

# Knowledge Structuring and Retrieval for Intracranial Aneurysm Research

Susanne HANSER <sup>a,1</sup>, Juliane FLUCK <sup>b</sup>, Laura Inés FURLONG <sup>c</sup>,  
Christoph M. FRIEDRICH <sup>b</sup>, Martin HOFMANN-APITIUS <sup>b</sup>,  
Holger STENZHORN <sup>a</sup> and Martin BOEKER <sup>a</sup>

<sup>a</sup> *Dept. of Medical Informatics, University Medical Center Freiburg, Germany*

<sup>b</sup> *Fraunhofer Institute SCAI, St. Augustin, Germany*

<sup>c</sup> *Research Unit on Biomedical Informatics (GRIB) IMIM/UPF, Barcelona, Spain*

**Abstract.** The main goal of the European project @neurIST is to link and integrate highly heterogeneous, distributed data and knowledge from various scientific disciplines to improve information access for personalized disease management and scientific discovery. To this end, a major activity of the project is centered around the development of a disease-specific ontology for exhaustively representing all entities pertaining to the management and research of intracranial aneurysms. The ontology is subsequently used for various tasks, such as terminological control, text annotation, as well as data mediation. This paper describes the background of the ontology development and further focuses on its actual usage in the context of text mining applications to retrieve biomolecular information related to intracranial aneurysms.

**Keywords.** Knowledge Retrieval; Ontology Design; Intracranial Aneurysm; Terminology

## Introduction

The European project @neurIST<sup>2</sup> (Integrated Biomedical Informatics for the Management of Cerebral Aneurysms) aims at developing an integrated IT infrastructure for the management of data from patients with an intracranial aneurysm. The ultimate aim of the project is a system that supports individualized healthcare and in particular allows for personalized risk assessment and a subsequent recommendation for treatment.

The project brings together partners from various domains including computer scientists, biomedical informaticians, clinicians and experts in biomolecular research. Due to the broad coverage of the involved scientific disciplines, we are dealing with highly heterogeneous data:

- Clinical data which encompass electronic patient records and data gathered via a dedicated clinical reference information model (CRIM)
- Epidemiological data originating from regional data collections including databases of general practitioner as well as of clinical studies

---

<sup>1</sup> Corresponding Author: Susanne Hanser, Dept. of Medical Informatics, University Medical Center Freiburg, Stefan-Meier-Str. 26, 79104 Freiburg, Germany; E-Mail: susanne.hanser@uniklinik-freiburg.de

<sup>2</sup> @neurIST website: <http://www.aneurist.org>

- Experimental data from molecular biology experiments done with patient samples including microarray data and genetic association studies
- Clinical imaging data which are processed for hemodynamic simulation
- Derived data such as a shape index that has been computed from processed image data (this index describes the topology of an intracranial aneurysm)
- Representations of literature mining approaches by experts (e.g., Cochrane-type of expert knowledge) or by machines (text mining results)

These data span over a wide range of various granularity types, e.g., from the molecular level over the cell and organ level up to the population level on a spatial scale and from bio-molecular conversions to evolutionary processes at a temporal scale. Data syntax and structures are highly heterogeneous: information is presented via text, images and other symbolic structures. They are also diverse in context and pragmatics, from clinical guidelines based on broad epidemiological studies, and knowledge gained from disease-specific research studies, up to patient-specific data in electronic health records.

As this diversity leads to a tremendous difficulty to generate and to maintain, the overview on the knowledge already present, one primary theme of @neurIST is to develop an horizontal data integration at every level of abstraction, from access to information sources, to complex information processing, knowledge representation, structuring and fusion of data. For this purpose we undertook a survey of all relevant data terminology and knowledge sources used in the consortium (e.g., data base schemata, classifications used, disease models) and generated an overall ontology covering and structuring the conceptual space relevant for intracranial aneurysm. In the following, we will give examples for how the ontology can support the data integration.

