



## **SemanticMining**

*NoE 507505*

### **Semantic Interoperability and Data Mining in Biomedicine**

## **Dissemination Activities 2006**

**D7.4 Joint European Summer School**

**D9.4 Dissemination of results**

**D40 Training program in biomedical ontology**

**D41 Terminology conferences**

Report Version: 1

Report Preparation Date: 2006-12-20

Classification: RE

Contract Start Date: 2004.01.01    Duration: 3.5 years

**Project funded by the European Community under the FP6 Programme “Integrating and Strengthening the European Research Area” (2002-2006)**

# Table of Content

<b>FOREWORD</b>	<b>4</b>
<b>JOINT EUROPEAN SUMMER SCHOOL IN BIOMEDICAL INFORMATICS</b>	<b>5</b>
Administrative information	5
Objectives	5
Results in numbers	5
Overview of the program	6
Doctoral consortium	8
Tutorials and workshops	10
Scientific presentations in the “best of” session	10
NoE cluster meeting	11
Assessment and Conclusion	11
Dissemination of Information	12
<b>INTERNATIONAL SYMPOSIUM ON SEMANTIC MINING IN BIOMEDICINE (SMBM)</b>	<b>13</b>
Administrative information	13
Objectives	13
Preparation of the SMBM (October 2005 – April 2006)	13
Realization of the SMBM (April 9th – April 12th 2006)	16
Revision of the SMBM (April – August 2006)	16
Statistics	16
<b>TRAINING COURSE ON BIOMEDICAL ONTOLOGY</b>	<b>17</b>
Administrative information	17
Summary	17
Introduction	17
Description of Work	18
Result	18
Assessment and Conclusion	21
List of Participants	22

<b>TERMINOLOGY CONFERENCE, SWEDEN</b>	<b>23</b>
Administrative information	23
Objectives	23
Conference program	23
Assessment and Conclusion	24
<b>SEMANTICMINING CONFERENCE ON SNOMED CT (SMCS)</b>	<b>25</b>
Administrative information	25
Objectives	25
Preparation	26
Conference	26
Follow-up	27
Assessment and Conclusion	27
Appendix I - Participant Countries	28
Appendix II - Full Program	29
<b>WORKSHOP ON FOUNDATIONS OF CLINICAL TERMINOLOGIES AND CLASSIFICATIONS</b>	<b>33</b>
Administrative information	33
Objectives	33
Preparation	33
Workshop	34
Follow-up	35
Assessment and Conclusions	35
<b>EXPLOITATION AND DISSEMINATION</b>	<b>36</b>
Administrative information	36
Dissemination through the web	36
Exploitation	36

## Foreword

This report summarises major outreach activities and events during 2006 within SemanticMining. Apart from these specific dissemination activities, a stream of high-impact conference presentations and scientific publications is coming out of the network and will continue beyond the funded period of this project.

# Joint European Summer School in Biomedical Informatics

## Administrative information

Lead contractor: National Institute for Strategic Health Research

Assisting partners: The Board

Participants: All partners

Author of report: György Surján (Summer School)

## Objectives

In 2006, the Third Semantic Mining Summer School was organised in cooperation with two other NoEs, INFOBIOMED and BIOPATTERN under the name "**European Summer School in Biomedical Informatics**"

<i>Objectives</i>	<i>Progress towards achieving objectives</i>
To plan and organise summer school 2006 together with two other IST NoE projects running in the field of Biomedical informatics (BIOPATTERN and INFOBIOMED) where researchers and doctoral students in either of the three networks gather in one place for an extended period (a full week) for a program of workshops, tutorials, demonstrations and social activities.	<ol style="list-style-type: none"><li>1. Planning including design of program in cooperation with INFOBIOMED and BIOPATTERN</li><li>2. Setting up the summer-school website</li><li>3. Realisation by partner ESKI</li><li>4. Running the summer-school: doctoral consortium, text and web mining WS, Data visualisation WS, Ontology engineering and terminology WS, integrative biomedical informatics supporting drug discovery WS, "best of" session</li></ol>

## Results in numbers

Number of participants according to NoE-s and countries.

Country	Semantic Mining	BIOPATTERN	INFOBIOMED	Grant	Total
Belgium		1			1
Denmark	7				7
Finland	4				4
France	3				3
Greece		1			1
The Netherlands	1		2		3
Israel				5	5
Hungary	2				2

Germany	12			2	14
Italy		2			2
Portugal			1		1
Spain			9		9
Switzerland	2				2
Sveden	17	1	1		19
UK	11	7	1		19
<b>Total</b>	<b>59</b>	<b>12</b>	<b>14</b>	<b>7</b>	<b>92</b>

## Overview of the program

Date	Programme
<b>Sunday, 2 July 2006</b>	Travel day (Meet in Budapest, bus Transfer to Balatonfüred)
<b>Monday-Tuesday 3-4 July</b>	"Doctoral consortium" - A two days training program for PhD Students, where they present their work, and get feedback from senior researchers. With help of discussants they can learn how to focus on the essential scientific problems of their field and how to present their work. 37 students and 8 researchers (for leading discussion groups, commenting on student presentations, etc.).
<b>- Wednesday 5-July</b>	<b>Text and web mining</b>
<b>8h30 - 9h15</b>	<b>: Introduction to biomedical text mining (U.Hahn)</b> - What is the benefit from text mining? - Categorisation of text mining approaches - Introduction to the evaluation of text mining results - Overview on resources to do text mining: corpora, training and test sets
<b>9h15 - 10h00</b>	<b>: Biological, chemical, medical, ontological named entities</b>  (D. Rebholz-Schuhmann) - Term variability in the scientific literature - Why is named entity recognition still a problem? - Dictionaries versus classifier based approaches - Integration of semantics in the text mining approach. 10h00 - 10h15 Break
<b>10h15 - 11h00:</b>	<b>Natural language processing (U.Hahn)</b> - Techniques to analyse natural language - What is required for a comprehensive solution? - Analysis of newswire text in comparison to biomedical literature - Which solutions are available now and in the near future
<b>11h00 - 12h00:</b>	<b>Knowledge discovery (E.v.Mulligan)</b>

<p><b>14h30 - 16h00:</b></p> <p><b>16h00 - 17h15:</b></p> <p><b>17h30 - 18h30:</b></p>	<ul style="list-style-type: none"> <li>- Documents represented in a "knowledge space" (fingerprints)</li> <li>- Inference of new findings from document representations</li> <li>- Use of knowledge discovery as part of the drug discovery process</li> </ul> <p><b>Data visualisation</b></p> <p><b>Introduction to Data Visualisation</b> (Ian T. Nabney)</p> <ul style="list-style-type: none"> <li>- Role of visualisation in data analysis - understanding data; finding clusters; detecting outliers; feature selection; interpreting</li> <li>- Visual data mining - information visualisation and data projection</li> <li>- Basic data projection methods - Principal Component Analysis (PCA); probabilistic PCA; Generative Topographic Mapping and Self-Organising Map; Neuroscale</li> <li>- Advanced data projection methods - hierarchical visualisation; automated structure generation; local predictive models; modelling uncertainty; time course data</li> </ul> <p><b>Practical Data Visualisation</b> (Ian T. Nabney, Dharmesh Maniyar)</p> <ul style="list-style-type: none"> <li>- Using a software tool to visualise datasets; training projection models (PCA, GTM and Neuroscale); interpreting projection results; interacting with data visualisation</li> </ul> <p style="text-align: center;">17h15 - 17h30: Break</p> <p><b>Bioinformatics Application of Data Visualisation</b> (Xose Fernandez)</p> <ul style="list-style-type: none"> <li>- Ensemble - Annotation of whole genomes of different species</li> <li>- Reactome - Navigation in Pathway representations for exploration and discovery</li> </ul>
<p><b>Thursday 6 July</b></p> <p>8.30-12.00</p> <p><b>15:00 – 15:30h.</b></p> <p>15:30 – 16:00h.</p> <p>16:00 – 16:30h.</p> <p>17:00 – 17:30h.</p>	<p><b>Ontology engineering, reference terminologies</b></p> <ol style="list-style-type: none"> <li>1. Reference Terminologies and formal ontological principles (Stefan Schulz, 45 min)</li> <li>2. Ontology / Terminology quality assurance and maintenance(Alan Rector, 45 min)</li> <li>3. SNOMED CT as a new generation of reference terminology, viewed under ontological and engineering aspects (Arne Kvaerneland, Ulrich Andersen, Ole Larsen)</li> </ol> <p><b>Points of view and experiences on integrative biomedical informatics supporting drug discovery and development.</b> (Round table session) Chaired by: Ferran Sanz</p> <p>Knowledge Management in the Innovative Medicines Initiative (IMI). The Research Unit on Biomedical Informatics (GRIB-IMIM/UPF) experience and future perspectives. Speaker: Ferran Sanz (Municipal Institute for Health Research) INFOBIOMED</p> <p>Speaker: Scott Boyer (Astra Zeneca) - INFOBIOMED Speaker: Johan van der Lei (Erasmus Medical Center) - INFOBIOMED</p> <p>16:30 – 17:00h. Coffee break</p> <p>Adverse drug reaction signaling and drug/disease annotation in Swiss-Prot. Speaker: Patrick Ruch (Geneve University Hospital) - Semantic Mining</p>

