling for query pattern mining. In: Proceedings of ACL 2008 BioNLP Workshop, Columbus, Ohio, USA (2008)
12. Wennerberg, P., Zillner, S.: Towards context driven modularization of large biomedical ontologies. In: Proceedings of the International Conference of Biomedical Ontology (ICBO), Buffalo, New York, US (2009)
13. Dameron, O., Roques, E., Rubin, D., Marquet, G., Burgun, A.: Grading lung tumors using owl-dl based reasoning. In: Proceedings of 9th International Protégé Conference. (2006)
14. Marquet, G., Dameron, O., Saikali, S., Mosser, J., Burgun, A.: Grading glioma tumors using owl-dl and nci-thesaurus. In: Proceedings of the American Medical Informatics Association Conference AMIA. (2007)
15. Golbreich, C., Dameron, O., Bierlaire, O., Gibaud, B.: What reasoning support for ontology and rules? the brain anatomy case study. In: Proceedings of the Workshop on OWL Experiences and Directions, Irlande (2005)
16. Zillner, S., Hauer, T., Rogulin, D., Tsymbal, A., Huber, M., Solomonides, T.: Semantic visualization of patient information. In: Proceedings of the 21th IEEE International Symposium on Computer-Based Medical Systems (CBMS), Jyvskyl, Finland (2008)