For the representation of the public knowledge, primarily in the biomolecular field, scientific publications are the most comprehensive source for knowledge; e.g., for the MeSH term “Intracranial Aneurysm” approximately 16,000 articles and 1,400 reviews are listed in the Medline database with this number steadily growing over time. Extending the search by diseases known to be associated with intracranial aneurysms as risk factors (e.g., polycystic kidney disease) or biological processes having an impact for the risk of the formation of cerebral aneurysm (e.g., mechanical stability of the vessel) will even multiply the number of relevant publications. Text mining systems can be used in this task to automatically extract both relevant disease-specific information as well as the associated molecular entities related to aneurysms and further provide a condensed overview about the disease area.

In the following, we describe our joint work of ontology-based knowledge retrieval in the context of the disease intracranial aneurysm. The aneurysm-specific terminology integrated in the @neurIST ontology is used for disease retrieval of information. Additional recognition of other entities like genes and proteins and their sequence variants (SNPs) allow the extraction of relationships between aneurysm-related concepts and molecular entities. The retrieved information was subsequently structured and categorized in the web portal @neuLink. Some examples for the retrieval of information relevant to researchers and clinicians are shown in the following as well.

## 1. System Overview

### 1.1. The @neurIST Ontology

An ontology describes a specific, well-defined domain, introducing the proper definitions and descriptions for all relevant entity types (also called concepts) and relations. The entity types are arranged in a taxonomy and associated via relations by applying formal axiomatic definitions. Those axioms constrain the interpretation and well-formed use of the entity types. In addition, human-understandable textual definitions are added to describe their respective meaning in plain language.

#### 1.1.1. Requirements and Use Cases

The development of any domain ontology should be driven by a comprehensive collection of both requirements and use cases towards the ontology:

- Primarily, the ontology is intended to serve as a terminology in all parts of the project. It identifies the allowed terms through its associated vocabulary and assigns their respective meanings in the clinical and experimental documentation as well as in the development of databases and graphical user interfaces.
- An important activity in the @neurIST project is knowledge discovery. Therefore, our ontology provides several term lists (dictionaries) which can be employed in text mining and annotation. Semantic analyses are supported through the given relations between the entities.
- Finally, the ontology is intended to support data integration. One of the most ambitious objectives of the @neurIST project is the integration of heterogeneous data sources through semantic mediation.

#### 1.1.2. Design Principles

The @neurIST ontology integrates different scientific disciplines and views on cerebral aneurysms and subarachnoid hemorrhage. Our sources in the ongoing acquisition of entity types, relevant terms and relations in the domain are

- clinical databases and a clinical reference information model (CRIM),
- literature (i.e., high frequency domain specific terms extracted from literature),
- direct communication with domain experts,
- the UMLS Metathesaurus , and
- open-access databases (e.g., in the molecular biology field).

The @neurIST ontology is based on and adapted to several existing ontologies that are commonly accepted as standard in the field. For example, the choice of an appropriate upper ontology is of crucial decision in the development of any more complex ontology. In contrast to “lightweight” ontologies which focus on a minimal terminological structure (i.e., often just a pure taxonomy) fitting the needs of a specific and mostly smaller community, the main purpose of “foundational” ontologies is to *negotiate* meaning on a larger scale, either to enable effective cooperation among multiple artificial agents or to *establish consensus* in a mixed society where such artificial agents need to cooperate with human beings [1]. All derived entity types of the @neurIST ontology were categorized according to the *Descriptive Ontology for Linguistic and Cognitive Engineering* upper ontology (DOLCE [2] and are therefore subclasses of one of its basic entities: *endurant* (independent essential wholes, e.g., object-

and substance like entities), perdurant (events, processes, activities, and states), quality (entities which can be perceived or measured, e.g., color, length) and region (abstract entities, i.e., spatial, temporal and abstract regions).

For the representation of anatomical entities of interest, the respective parts of the *Foundational Model of Anatomy (FMA)* [3] were included in our ontology as well and (re-)classified along the lines of the DOLCE top-level concepts. Information of several other biomedical terminological resources and databases are either directly linked into the @neurIST ontology (i.e., via OWL imports) or parts of them were manually incorporated. The *UMLS Metathesaurus* provides mainly taxonomic information about concepts, as well as synonyms and definitions for a large number of entities [4]. The entities of the ontology were mapped to UMLS Concept Identifiers (CUIs) whenever this was feasible using MetaMapTransfer MMTX<sup>3</sup>. The mapping to UMLS CUIs provided a preliminary classification of entity types according to the top-level categories of the UMLS Semantic Network as well as a mapping to existing biomedical vocabularies containing adequate entities. Biomolecular entities are directly linked to SwissProt IDs, Entrez-Gene IDs and Gene Ontology IDs.