17:30 – 18:00h	Clustering and Visualisation for Drug Discovery. Speaker: Ian Nabney (Aston University) and Paulo Lisboa (Liverpool John Moores University) - Biopattern
<b>Friday 7 July from</b>	<b>“Best of” sessions 1 day, scientific presentations from all NoE-s</b> <i>This session gives an overview of the research programme of the three NoE projects through a selection of papers.</i>
<b>Saturday 8 July</b>	<b>NoE cluster meeting</b> <i>Lessons learned from the three NoEs. Strategic long-term discussions including FP7.</i> <b>NoE specific meeting, Assembly meetings</b>

## Doctoral consortium

The objective of the doctoral consortium is expressed through the presentation of the call for participation below. A separate call for discussants was also circulated through e-mail in the three networks.

We are planning a 2-day workshop for doctoral students, which will provide opportunities for exchanging their experiences with other students, getting feedback on their work from authorities in the field, and developing skills which are essential for gaining a PhD. The event will be a mix of group activities, presentations and discussions on research issues and research skills. All participants are invited to present a poster about their research work and some students will also have the opportunity to present their work in short seminars. All participants are expected to submit a one-page to two-pages research abstract by June 15<sup>th</sup>. The abstracts will be collected into a booklet to be distributed during the workshop.

To confirm their participation, students should e-mail [C.Mancini@open.ac.uk](mailto:C.Mancini@open.ac.uk) putting in the subject “Summer School: Doctoral Consortium” and attaching their research abstract.

### Registration for Summer School and funding

In order to participate to the Doctoral Consortium, students will \*ALSO\* have to register for the Summer School itself. The registration form can be downloaded from the website of the Summer School at <http://www.medinfo.hu/semi/joint/>. Students will be able to obtain funding for attendance from their Network of Excellence. For more detail they should contact the representatives of their own network, whose e-mails can be found on the website of the Summer School.

### Research abstract guidelines

By June 15<sup>th</sup>, participants should send a one-page to two-pages research abstract (file format: plain text) to [C.Mancini@open.ac.uk](mailto:C.Mancini@open.ac.uk). The research abstract should contain a clear statement of their research question, indicate where their research fits among other key work in their area, briefly explain their methods, and summarise any preliminary findings they may have. The title of the abstract should be descriptive of the topic. The student’s institution, supervisors and stage of the degree programme (e.g., first year in a three-year programme, nearing completion, etc.) should be specified underneath.

The outline of the program during the two day doctoral consortium is presented below.

<b>Monday 3<sup>rd</sup> July 2006</b>		
Time	Duration	Activity
9:30	45min	Welcome and Ice-Breaker
10:15	75min	Discussion 1: Research Questions
11:30	30min	Coffee
12:00	60min	Presentations:
12:00		Joel Pedriz Arrais, <i>Information Systems for microarrays and their integration in clinical environments</i> , p.8
12:20		Karin Friberg, <i>Decomposing Swedish medical compounds during query expansion and indexing in information retrieval</i> , p.10
12:40		Julie Niès, <i>A clinical decision support system devoted to prescription integrated to the HEGP information system</i> , p.16
13:00	60min	Lunch
14:00	60min	Presentations:
14:00		Erik Sundvall, <i>Patient Record Overviews and Navigation</i> , p.27
14:20		Lena Hansson, <i>Regions of Increased Gene Expression, RIDGES, in the mouse MHC loci upon infection with mCMV and/or priming with IFN-gamma</i> , p.14
14:40	60min	Cocktail Party Introduction
15:30		End
<b>Tuesday 4<sup>th</sup> July 2006</b>		
Time	Duration	Activity
9:30	60min	Discussion 2: Methods and Evidence
10:30	60min	Coffee and Poster Session
11:30	10min	Plenary Session on Posters
11:40	80min	Presentations:
11:40		Dima Stopel, <i>Brief introduction to artificial neural networks</i>
12:00		Erez Shalom, <i>An evaluation of a methodology for specification of clinical guidelines</i>
12:20		Avner Hatsek, <i>A graphical framework for specification of clinical guidelines</i>
12:40		Holger Stenzhorn, <i>Best ontologies design practices and their technical application</i>
13:00	60min	Lunch
14:00	60min	Discussion 3: Significant Contributions
15:00	60min	Up Against the Wall
16:00		End

The doctoral consortium was attended by 37 students and 8 researchers. A booklet containing the abstracts of the student's presentations was produced and is available on [www.semanticmining.org](http://www.semanticmining.org).

The doctoral consortium was evaluated through feedback forms. In total we received 28 feedback forms (25 on paper and 3 in electronic format). In general students found the workshop very useful. Some liked or disliked certain activities while others liked or disliked other activities, but in general it was evident from the forms that they appreciated having the opportunity of reflecting on the questions underlining their research and the methods they used. They appreciated learning from the experience of other students and getting advice from the discussants on how to tackle and present their work. Finally, they appreciated having an

opportunity for networking and exchange ideas. Both students and discussants provided useful observations on how we can improve the next doctoral consortium.

## Tutorials and workshops

The topics for the tutorials Text and Web Mining, Data Visualisation, and Ontology Engineering were chosen as they represent core knowledge areas of the networks. Experts from all three networks were involved in the preparation and presentations. The results of this part of the summer school were twofold; sharing of in-depth knowledge between participants, and insight into different application areas of the three networks.

## Scientific presentations in the “best of” session

Members of the three NoEs were invited to submit scientific presentations to the summer school. Selection of papers was done by a program committee of three experts. Prize for best papers were given to Ekaterina Buyko, Joachim Wermter, Michael Poprat. University of Jena, Germany and Julie Nièsac, Olivier Steichen, Christel Le Bozec. René Descartes University, INSERM Paris.

Ulf Schwarz, Mathias Brochhausen. IFOMIS, Univ Saarland, Germany	Ontology and layman’s interpretation of biomedical terminology
Anna Vikström, Gunnar Nilsson. Karolinska Institutet, Stockholm	SNOMED CT in multidisciplinary clinical practice - evaluation of usefulness for classification and coding of care planning
Erik Sundvall, Mattias Forss, Johan Hjalmarsson, Mikael Nyström, Håkan Petersson. Linköpings universitet, Sweden	Development of an Archetype Editor - A Tool for Modelling Structure in Electronic Health Records
Julie Nièsac, Olivier Steichen, Christel Lebozec. René Descartes University, INSERM Paris	Using Archetypes to Share Information between Computerized Clinical Decision Support Systems and Electronic Healthcare Records
Christel Daniel-Le Bozec, Olivier Steichen, Thierry Dart, Bruno Frandji, Marie-Christine Jaulent, Patrice Degoulet. INSERM, Paris Lunch	Terminological System for EHR: local dictionary of concepts and reference terminologies
Ekaterina Buyko, Joachim Wermter, Michael Poprat. University of Jena Julien Gobeill, Imad Tbahriti, Anne-Lise Veuthey, Patrick Ruch. Medical Informatics, Geneva University Hospital	Adapting an NLP Tool Suite to the Biology Domain: Evaluating OpenNLP Tools on the GENIA and PennBioIE Corpora Document Frequency Mixture for Effective Functional - Annotation in Swiss-Prot
Arrais, J.P., Oliveira, J.L., Campos, G.L., Carreto, L., Santos, M. Universidade de Aveiro, Portugal	Microarray data: from the hybridisation to the analysis Fetal reactivity bioprofiling
KG Rosen. Neoventor, Sweden Emmanuel Ifeachor. University of Plymouth, UK	Bioprofiling over Grid for e-healthcare

## NoE cluster meeting

The objective of the NoE cluster meeting was to share experiences of management, examples of good practise, and to discuss future initiatives from the biomedical informatics community in Europe. The following aspects were all addressed by the coordinators from the three NoEs.

Integration issues: indicators of "durable integration", forms of "integration" (virtual institute, other), self-sustainability after EC funding period, etc.

Contractual issues: Signature and entry into force, amendments, consortium agreement, collective responsibility, etc.

Management: burden of management tasks, autonomy/flexibility, distribution of budget among participants, conflict resolution, introducing new participants/competitive calls, gender, etc.

Dissemination/exploitation issues: spreading of excellence, links to industry/exploitation of results, IPR issues, etc.

Future approaches: Themes and orientation of FP7, possibilities for a joint NoE in the eHealth area, initiatives and proposals.

As a result of the discussions, the following areas were identified as a common priority list for the NoEs.

- Information sharing (deliverables, events etc.)
- Student mobility
- Sharing of tools and resources
- Event organisation and participation
- Sharing of best practice management, gender balance, collaborative research etc.
- New funding opportunities
- Collaborative opportunities
- Shared research activities

Discussions on future opportunities focused on the upcoming FP7 call its potential priority areas. The SYMBIOmatics priority list presented at the ICT for BIO-Medical Sciences Conference, June 29-30, 2006, were discussed, as well as different opportunities for joint applications in the form of NoEs, IPs or STREPs.

## Assessment and Conclusion

The evaluation of the summer school is based on panel discussions with participants from all three networks, and are summarised as follows.

Representatives of all involved NoEs agreed, that the organisation of the common summer school was an excellent occasion to share ideas and exchange experiences between senior researches and PhD students working in various fields of biomedical informatics.

The number of participants was lower than expected, but nearly optimal to manage the summer school in an efficient and practical way.

Exchange of ideas between various groups of experts working in different fields was satisfactory, as was the dissemination of knowledge between senior researchers and students.

The two day doctoral consortium was appreciated as an extremely successful and attractive part of the program.

As an appreciation of the success of the summer school, a similar event will be organised by Semantic Mining and INFOBIOMED and to be held in Barcelona, Spain in June 2007.

## **Dissemination of Information**

Presentations are available on MERMIG ([www.semanticmining.org](http://www.semanticmining.org)).

# International Symposium on Semantic Mining in Biomedicine (SMBM)

## Administrative information

**Lead contractor:** JENA (Friedrich-Schiller-Universität Jena)

**Responsible:** JENA

**Participants:** JENA, EMBL-EBI, UKLFR

**Deliverable:** D41

**Authors:** Michael Poprat, Udo Hahn (JENA)

## Objectives

Medicine and molecular biology rapidly develop into data-driven sciences. Therefore, electronic access to terminologies, ontologies, information resources such as databases, and computer tools is getting more and more crucial for successful applications like information retrieval, information extraction and text mining. One of the persistent goals of the SemanticMining NoE relates to the integration of knowledge-based terminological and ontological resources into access and retrieval methods for natural language documents and biomedical databases in order to support researchers and information providers in the biomedical domain.

Hence, several member institutions of the NoE (JENA, EMBL-EBI, UKLR) decided to launch an international symposium in order to address and discuss these topics at a scientific level. SMBM 2006 was the second event on this kind, carrying on the experience of the successful SMBM 2005, organized in April 2005 at EBI. The objective of SMBM 2006 was to enhance the networking between the following scientific and industrial communities: Molecular Biology, Bioinformatics, Chemoinformatics, Medical Informatics, Biological Text and Data Mining, Biological Knowledge Representation, Biological Ontology Design and Engineering, Human Language Technology and Computer Science, in general.

The scope of SMBM 2006 included the following topics applied to the domain of Biomedicine: Information extraction, information retrieval, text mining, knowledge discovery and data mining, term engineering, named entity recognition and interpretation, evaluation standards, ontological foundations of molecular biology and related areas, automated corpus/lexicon construction for Biomedicine. We offered four tutorials on Sunday, April 9. The scientific program started on Monday, April 10th and ended on Wednesday, April 12th. It included three keynote talks, panel discussions and an industry exhibition. An opening reception and a conference dinner took place on Monday and on Tuesday, respectively.

## Preparation of the SMBM (October 2005 – April 2006)

The work package proposal was formulated in October 2005 and accepted at the Assembly meeting in Manchester on December 6, 2006. With regards to the duration of the symposium, we followed the experience made for the first SMBM in 2005 (a 2.5-day event). Besides the general program (tutorials, scientific part, invited speakers) the following activities were carried out: installation of a university-wide online banking payment system for FSU, invitation letters to local authorities (mayor, deans, rector) as well as various political and industrial representatives, addressing sponsors from the industry and other institutions, planning of an industrial track, a panel discussion, a poster session, offerings of student grants. Social events included an informal get-together party in a student pub and the conference dinner.

**Program Chair and Committee.** The local organizing committee (LOC) consisted of Udo Hahn (head), Michael Poprat (assistant) and Anne Schneider (student helper), all from FSU-JENA. The program chairs (PC) were Sophia Ananiadou (National Centre for Text Mining, UK) and Juliane Fluck (Fraunhofer SCAI, Germany). Their program committee was composed of 18 highly esteemed international researchers (<http://supreme.coling.uni-jena.de/content/view/62/109/>). The PCs prepared a call for papers (maximum of 8 pages, deadline 23rd of January 2006) and a call for posters (maximum of 4 pages, deadline 18th of February 2006). Together with the local team, they broadcasted it to several newsgroups (BioNLP, corpora list, linguist list) and internal mailing lists (SMBM 2005, ECCB 2005, EMBL-EBI, SemanticMining) (first, second and final call). Moreover, hardcopy printed flyers were distributed at several meetings and conferences (SCAI text mining symposium (Oct. 2005), SemanticMining OWL workshop in Manchester (Dec. 2005), SemanticMining WP20 meeting in Geneva (Jan 2006), SemanticMining WP27 kick-off in Milton Keynes (Feb. 2006) by the local team, the program chairs and the other partners of this project.

**Tutorials.** The LOC and the PC put together a tutorial program which covered the most important topics of this symposium (text mining, ontologies, bioinformatics). We asked Dietrich Rebholz-Schuhmann (EBI), Alex Morgan (MITRE), Robert Stevens (U Manchester), Steffen Staab (U Koblenz-Landau) and Stefan Schuster (U Jena) to give a half-day tutorial. Dietrich Rebholz-Schuhmann and Alex Morgan delivered joint work with Jim Dong-Kim (U Tokyo) and Martin Krallinger (CNB/CNIO), respectively (see <http://supreme.coling.uni-jena.de/content/view/73/103> for the abstracts). The tutors were reimbursed for their travelling costs, their hotel costs (2 nights) and received the amount of 250 Euros for the tutorial.

**Invited Speakers.** Similar to the choice of tutorials, we tried to cover a wide range of issues relevant to semantic mining. Martin Romacker from *Novartis* talked about the role of text mining in the pharmaceutical industry while Lynette Hirschmann (MITRE) addressed the crucial role of evaluation effort in biomedical text mining. Unfortunately, the talk by Judith Blake (*Jackson Lab*) had to be cancelled. The invited speakers were reimbursed for their travelling costs, their hotel costs and were offered free participation of the workshops, the conference and the conference dinner.

**Panel Discussion.** The panel discussion was titled “Why do we need text mining centres for the Life Science domain and how could they support us?”. Representatives from international Text Mining centres and initiatives (Sophia Ananiadou, UK; Lynette Hirschman, MITRE; Udo Hahn, Germany) and the industry (Thérèse Vachon, Novartis) discussed this topic with the audience.