### 1.1.3. Ontology Implementation

As technical framework for the ontology development, the Protégé ontology editor<sup>4</sup> was utilized to simplify OWL ontology development. The @neurIST ontology is fully conformant with the OWL-DL specification and therefore we could use the reasoner Pellet<sup>5</sup> to continuously check the consistency of the ontology and classify it.

For the generation and especially for the adaptation of the ontology and the dictionaries containing the relevant medical terms describing the disease intracranial aneurysm, it was necessary to attune the different approaches to biomolecular research and medical terminology. An ontology intends to represent the “world” (the reality) or parts of it, respectively. The emphasis of this kind of terminology therefore lies on a correct and *consistent* presentation of the entity types relevant in a field, while for text mining collections of words (strings) used to communicate over (part of) the world are important. The @neurIST ontology incorporates a subtree to the ontology which extracts, reclassifies and displays the types of interest from the ontology as such to user required conceptual models which are knowledge or context dependent (Figure 1).

Giving both textual and formal (i.e., description logic) definitions, the current version of the @neurIST ontology specifically represents the entities describing patients, treatment and molecular etiology of aneurysms and subarachnoid hemorrhage, it also integrates the genes and proteins related to intracranial aneurysms found in literature. To provide an adequate terminological coverage, the ontology is linked to a separate lexical resource which is not part of the ontology proper.

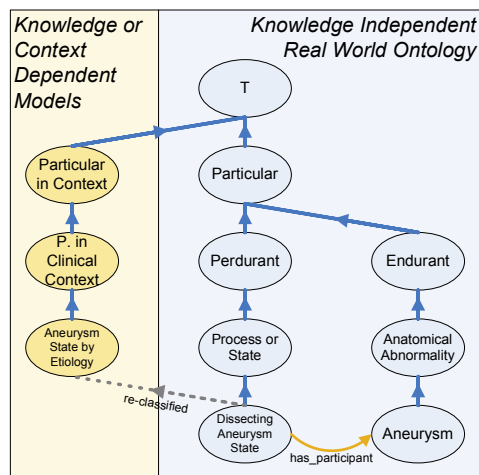
A web-based ontology-browser was developed to provide easy access to the content; the hierarchical and lexical information displayed is accessible through a link from the @neuLink Suite, a similar link from the project’s patient data entry tool @neuQUEST is planned.

---

<sup>3</sup> MetaMap Transfer website: <http://mmtx.nlm.nih.gov>

<sup>4</sup> Protégé website: <http://protege.stanford.edu>

<sup>5</sup> Pellet website: <http://pellet.owldl.com>



**Figure 1.** The @neurIST ontology not only consists of a knowledge independent real world ontology (the “ontology as such”) but allows the inclusion of models which are dependent on knowledge or context as required by the @neurIST project. Entity types are assigned strictly ontological as subtypes of the DOLCE upper level ontology (right side). The association to subtypes of “Particular in Context” is a secondary re-classification by a description logic reasoner according to the restrictions of the types.

## 1.2. Recognition of Different Biological Relevant Entities in Scientific Text and their Semantic Integration to Corresponding Data Resources

For the recognition of aneurysm-relevant entities in texts, like disease or anatomic terms but also genes, proteins or drugs, the named entity recognition tool ProMiner [5, 6] is used. ProMiner is already established for the recognition of gene and protein names and their corresponding sequence database entry (e.g., “human BCL2” in text should correctly be associated to “GeneID:596” in Entrez Gene). The software performed very well on two independent international assessments for this task (Biocreative I [7] and II [6])<sup>6</sup> and reached top ranking positions in both competitions for gene normalisation. In ProMiner, a gene and protein dictionary extracted from the sequence databases EntrezGene and UniProt is used and extensive curation strategies have been developed for the generation of its dictionaries, sophisticated rules for disambiguation of ambiguous names and adaptive strategies for the handling of acronyms.