**Local Conference Organization.** The LOC built an official webpage (<http://supreme.coling.unijena.de/content/blogcategory/32/103/>) where all the information relating to the SMBM 2006 symposium was made accessible online. Together with the administrative branch of the Friedrich Schiller-Universität (FSU) in Jena, an online payment system was installed and became operational in February 2006. A registration form was created which reflected the terms and conditions for the participation as checked and translated by the legal department of the FSU. Additional infrastructure support could not be offered by the FSU. So it was up to the LOC to compose a list of selected accommodations (9 hotels, 3 guesthouses, 1 youth hostel), to choose a catering service (6 offers for catering services and the conference dinner) and to find out a suitable location (3 choices). After carefully balancing costs and quality, the LOC decided on the *Scala* tower restaurant which has an excellent reputation and offered an outstanding all-in-one solution (location, catering, conference dinner). Besides the conference dinner, the framework program consisted of an organized barbecue evening in one of the oldest student pubs in town. The band, hired for the background music, was casted and chosen by the LOC team as well.

**Sponsoring and Industrial Track.** In order to offer acceptable conference fees, in particular for students, the LOC decided to ask interested partners from the industry, other institutions

and foundations to support the SMBM 2006 financially. The *Jena Centre for Bioinformatics* (JCB) offered to pay the costs for the invited speakers. *Intershop* waived the costs for renting the rooms (about 2800 Euros). *Sun* offered 1000 Euros, merchandising products (bags and biros) and computer equipment for the conference (Internet room). *BioMAX* und *Temis* offered 500 Euros, while *Novartis* supported the conference with 5000 SFR. In a gentle response to these substantial financial contributions, we offered all companies to present their work and products during the conference including an oral presentation round and a booth for each company. *BioMAX*, *Temis* and *Sun* each gave a presentation and thus formed the industrial track of the symposium.

**Student Grants.** Two external students received a grant for participation (workshop and conference). Three local student helpers who were responsible for organizational matters of the event received a grant for participation as well. In particular, Alex Morgan from *MITRE* and Martin Romacker from *Novartis* refused the main part of their reimbursement and donated this money to the students.

**Conference Fees.** The LOC decided to establish an early-bird rate fee for those who registered before 15th of March (late registration after 15th of March until 9th of April). For students, the rates were 200 Euros (300 Euros), for full payers 300 Euros (400 Euros). The conference dinner was 50 Euros (75 Euros) and each tutorial was 75 Euros or 125 Euros when attending two tutorials.

**Paper and Poster Submissions and Acceptances.** The PC received 13 paper submissions. Each paper was reviewed by 3 reviewers who evaluated the quality of the work by a given form. 10 papers were accepted (36 different authors, 9 different countries, 19 different organizations). Out of the 10 accepted papers, 5 papers were proposed to be published in “BMC Bioinformatics” (4.96 impact points). A second reviewing process will decide on the final choice of papers. For the posters, 10 were submitted and 6 were accepted (at least 2 reviewers) coming from 24 different authors, 6 different countries and 11 different organizations. None of them was upgraded to a paper for publication.

**Number of Registrations.** The LOC received registrations from 57 persons (34 full payment, 21 student rates, 2 grants). As far as the tutorials are concerned, the number of participants was well balanced (12 for Rebolz/Kim: “TextMining”, 15 for Morgan/Krallinger: “Evaluation of Text Mining Systems”, 13 for Stevens: “Ontologies in the Life Sciences” and 11 for Staab: “Ontologies and the Semantic Web”). Furthermore, several official representatives from the FSU JENA were invited for the opening session.

**Communicating with the Participants.** A continuous contact (email and phone) with participants, invited speakers, tutors and the industrial participants was necessary in order to answer questions and to control organization issues. During the weeks preceding the conference, 20 to 30 mails per day had to be answered. All in all, the LOC received about 1000 emails.

**Composition of the Program.** After the decisions were made in terms of accepted papers and posters and all speakers were acknowledged and registered (for one poster, none of the authors registered; so the poster was not published), the program was finalized. Given several time constraints by presenters and invited speakers, the presentations were grouped thematically. People from the participant lists were asked to act as session chairs (for the poster, paper and industrial presentations).

**Print Media.** After all final papers and posters were received (at least one author had to register for the conference), the papers and posters were published as a hardcopy version and distributed to the participants. These proceedings also contained a foreword by S. Ananiadou and J. Fluck, the abstracts of the invited and industrial talks and a printed version of the program. Furthermore, all tutorial attendees received a print version of the slides if they were sent to us by the tutors. A printed version of the program was published as well. For the

information about the venue, the LOC composed a flyer with useful information (restaurants, bars, banks, post offices) and a map of the city center. Furthermore, badges, vouchers and receipts were given personally to each participant.

## **Realization of the SMBM (April 9th – April 12th 2006)**

The preparations on site started on Friday, April 7th with the installation of poster walls that were provided by the FSU. *Sun* installed 6 clients and 1 server for Internet access. *IKS*, the company responsible for the network installation at the venue, patched several network boxes and a WLAN connection. The technical equipment (digital projector, sound and microphones) were checked. The bags containing all the necessary information were packed. For every day of the conference, the LOC established a registration and information desk that was run by the local organizers, the student helpers as well as the other members of the JULIE Lab. Their job was to welcome the participants, hand out their conference documents, receive on-site registrations and be available for all kinds of questions. The opening session was composed of talks by the local organizing committee (U. Hahn), the program chairs (J. Fluck), the *JCB* (J. Sühnel) and the rector of the FSU (K. Dicke). During the conference, the venue had to be re-arranged several times (installing the rooms for the tutorials, placing the posters for the poster session, installing tables for the industrial track and for the panel discussion etc.). See also the conference program under <http://supreme.coling.uni-jena.de/content/view/84/143/>.

## **Revision of the SMBM (April – August 2006)**

**Documentation.** The experiences made by the LOC and the PC before and during the symposium were documented. Furthermore, pictures made during the conference were processed and made available online at the conference. As a wish of several attendees, a list of participants was composed (everyone's permission was asked) and published later by email.

**Publications.** All papers (except those that were selected for a journal reviewing process) and posters were published online at CEUR-WS (<http://ftp.informatik.rwth-aachen.de/Publications/CEURWS/Vol-177/>).

**Reviewing Process for the Journal Publication.** 5 papers were selected for a second reviewing process in order to be published in the *BMC Bioinformatics* journal. The reviewing and publishing process is still ongoing.

## **Statistics**

Number of papers (accepted): 10

Number of posters (accepted): 6

Number of participants: 57 (23 students) - industry: 14 - national centers: 11 - academia: 32  
Nations: - Germany (27) - U.K. (9) - USA (5) - France (4) - Japan (4) - Switzerland (3) - Sweden (2) - Spain (1) - Finland (1) - Taiwan (1)

# Training Course on Biomedical Ontology

## Administrative information

**Lead Contractor:** IFOMIS

**Deliverable:** D40

**Author:** Barry Smith, Michelle Carnell, Mathias Brochhausen

## Summary

This deliverable reports on a Training Course on Biomedical Ontology held at Schloss Dagstuhl, Wadern, Germany 12-14 May 2006. It was an International Event Organized by Barry Smith in collaboration with the European Network of Excellence SemanticMining and the US National Center for Biomedical Ontology.

The deliverable audience is the Commission and other project participants. It is not intended for public distribution. The main evaluation criteria for this WP are Valorisation and Dissemination as well as Education and Training.

## Introduction

In May 2003, an International Workshop on Semantic Ontology vs. Ontological Semantics was held in Leipzig, referred to as Ontological Spring.

As a follow-up to this event with a specific slant on the biomedical goals of the NoE, it was decided to offer a similar event with the title “Training Course on Biomedical Ontology”, or Ontological Spring II.

The three-day training course was designed to provide a basic introduction to the field of biomedical ontology and to enhance awareness of current developments and best practices in ontology in the life sciences.

It hoped to gain the interest of participants with the following backgrounds: developers and users of biomedical ontologies, terminologies and coding systems, developers and users of electronic patient record systems, biologists and physicians interested in the possibilities of modern ontologies, and targeted advanced doctoral students, but also interested post-doc and industrial participants or people from hospitals for synergetic effects. The number of participants was to be restricted to about 30 to maximize possibilities for intense discussion. All participants should receive from their attendance in this tutorial hands-on training in ontology design and use.

The course was set up in groups of block lectures by the speakers (Barry Smith: Introduction to Biomedical Ontologies; Werner Ceusters: Biomedical Ontologies and the Electronic Health Record; Olivier Bodenreider: On Mapping, Aligning and Integrating Biomedical Ontologies; Mark Musen: Case Studies in Ontology Development.) It also included a discussion session on the final day.

Short biographies of the speakers and summaries of the course contents are given in the **Results** section below.

## Description of Work

Barry Smith (IFOMIS) was invited by the partners in the SemanticMining Network of Excellence to organise the Training Course within the goals of Workpackage 21 and 26. An initial program was published on the web in Autumn 2005, and a definitive slate of speakers was established in the weeks following, drawing from the leading figures in biomedical ontology throughout the world, with special reference to our NoE partners. The number of participants was capped at about 35 in order to ensure the best learning and discussion environment.

At the event, we collected feedback from all the participants. A summary of their replies is included as an attachment (Dagstuhl course evaluation) to this document.

## Result

The workshop took place 21 – 24 May at Schloss Dagstuhl.

There were four (4) main lecturers, each covering a different aspect of the need for, use and application of biomedical ontologies.

In addition to the speakers, there were 36 further participants at the event, 6 of which were from hospitals and industry.

Participants (speakers and delegates) comprised an international audience:

Germany 12	USA 6
United Kingdom 6	Canada 1
France 4	Australia 1
Sweden 4	The Netherlands 1
Norway 2	Belgium 1
Iceland 1	Italy 1

Several partners from the network sent participants to the meeting (IFOMIS, LIU, EBI, Jena, Manchester).

The workshop announcement and the programme with links to all presentations remain visible at <http://ontology.buffalo.edu/06/os2/>

Short biographies of the speakers follow:

**Barry Smith**

Barry Smith is Julian Park Distinguished Professor of Philosophy in the University at Buffalo (New York, USA) and Director of the Institute for Formal Ontology and Medical Information Science in Saarbrücken, Germany. He is the author of some 400 scientific publications, including 15 authored or edited books, and editor of *The Monist: An International Quarterly Journal of General Philosophical Inquiry*. His research has been funded by the US, Swiss and Austrian National Science Foundations, the Volkswagen Foundation, and the European Union. In 2002 he received in recognition of his scientific achievements the 2.2 Million Euro Wolfgang Paul Award of the Alexander von Humboldt Foundation.

**Werner Ceusters**

Werner Ceusters is Professor of Psychiatry in the University at Buffalo and Director of the Ontology Research Group in the New York State Center of Excellence in Bioinformatics and Life Sciences. He has degrees in medicine, neuropsychiatry, informatics and knowledge engineering and served as coordinator of a series of international research projects in medical natural language processing under the Third, Fourth and Fifth Research Frameworks of the European Commission. Since then, he has also been active in standardisation bodies related to medical terminology such as CEN/TC251/WG2 and ISO/TC215/WG3. In April 1998, he started a new company - Language & Computing nv (L&C) - to exploit the results of his research. He left L&C in 2004, his main interest being now applying and testing a new theoretically-grounded approach to ontological engineering.

**Olivier Bodenreider**

Olivier Bodenreider, MD, PhD, leads Medical Ontology Research in the National Library of Medicine, Bethesda MD, USA. He studied Medicine at the University of Strasbourg and Informatics, Statistics and Epidemiology at the University of Nancy. He received further degrees in Computer Science, Medical Informatics, and Medical Information. His research activities focus on biomedical ontologies, the UMLS and he is the author of many influential publications on these and related topics.

**Mark Musen**

Mark Musen is head of Stanford Medical Informatics, an interdisciplinary research group at Stanford University, with a special focus on the development and use of ontologies within biomedicine. It studies new methods for acquiring, representing, processing, and managing information and data within health care and the biomedical sciences. Musen's work on the Protégé ontology system has led to an open-source technology now used by thousands of developers around the world.

The course blocks offered were:

### **I. Introduction to Biomedical Ontologies (Barry Smith)**

Biomedical Ontologies have developed in an uncoordinated way, often reflecting mere relations of 'association' between what are called 'concepts', and serving primarily the purposes of information extraction from on-line biomedical literature and databases. In recent years, we have learned a great deal about the criteria which must be satisfied if an ontology is to allow true information integration and automatic reasoning across data and information derived from different sources. Such criteria would ensure that ontologies which satisfy them would be automatically interoperable.

We will survey existing biomedical ontologies and show how they fall short of meeting these criteria and how they can be reformed in such a way as to allow true information integration. We also survey the current reform efforts under the auspices of the Open Biological Ontologies (OBO) consortium.

### **II. Biomedical Ontologies and the Electronic Health Record (Werner Ceusters)**

The future of biomedical informatics, in an era of personalized, evidence-based medicine, will increasingly involve reasoning with the sorts of temporally indexed instance data we find in the Electronic Health Record (EHR). We will survey the problems which confront us when we try to use current EHR systems as a basis for such reasoning, and show how an adequate ontology of the biomedical domain can lead to a new and better management of EHR data.

Current Electronic Health Records (EHRs) are organized around two kinds of statements: those reporting observations made, and those reporting acts performed. In neither case does the record involve any direct reference to what such statements are actually about. They record not: what is happening on the side of the patient, but rather: what is said about what is happening.

In addition, existing EHRs embody information primarily in the form of general concept codes, which are tied to instances in reality only indirectly, via designations of persons and times. In contrast to this, we shall describe a regime in which the EHR incorporates direct representations of all clinically salient individuals as they are related together in reality. We show how this will allow us to achieve interoperability among different EHR systems at the level where it really matters: in regard to what is happening in the real world. It will allow us to keep track of particular disorders and of the effects of particular treatments in a precise and unambiguous way, and to engage in new types of reasoning and error checking in relation to the data encoded. We will also show a prototype implementation of an EHR/terminology system conforming to our methodology for ontology design, focusing on how such an implementation can be used to verify data entry in the EHR, to reason with the data, and to use the resultant EHR / terminology system for statistical and other purposes.

### **III. On Mapping, Aligning and Integrating Biomedical Ontologies (Olivier Bodenreider)**

Unlike other domains, biomedicine has a long tradition of developing terminological resources and ontologies for organizing both documents and thoughts. With the increasing need to integrate vast, disparate clinical and biological information resources, these efforts are becoming increasingly important. One result is the development of repositories of cross-referenced ontologies such as the Open Biomedical Ontologies (OBO) library.

We will survey these developments in a critical manner, focusing on mappings, alignments and integrations in various domains of ontology, and paying attention also to the way biomedicine is currently being used as a testbed for Semantic Web technologies and on the associated formalisms for representing ontologies (RDF/S, OWL). Thus for example we will examine how relations among types in single ontologies can be detected from patterns of associations in annotations, and draw parallels between terminology integration in the UMLS and data integration through RDF in the Semantic Web.

### **IV. Case Studies in Ontology Development (Mark Musen)**

We will examine a series of existing biomedical ontologies, indicating how they were initially developed and subsequently maintained, and outlining common problems and solutions. The evolution of ontologies to represent clinical practice guidelines and protocols will provide a foundation for much of the discussion. We will survey early efforts to define entities of clinical care, and trace the expansion of the corresponding models as investigators incorporated increasing numbers of distinctions concerning clinical protocols and guidelines. We will discuss trade-offs in modeling as well as work on ontology evaluation.

We will discuss how the need to create biomedical ontologies has stimulated the creation of new technology for ontology editing and management. Special attention will be paid to the Protegé ontology development environment, and to its OWL plug-in, and also to the new technologies developed by the US National Center for Biomedical Ontology.

The detailed programme may be found here: <http://ontology.buffalo.edu/06/os2/program.htm>

Reading materials to prepare for the course were suggested for the participants along with the programme. Printed readers were supplied to all participants at the workshop.

There was plenty of time for discussion during and after all the lectures so that it was deemed that the discussion session planned for the final day could be dropped.

A full list of the participants may be found in the appendix.

## **Assessment and Conclusion**

The quality and composition of the training course pleased most participants as reflected in the feedback we collected.