Here we use the ProMiner technology to identify not only gene and protein names but also the aneurysm relevant medical entities gathered in the @neurIST ontology, drugs or chromosomal locations in a text corpus specific for the knowledge area of intracranial (cerebral) aneurysms. The advantage of this dictionary based approach is that for named biomedical entities that are detected in text the association to database entries is given. For the recognition of chemical active molecules we used chemical dictionaries with compound names compiled from reference names and synonyms in DrugBank [8]. As with biomedical entities, we can use this approach to link chemical entities in text to their respective entries in DrugBank. The medical entities link directly to the ontology and allow direct integration with other data within the aneurysm context. In such a way the extracted information can be directly combined with available database information or experimental results. The latter is of particular importance in the case that experimental data such as microarray data or proteomics data should be used for the evaluation of quantitative models in the context of aneurysm.

<sup>6</sup> [http://biocreative.sourceforge.net/biocreative\\_2.html](http://biocreative.sourceforge.net/biocreative_2.html)

In addition to anatomical, drug, gene and protein terms, the recognition of allelic variants of genes from text is a topic of interest to the @neurIST project. Gene allelic variants such as Single Nucleotide Polymorphisms (SNPs) may underlie the genetic aetiology of the disease or modulate the action of environmental risk factors, as evidenced by a plethora of genetic association studies carried out in the last years. OSIRIS was applied for the specific aim of recognition of gene and protein sequence variants in the biomedical literature. Performance evaluation showed satisfactory results on both precision (99 %) and recall (89 %) for the recognition of variation terms and disambiguation to dbSNP identifiers [9].

Similarly to ProMiner, OSIRIS uses a dictionary based approach to find terms representing sequence variants in text. The dictionary is compiled from biological databases collecting sequence variant information (dbSNP from the NCBI [10]) following manually devised rules developed after inspection of literature and nomenclature guidelines. The use of OSIRIS allows the linkage of a textual representation of a SNP with its sequence database identifier from dbSNP. This linkage, as already mentioned for genes and proteins, provides relevant contextual information to the entities identified in the text. For instance, mapping a variation entity with a dbSNP identifier allows its unambiguous identification, and in consequence all the information available for this variation can be gathered: organism, location in the genome, validation status, populations in which the variant has been sequenced, biological sequences where the variant has been mapped, and so forth.

### *1.3. The @neuLink Application Suite*

The @neuLink application suite is one of four visible interfaces to the users within the @neurIST IT infrastructure. It aims at linking genetics to disease and Knowledge Discovery. The target audience of @neuLink are genetics-oriented researchers, but the final suite will also address questions of interest to general biomedical researchers. Complex diseases like intracranial aneurysms are affected by heritable effects, life style and environmental risk factors. One of the targets of the @neuLink application suite is therefore to support the genetics oriented researcher in finding disease or phenotype associated genes, proteins and SNPs. There are two access scenarios that are supported, a web based application interface for humans (cf. figure 2) and a grid based data access service for other modules of the @neurIST infrastructure. Most of the information on associations between diseases and genes is stored as unstructured text. The data source for the text-mining applications described in section 1.2 is the Medline repository with information on approximately 18 Mio. documents at the moment. The data repository for the @neuLink suite consists of a sub-corpus with approximately 2,000,000 documents from the disease domain and the corresponding information on extracted entities (cf. section 1.2). The interface allows full text and semantic search and provides the user with a visualization of aggregated search results.

## **2. Retrieval of Aneurysm-Related Public Knowledge**

For the retrieval of intracranial aneurysm-related public knowledge, we searched in the Medline abstract database. In the presented examples of retrieved results we selected a corpus using the search string “intracranial AND aneurysm”. For searching all aneurysm-related information in this corpus we used the @neurIST ontology associated

terminology together with other dictionaries like gene and protein names or drug names in ProMiner. The retrieved gene and protein information is used in addition to find SNP-related information in the texts with the OSIRIS software. The results together with statistical information, selection and visualisation possibilities were included in the @neuLink front end. Here we will give some examples how the @neurIST terminology together with other recognition of relevant entities classes like genes and proteins or SNP information can be used for disease specific retrieval.