Due to large number of interested researchers who were unable to participate due to the limited numbers and due to the positive feedback from the participants, we intend to carry out another similar event at Dagstuhl in June 2007. The preparations for this are already in full swing. The topic is: Reasoning with Biomedical Information: Training Course in Logic for Biomedical Research (cf. <http://www.ifomis.uni-saarland.de/Events/OntospringIII.html>)

## List of Participants

The following people were present at the workshop (40 + 1 no-show)

Name	Status	Affiliation	Place	E-mail	
Bodenreider	Olivier	Speaker	Natlional Library of Medicine	Bethseda, MD	olivier@nlm.nih.gov
Ceusters	Werner	Speaker	NCOR	Buffalo, NY	ceusters@buffalo.edu
Musen	Mark	Speaker	Stanford	CA	musen@SMI.Stanford.EDU
Smith	Barry	Speaker	SUNY	Buffalo, NY	<a href="mailto:phismith@buffalo.edu">phismith@buffalo.edu</a>
Brochhausen	Mathias	Post-doc	IFOMIS	Saarbrücken, D	<a href="mailto:mathias.brochhausen@ifomis.uni-saarland.de">mathias.brochhausen@ifomis.uni-saarland.de</a>
Jansen	Ludger	Post-doc	IFOMIS	Saarbrücken, D	<a href="mailto:ludger.jansen@ifomis.uni-saarland.de">ludger.jansen@ifomis.uni-saarland.de</a>
Manzoor	Shahid	Post-grad	IFOMIS	Saarbrücken, D	shahid.manzoor@ifomis.uni-saarland.de
Schwarz	Ulf	Post-doc	IFOMIS	Saarbrücken, D	ulf.schwarz@ifomis.uni-saarland.de
Spear	Andrew	PhD student	IFOMIS/Buffalo	Saarbrücken, D	adspear@buffalo.edu
Stenzhorn	Holger	PhD student	IFOMIS	Saarbrücken, D	<a href="mailto:holger.stenzhorn@ifomis.uni-saarland.de">holger.stenzhorn@ifomis.uni-saarland.de</a>
Hovenga	Evelyn	Prof.	Health Informatics, CQU	Australia	e.hovenga@cqu.edu.au
Antezana	Erick	PhD student	Ghent University	BE - Belgium	erant@psb.ugent.be
James	Andrew	Prof.	Paediatrics, Univ. of Toronto	Canada -Toronto	andrew.james@sickkids.ca
Wächter	Thomas	PhD student	Bioinformatics, TU Dresden	D	thomas.waechter@biotec.tu-dresden.de
Hanik	Michael	Dr.	GAIA AG	D - Hamburg	Michael.Hanik@gaia-group.com
Beisswanger	Elena	PhD student	Coling, Univ. Jena	D - Jena	beisswanger@coling-uni-jena.de
Mudunuri	Raj	PhD student	ICCAS, Medicine, Univ. Leipzig	D - Leipzig	Raj.Mudunuri@medizin.uni-leipzig.de
Lengger	Christoph	Dr.	GSF FZ für Umwelt und Gesundheit	D - Neuherberg	lengger@gsf.de
Weiler	Gabriele	PhD student	Fraunhofer Institute	D - St. Ingbert	gabriele.weiler@ibmt.fraunhofer.de
Karoui	Lobna	PhD student	Supélec	F	Lobna.Karoui@supelec.fr
Simonet	Michel	Researcher	Faculté de Médecine, La Tronche	F	michel.simonet@imag.fr
Messai	Radja	PhD student	TIMC-IMAG laboratory	F - Grenoble	radja.messai@imag.fr
Jossinet	Fabrice	Researcher	Lab. Bioinf., Univ. Louis Pasteur	F - Strasbourg	f.jossinet@ibmc.u-strasbg.fr
Hardardottir	Arna	Post-grad	University Hospital	Iceland-Reykjavik	arnah@landspitali.is
Keet	Marijke	Researcher	University Bolzano/Bozen	IT	keet@inf.unibz.it
van Mulligen	Erik	Researcher	Med. Informatics, Erasmus MC	NL - Rotterdam	mulligen@gmail.com
Faxvaag	Arild	Prof./MD	NSEP	NO - Trondheim	arild.faxvaag@ntnu.no
Kusnierczyk	Waclaw	MD, PhD stud.	Informatics, NUST	NO - Trondheim	Waclaw.Marcin.Kusnierczyk@idi.ntnu.no
Nyström	Mikael	PhD student	Dept of Biomedical Engineering	SW - Linköping	mikny@imt.liu.se
Petersson	Håkan	Post-doc	Dept of Biomedical Engineering	SW - Linköping	hakan.petersson@imt.liu.se
Sundvall	Erik	PhD student	Dept of Biomedical Engineering	SW - Linköping	erisu@imt.liu.se
Berzell	Martin	PhD student	Linköpings Universitet	SW -Linköping	marbe@ihs.liu.se
Hu	Bo	PhD student	University of Southampton	UK	bh@ecs.soton.ac.uk
Napolitano	Giulio	Post-grad	NICR, Queen's Univ.	UK - Belfast	g.napolitano@qub.ac.uk
Rogulin	Dmitry	Post-grad	Univ. West of England	UK - Bristol	Dmitry.Rogulin@cern.ch
Rocca-Serra	Philippe	Researcher	EBI	UK - Cambridge	rocca@ebi.ac.uk
Schober	Daniel	PhD student	EBI	UK - Cambridge	schober@ebi.ac.uk
Gibson	Andrew	Post-doc	Manchester	UK - Manchester	andrew.p.gibson@manchester.ac.uk
Grey Cowell	Lindsay	Assist. Prof.	Duke University	USA	lgcowell@duke.edu
Grasela	Ted	Dr.	Cognigen Corporation	USA - Buffalo, NY	ted.grasela@cognigencorp.com

# Terminology Conference, Sweden

## Administrative information

Lead contractor: Linköping University

Assisting partners: Karolinska Institute, Swedish Federation for Medical Informatics

Author of report: Hans Åhlfeldt

## Objectives

The objective of the Swedish Terminology Conference, is to establish a forum and meeting place for health care professionals, system developers, informaticians and researchers with an interest in the further development of documentation and sharing of patient information, the multi-professional health record, and follow-up and quality assessment of health care. Both national and international perspectives should be presented, as well as perspectives of the patient, the health care professional, and ICT system providers.

## Conference program

The principle organiser of the two day conference was the Swedish Federation of Medical Informatics (see [www.sfmi.org](http://www.sfmi.org)) together with SemanticMining. The conference was attended by 120 participants, representing the major health care regions in Sweden, the major providers of electronic health record systems (EHRs), and public health care organisations such as National Board of Health and Carelink (national network of health care providers), and universities. The date and location of the conference was September 28-29, 2006, in Kalmar, Sweden.

Program overview (in part translated from Swedish):

National ICT strategy, Lotta Holm-Sjögren, Carelink

eHealth in Europe – language and terminology challenges, key-note by Silas Olsson, European Commission

National projects at Carelink and National Board of Health, Ulla Gerdin, KG Nerander  
Contsys 2, Maria Areblad, Cambio Systems

Terminology issues in national quality registries, Helena Palm, TNC

Experiences from patients having direct access to the EHR, Daniel Carlsson, Linköping University

Poster session

Reference terminologies and multi-professional health records, Hans Åhlfeldt, Linköping University/SemanticMining

- International Classification of Functioning (ICF)

- SNOMED CT

- The openEHR Foundation, Archetypes and CEN standards for EHR (EN 13606)

- Demonstration of tools: SNOMED CT browser, archetype editor, archetype-based EHR

The implication of standards and terminology for the clinician, key-note by Nicholas Hardiker, Salford Health Informatics Research Environment, UK

Group discussions/seminars on

- Patient safety

- Naming of health care specialities (terminology issues)

Education in medical terminology  
Multi-professional health records and reference terminologies  
Panel discussion on national ICT strategies

## **Assessment and Conclusion**

The conference provided an excellent opportunity for experts, system providers and health care professionals to meet, exchange experience and discuss ongoing national and international projects in the area of medical terminologies and next generation of EHRs. Ongoing work within SemanticMining, particularly in work packages WP22 SNOMED CT and WP26 The Electronic Health Record were presented and received a lot of interest. This national event together with the following SemanticMining Conference on Experiences from SNOMED CT in Copenhagen, October 1-3, had a significant impact on several countries negotiations with the upcoming international organisation for carrying SNOMED CT further on, as well as the general awareness and uptake of international standards in the area. As result of the Swedish conference, workshops are being planned for 2007, where the models and tools of openEHR (archetype editor, repository, terminology binding etc.) and SNOMED CT (conceptual framework, browsers) will be further presented and scrutinised.

# SemanticMining Conference on SNOMED CT (SMCS)

## Administrative information

**Lead contractor:** UKLFR (Freiburg University Hospital)

**Participants:** UKLFR, KI, NBH

**Deliverable:** D33

**Author:** Stefan Schulz (UKLFR), Ulrich Andersen (NBH)

## Objectives

The objective was to organize an international forum for discussing achievements and actual experiences with reference terminology, framework, terminology contents and organizational issues in relation to SNOMED CT®. A broad range of topics were to be addressed:

- Formal and ontological aspects of SNOMED CT®
- Mapping between SNOMED CT® and legacy terminologies and classifications
- Partial and local use of SNOMED CT®
- SNOMED CT® and the Electronic Health Record
- Questions of terminology maintenance
- SNOMED CT® and nursing terminology
- SNOMED CT® support for Coding and epidemiology
- SNOMED CT® viewers and browsing tools
- Worldwide semantic interoperability
- Translation issues
- Terminological standards
- Clinical use cases
- SNOMED CT® in life science research
- Post-coordination in practice.

SNOMED CT® represents the vision of a universal clinical terminology, covering a broad range of health-related domains and meeting the needs of all health professionals has stimulated numerous health informatics research activities in the last two decades. During this period, SNOMED grew from a pathology-centered vocabulary to a comprehensive clinical terminology. SNOMED Clinical Terms (CT) is the result of a joint development between the UK NHS and the College of American Pathologists (CAP). Some countries and organizations have already licensed SNOMED CT®, and there is an increasing awareness of SNOMED CT® development and implementation all over the world. Still there are only few prototypical implementations of SNOMED CT® in clinical settings, and there are many concerns about the feasibility of such a comprehensive terminology as a basis for the entire health delivery process.

This event, called Semantic Mining Conference on SNOMED CT® (SMCS 2006), was intended to be the first of several European fora for health policy makers, clinicians, nurses, system developers, computer scientists, terminologists and translators. It should embrace both scientific presentations and invited presentations which provide an overview of current efforts and developments in the context of SNOMED CT®.

## Preparation

The entire preparation process of SMCS 2006 took less than seven months. It initiated in March 2006 when the Semantic Mining Board, motivated by discussions with the reviewers at the 2006 ATR session, decided to prepare a European meeting on SNOMED CT®.

The Danish National Board of Health which had previously organized SNOMED CT® meetings at a national level volunteered as local organizer of SMCS 2006, UKLFR and KI took up the organization of the scientific part of the meeting.

A Scientific Program Committee was constituted by the following experts: Stefan Schulz (co-chair), Freiburg, Gunnar Klein (co-chair), Stockholm, Hans Ahlfeldt, Linköping, Ulrich Andersen, Copenhagen, Robert Baud, Geneva, Olivier Bodenreider, Bethesda, Anita Burgun, Rennes, Werner Ceusters, Buffalo, Chris Chute, Rochester, James Cimino, New York, Joachim Dudeck, Giessen, Kin Wah Fung, Bethesda, Kathy Giannangelo, Chicago, Udo Hahn, Jena, Marie-Christine Jaulent, Paris, John Kilbourne, Bethesda, Rüdiger Klar, Freiburg, Arne Kverneland, Copenhagen, Victor Maojo, Madrid, David Markwell, Reading, Erik van Mulligen, Rotterdam, Jean-Marie Rodrigues, St. Étienne, Patrick Ruch, Geneva, Gunther Schadow, Indianapolis, Barry Smith, Buffalo, Kent Spackman, Portland.

26 scientific papers were submitted. Each of them was seen by at least three reviewers.

Sixteen papers were accepted as oral presentations (after modifications), two as demo presentations, and eight as posters.

The following keynote speakers were invited: Alan Rector (University of Manchester), Kent Spackman (University of Oregon), Peter Zanstra (University of Nijmegen), Andrew Wiesenthal, Kaiser Permanente, Anne Casey, (NHS, UK), William Goossen, (Acquest Consultancy, NL), Kevin Donnelly, SNOMED CT®, Martin Severs (NHS UK), Tevfik Bedirhan Üstün (WHO). For the organization of tutorials we invited David Markwell (UK). The conference was advertised in several mailing lists, folders were printed and distributed, and a website was set up at <http://www.hiww.org/smcs2006/>

## Conference

The conference was held between October 1 and 3 in Copenhagen, Denmark. The number of registrations (170) by far exceeded our expectations. The participant countries are listed in Appendix I.

### **The program headings were:**

- Key Note Introductory Presentation
- Framework for Terminology
- Contents of Terminology
- Demos and Poster Presentations of Current Projects
- Organizational Issues

### **Main topics of the conference were:**

- Common terminology/exchange is necessary in order to support electronic health systems
- SNOMED CT® is a comprehensive interdisciplinary clinical terminology for use in electronic health records and communication systems
- SNOMED CT® covers issues within all health professional areas
- SNOMED CT® is not complete but can be extended as required
- There is a need for considerably more coordination of the quality assessment of SNOMED CT®

- Health professionals must share in the responsibility of the quality assessment within specific areas
- There is a need for pilot projects, where SNOMED CT® is used
- There is a need to transfer SNOMED CT® from the College of American Pathologists to a non-profit and internationally rooted terminology organization.

All conference papers were published in the SMCS 2006 proceedings volume which was available at the conference. The full program can be seen in Appendix II and a list of the proceedings in Appendix III.

## Follow-up

After the conference all manuscripts and presentation slides were made accessible through the event's website <http://www.hiww.org/smcs2006/>.

The Scientific Program Committee selected the five best contributions to be included in a special issue of BMC Medical Informatics and Decision Making. They are currently under review.

When a decision has been made about the acceptance of these selected manuscripts, the remaining papers and poster abstracts, together with the presentation slides will be made available on the Open Access platform CEUR-WS.

## Assessment and Conclusion

Potential Charter Members (PCM)<sup>1</sup> of the SNOMED CT® SDO (Standard Developments Organization) had meetings in Copenhagen in connection to the SMCS 2006 Conference with representatives of the College of American Pathologists in order to set up a timeframe for transfer of the IPRs of SNOMED CT®. Due to this development and the need to make decisions on national level about whether or not to join the SNOMED CT® SDO, the SMCS 2006 Conference was offered at the right time and met a high demand of terminology makers, researchers, and health policy makers. This may explain the extraordinary response to this conference, the high level of scientific contributions and the readiness of top experts to give tutorials and keynote presentations.

The feedback of participants was positive to enthusiastic. Excellent keynote presentations showed not only achievements in the SNOMED CT® development, but also shortcomings, concerning the content and maintenance quality. The Danish SNOMED CT® localization experience was presented and shown high interest by representatives from other European countries. The new route SNOMED CT® is taking by the foundation of the SNOMED CT® SDO and its implications were extensively discussed.

SNOMED CT® was recognized as a framework suitable for the content of a multilingual terminology necessary for true semantic interoperability in healthcare.

Summing up SMCS 2006 was a highly satisfying event which took place at the right place, with the right content and at the right time.

---

<sup>1</sup> UK, USA, Lithuania, Australia, Canada, New Zealand and Denmark. Since then two more EU member states, Sweden and the Netherlands, have joined the group of PCMs.