### *2.1. Retrieval Based on the @neurIST Ontology*

In the ontology tree ‘Risk Factor for Intracranial Aneurysm State’ all known proven and suspected risk factors are listed (cf. figure 2, left part). This selection of proven or suspected risk factors in the ontology was based on a first epidemiologic review of literature [11, 12, 13]. In the result set all 31 found entries of the selected subtree are shown (cf. figure 2, right part). Most frequent hits are the proven risk factors ‘Arteriovenous Malformation State’, ‘Polycystic Kidney Autosomal Dominant’ and ‘Moyamoya’ (cf the first two text examples at the right bottom of Figure 2).

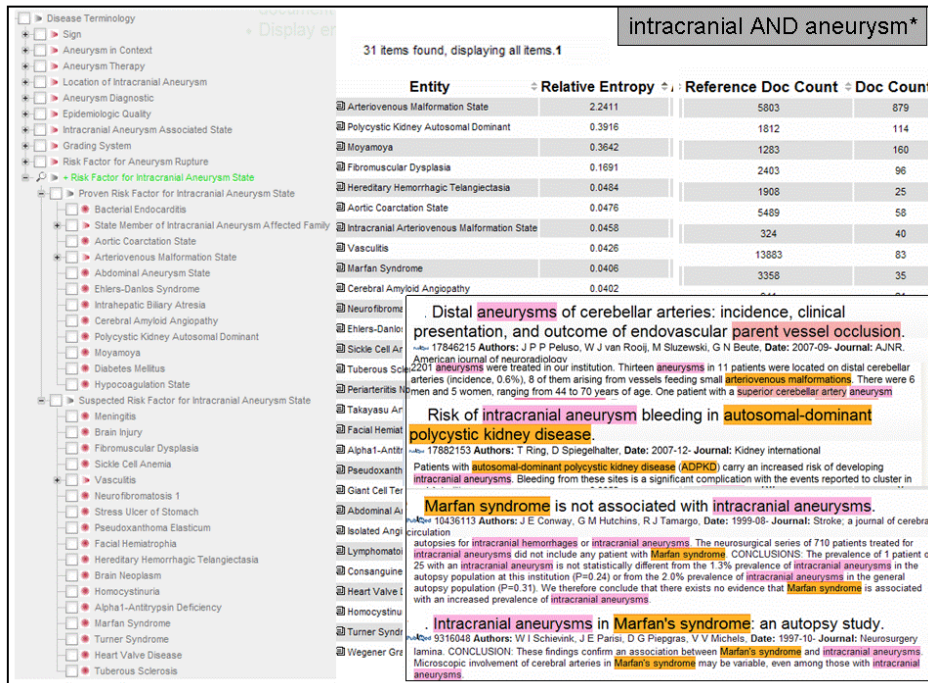
But we find also suspected risk factors like ‘Fibromuscular dysplasia’, ‘Hereditary haemorrhagic telangiectasia’ or ‘Marfan syndrome’ among the most frequent ones. For ‘Fibromuscular dysplasia’ a high number of case reports show intracranial aneurysms in these patients, the ‘Hereditary haemorrhagic telangiectasia’ causes arteriovenous malformations and thus are a suspected risk factor for intracranial aneurysm. For the ‘Marfan Syndrome’, which is a heritable disorder of connective tissue, we find contradictory publications (cf. Figure 2, bottom right). In such a way, @neuLink enables a fast overview over the published work and can serve as first retrieval step for epidemiologic studies to select relevant publications for further reading of the full texts.

In a second example, we retrieve information about different aneurysm therapies and their relation to different grading systems used for describing the disease. We selected the ontology category ‘Aneurysm Therapy’ and the category ‘Grading system’ and, by doing so, find publications talking specifically about the outcome of different therapies (e.g., ‘Glasgow Coma Score Region’, ‘Modified ranking scale’) or the success of a certain therapy within patient having a certain grade of the disease (e.g., ‘Hunt and Hess Grade’). With the selection of these recognized disease related terms it is easily possible to get a quick overview on the success and outcome of different therapies.

### *2.2. Retrieval Based on Additionally-used Dictionaries and Retrieval Methods*

In addition to the dictionaries derived from the ontology various other dictionaries covering genes and proteins, SNPs and drugs are used to find relevant aneurysm entities. No preselection is done at this point on relevant aneurysm genes or proteins but all genes and protein are searched for. 364 different genes/proteins are in 676 documents (abstracts including whole Medline till 12/2007) are shown in the result set.

For evaluation purposes the review on genetics [14] was also considered. They mention 18 associated genes in the context of intracranial aneurysms. In our evaluation [15] we found all these genes and ranked them to the first 238 hits with 7 hits among the top 16 candidates. High scoring genes not associated to genetics are for example PLAT (tissue plasminogen activator) which is often used in aneurysm therapy. Here the additional selection of the MeSH term ‘genetics’ reduced the amount of genes to 119.



**Figure 2.** Selection of all Medline abstracts retrieved by a string search for ‘intracranial AND aneurysm’ and recognition of all entities contained under the category ‘Risk Factor for Intracranial Aneurysm State’. On the left hand side all concepts belonging to ‘Risk Factor for Intracranial Aneurysm State’ are shown, on the right hand side all found entities with the accompanied relative entropy and the frequency of found documents in whole Medline (‘Reference Doc Count’) and the selected corpus (‘Doc Count’) is depicted. On the right bottom, some example text snippets are shown.

Now, under the 15 top ranking genes selected by relative entropy, twelve genes are described by Kricshek and Inoue [14]. The others are PKD1, APOE and PKD2 which are clearly also suspected to be directly associated with aneurysm. In comparison to reviews about genetics in intracranial aneurysms the text mining approach presented here is always up to date and provides a superset of genes involved in different aspects of the disease.

A further selection of relevant information can be done through the identification of SNP-related information. This resulted in the identification of SNP mentionings for five genes. Using this approach, SNPs mentioned in the text are identified and normalized to dbSNP identifiers with high precision. However, it was found that inconsistencies between naming conventions used in the literature and dbSNP (e.g., the eNOS SNP T-786C in the text example) impaired the normalization of SNP mentionings.

### 3. Data Integration Based on the @neurIST Ontology

The ontology represents relevant types and their relations in the domain of interest and can be used as a knowledge base providing textual and formal definitions of classes, which are linked to standard terminologies via UMLS identifiers. In combination with a lexicon, the ontology also provides services for synonym control and cross-lingual

access to classes. An ontology enables interoperability between data and applications by providing a common semantically-defined resource. The meaning of data and the operations on them are specified by linking data fields and values to ontology classes. In the @neurIST project, the linking of data to the ontology has been already achieved by mapping the CRIM attributes to the UMLS and consequently to the ontology.

The ontology defines all types (classes) and the relations between them explicitly in a formal logic providing computational usability in visual or conceptual navigation applications. The formalization through explicit semantics is also of major value in the design and development of databases, object models as well as whole architectures.

One of the objectives of the @neurIST project is the inclusion of distributed computing on a large scale. Major roles of the ontology in this scenario are semantic interoperability and data mediation which can be provided by mediating between heterogeneous databases on a semantic level.

#### **4. Summary and Future Prospects**

The @neurIST ontology models the conceptual space and current knowledge relevant for the research and management of intracranial aneurysms. Moreover, it fulfills all major requirements of the project partners as has been shown by a standardized questionnaire. These facts made it possible to successfully deploy the ontology in various applications and facilitate their respective development. The example text mining application allows the actual end-user (i.e., clinicians and researchers) to gain simplified access to current knowledge on intracranial aneurysms by employing dictionaries based on the ontology which is further enhanced by the visualization capabilities of the @neuLINK suite.

It has shown to be very beneficial both for the ontology developers as well as its users, that a set of principles for “good ontology design” were followed: The ontology relies on standard technology and terminology systems, and so developers can readily relate the ontology content to familiar tools and environments. The DOLCE upper level ontology provided an extremely valuable framework in that it made available a rigid and well-defined formal background on how to create and arrange entity types within the ontology in a coherent and correct fashion. To meet the user requirements, a separate “particular in context” branch of the ontology was added to represent knowledge or context dependent types (e.g., risk factors). To allow for a clean development and large scale of applications the entity types within the ontology linked to several well maintained separated lexicons with good coverage.

But despite all the given benefits of the @neurIST ontology, one problem is its tremendous size and intricate structure: The ontology covers a very broad scope of sub-topics within the domain and hence contains such a large amount of content that it is difficult to fully grasp and handle. This situation is further complicated by the fact that the domain experts are usually not experts in the logical formalization of ontologies - and current tools do not provide easy-to-use interfaces for such domain experts. Therefore, one major goal is to better communicate the structure, usability and the specific targets of the ontology to the actual users to integrate them even more in the ongoing ontology development. This development, including the definition of more entity types and relations, will lead to enhanced applicability of the ontology to specific user needs.

In conclusion, the @neurIST ontology provides a fully formalized resource which covers a broad range of types (concepts) in the domain of intracranial aneurysm from

molecular biology over clinical medicine and flow dynamics to epidemiology. The ontology design had both to satisfy the principles of good ontological practice and a large scale of user requirements. The @neurIST ontology has now been successfully integrated in several application contexts throughout the complete project.

## Acknowledgments

This work has been funded in the framework of the European integrated project @neurIST, which is co-financed by the European Commission through the contract no. IST-027703. In addition, we like to acknowledge the excellent work of our co-workers: Roman Klinger, Heinz-Theodor Mevissen, Holger Dach and Tobias Gattermayer. Special thanks go to Stefan Schulz for many fruitful discussions on ontology design.

## References

- [1] Grenon P. BFO in a Nutshell: A Bicategorical Axiomatization of BFO and Comparison to DOLCE. Grenon, Poerre IFOMIS Reports 2003. Leipzig, IFOMIS, Universität Leipzig
- [2] Masolo C, Borgo S, Gangemi A, Guarino N, and Oltramari A. WonderWeb Deliverable D18. Ontology Library (final). 2003
- [3] Rosse C., Mejino JLV. A reference ontology for biomedical informatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics* 2003;36:478-500
- [4] Bodenreider O. The Unified Medical Language System (UMLS): Integrating biomedical terminology. *Nucleic Acids Research*. 2004;32:D267-D270
- [5] Hanisch D, Fundel K, Mevissen HT, Zimmer R and Fluck J. ProMiner: Rule based protein and gene entity recognition. 2005. *BMC Bioinformatics* 2005, 6(Suppl 1):S14
- [6] Fluck, J., Mevissen T.H., Dach, H., Oster M., Hofmann-Apitius M. ProMiner: Recognition of Human Gene and Protein Names using regularly updated Dictionaries. Second BioCreAtIvE Challenge Workshop 2007, Critical Assessment of Information Extraction in Molecular Biology, Madrid Spain.
- [7] Hirschman L, Yeh A, Blaschke C, Valencia An Overview of BioCreAtIvE: critical assessment of information extraction for biology *BMC Bioinformatics* 2005, 6(Suppl 1):S1 (24 May 2005)
- [8] Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res*. 2006, 34(Database issue):D668-72
- [9] Furlong LI, Dach H, Hofmann-Apitius M, Sanz F. OSIRISv1.2: a named entity recognition system for sequence variants of genes in biomedical literature. *BMC Bioinformatics* 2008, 9:84.
- [10] Sherry ST, Ward M, Sirotkin K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res* 1999, 9(8):677-679.
- [11] Rinkel GJ, Djibuti M, Algra A, et al: Prevalence and risk of rupture of intracranial aneurysms: a systematic review. *Stroke* 29:251-256, 1998
- [12] International Subarachnoid Aneurysm Trial (ISAT) Collaborative Group. International Subarachnoid Aneurysm Trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptures intracranial aneurysms: a randomized trial. *The Lancet* 2002;360:1264-74
- [13] The International Study of Unruptured Intracranial Aneurysms Investigators: Unruptured Intracranial Aneurysms – Risks of Rupture and Risks of Surgical Intervention. *New England Journal of Medicine*. 1998;339:1725-33
- [14] Krischek, B. and Inoue, I. (2006). The genetics of intracranial aneurysms. *J Hum Genet*, 51(7):587–594.
- [15] Gattermayer, T. (2007). SCAIView: annotation and visualization system for knowledge discovery. Master's thesis, Life Science Informatics at Bonn-Aachen International Center for Information Technology (B-IT); Germany.