## **Appendix I - Participant Countries**

- Australia, Brazil, Canada, Denmark , Finland, France, Germany, Hungary, Iceland, Lithuania, Luxembourg, The Netherlands, New Zealand, Norway, Poland, Sweden, Switzerland, UK, USA

## Appendix II - Full Program

Sunday October 1 2006			
13.00-14.00	Registration and a light lunch		
	<b>Parallel Tutorials</b>		
14.00-16.15	<table border="0"> <tr> <td style="vertical-align: top;">           Room: 'Grand Ball'            Chair: Ole Terkelsen  <b>Tutorial 1</b>  <b>The fundamental features of SNOMED Clinical Terms®</b>  <b>David Markwell</b>            Consultant to the SNOMED CT® International Editorial Board         </td> <td style="vertical-align: top;">           Room: 'Amalienborg'            Chair: Arne Kverneland  <b>Tutorial 2</b>  <b>Concept Based Translation</b>  <b>Ulrich Andersen, Janni Lerche, Palle G. Petersen, Lene Asholm, Asta Høy, Jesper Sandberg, Linda Lindquist, Jacob Boye Hansen</b>            Danish National Board of Health, InterText Mus&amp;Pen &amp; CareCom A/S         </td> </tr> </table>	Room: 'Grand Ball' Chair: Ole Terkelsen <b>Tutorial 1</b> <b>The fundamental features of SNOMED Clinical Terms®</b> <b>David Markwell</b> Consultant to the SNOMED CT® International Editorial Board	Room: 'Amalienborg' Chair: Arne Kverneland <b>Tutorial 2</b> <b>Concept Based Translation</b> <b>Ulrich Andersen, Janni Lerche, Palle G. Petersen, Lene Asholm, Asta Høy, Jesper Sandberg, Linda Lindquist, Jacob Boye Hansen</b> Danish National Board of Health, InterText Mus&Pen & CareCom A/S
Room: 'Grand Ball' Chair: Ole Terkelsen <b>Tutorial 1</b> <b>The fundamental features of SNOMED Clinical Terms®</b> <b>David Markwell</b> Consultant to the SNOMED CT® International Editorial Board	Room: 'Amalienborg' Chair: Arne Kverneland <b>Tutorial 2</b> <b>Concept Based Translation</b> <b>Ulrich Andersen, Janni Lerche, Palle G. Petersen, Lene Asholm, Asta Høy, Jesper Sandberg, Linda Lindquist, Jacob Boye Hansen</b> Danish National Board of Health, InterText Mus&Pen & CareCom A/S		
16.15-16.45	Coffee Break		
16.45-17.45	Room: 'Grand Ball'. Chair: Ulrich Andersen <b>Key Note</b> <b>Clinical Terminology – Why – How – When?</b> <b>Professor Alan Rector</b> Department of Computer Science at Manchester University		
17.30-19.00	Registration		
19.00-20.00	Reception		

Monday October 2 2006	
8.00-8.30	Registration
8.30-8.40	<b>Welcome</b> <b>Arne Kverneland</b>
8.40-10.15	Room: 'Grand Ball'. Chair: Stefan Schulz <b>Framework for Terminology</b> <b>SNOMED CT® - the Principles, the Hierarchies and the Relationships</b>

	<p><b>Kent A. Spackman</b>, Scientific Director SNOMED CT® International</p> <p><b>What are the Expectations to an International Terminology in Health Care?</b></p> <p><b>Pieter Zanstra</b>, Radboud University Nijmegen Medical Centre, NL</p> <p>Discussion</p>
10.15-10.45	Coffee Break

10.45-12.50	<p>Room: 'Grand Ball'. Chair: David Markwell</p> <p><b>Contents of Terminology</b></p> <p><b>Experiences from Kaiser Permanente</b> <b>How to implement SNOMED CT®</b> <b>Andrew Wiesenthal</b>, Kaiser Permanente</p> <p><b>Is SNOMED CT® Really Multi Professional?</b> <b>Anne Casey</b>, MSc, RSCN, NHS, UK</p> <p><b>Applying SNOMED CT® codes for Electronic Health Records in the Netherlands</b> <b>William Goossen</b>, RN, PhD. Senior Researcher and Consultant, Health and Nursing Informatics, Acquest Consultancy, NL</p> <p>Discussion</p>	
12.45-14.00	Lunch	
	<b>Parallel Tutorials</b>	
14.00-15.15	<p>Room: 'Grand Ball' Chair: Ulrich Andersen</p> <p><b>SNOMED and Ontology</b></p> <ol style="list-style-type: none"> <li><b>Gergely Héja, György Surján, Péter Varga:</b> SNOMED CT and formal ontologies.</li> <li><b>Jon Patrick:</b> Aggregation and Generalization in SNOMED CT®.</li> <li><b>Stefan Schulz, Kornél Markó, Boontawee Suntisrvaraporn:</b> Complex Occurrents in Clinical Terminologies and their Representation in a Formal Language.</li> </ol>	<p>Room: 'Amalienborg' Chair: Hans Åhlfeldt</p> <p><b>SNOMED and Linguistics</b></p> <ol style="list-style-type: none"> <li><b>Hans Rudolf Straub, Maurus Duelli:</b> With Semantic Analysis of Noun Phrases to SNOMED CT and Classification Codes.</li> <li><b>Patrick Ruch, Julien Gobeill, Imad Tbahriti, Robert Baud, Antoine Geissbühler:</b> Automatic Assignment of SNOMED Categories: Preliminary and Qualitative Evaluations.</li> <li><b>Yefeng Wang, Jon Patrick, Graeme Miller, Julie O'Halloran:</b> Linguistic Mapping of Terminologies to SNOMED CT®.</li> </ol>
15.15-17.15	<p>Room: 'Amalienborg' Chair: Gunnar Klein</p> <p><b>Demos, Poster Presentations &amp; Coffee Break</b></p> <p><b>Demos</b></p> <ol style="list-style-type: none"> <li><b>Seth Rodriguez:</b> The NASA Longitudinal Study of Astronaut Health</li> </ol>	

	<p>Participant Summary and SNOMED CT® Coding Tool.</p> <p>2. <b>Brian Levy:</b> Demonstration of Terminology Maintenance and Distribution.</p> <p><b>Posters</b></p> <p>3. <b>Camilla Wiberg Danielsen:</b> Comparing the Danish Concept Council's Upper Ontology with SNOMED CT® and DOLCE.</p> <p>4. <b>Joachim Dudeck:</b> Concept of an Open SNOMED CT® Development Framework (OSDF).</p> <p>5. <b>Kathy Giannangelo, Mary Stanfill:</b> SNOMED CT® and Legacy Terminologies: Partnerships for Semantic Mining Through Maps.</p> <p>6. <b>Catalina Hallett:</b> Textual Rendering of Post-coordinated Clinical Concepts.</p> <p>7. <b>Asta Høy:</b> Issues Relating to the Translation of SNOMED CT® Concepts.</p> <p>8. <b>Janni Lerche, Lene Asholm:</b> The Use of a Multidisciplinary Terminology in the Electronic Health Record.</p> <p>9. <b>Ulla Magdal:</b> Mapping Between SNOMED CT and a European Laboratory Classification.</p> <p>10. <b>Anna Vikström, Ylva Skånér, Lars-Erik Strender, Gunnar Nilsson:</b> How useful is SNOMED CT® for diagnostic coding with the Swedish primary health care version of ICD 10?</p>
19.00	<b>Conference Dinner at the Cassiopeia Restaurant, next to the Tycho Brahe Planetarium</b>

Tuesday October 3 2006			
9.00-10.15	Parallel Sessions		
	<table border="1" style="width: 100%;"> <tr> <td style="width: 50%; vertical-align: top;"> <p>Room: 'Grand Ball' Chair: Pieter Zanstra</p> <p><b>Mapping of Terminologies</b></p> <p>7. <b>Rahil Qamar, Alan Rector:</b> MoST: A System to Semantically Map Clinical Model Data.</p> <p>8. <b>Iulian Alecu, Cedric Bousquet, Marie-Christine Jaulent:</b> Extraction of SNOMED-CT® Associative Relations to Improve Grouping of the Related WHO-</p> </td> <td style="width: 50%; vertical-align: top;"> <p>Room: 'Amalienborg' Chair: Stefan Schulz</p> <p><b>Miscellaneous</b></p> <p>10. <b>Ronald Cornet, Nicolette de Keizer:</b> Forty years of SNOMED: a Literature Review.</p> <p>11. <b>Ming Zhang, Jon Patrick, Donna Truran:</b> Deriving a SNOMED CT® Data Model.</p> </td> </tr> </table>	<p>Room: 'Grand Ball' Chair: Pieter Zanstra</p> <p><b>Mapping of Terminologies</b></p> <p>7. <b>Rahil Qamar, Alan Rector:</b> MoST: A System to Semantically Map Clinical Model Data.</p> <p>8. <b>Iulian Alecu, Cedric Bousquet, Marie-Christine Jaulent:</b> Extraction of SNOMED-CT® Associative Relations to Improve Grouping of the Related WHO-</p>	<p>Room: 'Amalienborg' Chair: Stefan Schulz</p> <p><b>Miscellaneous</b></p> <p>10. <b>Ronald Cornet, Nicolette de Keizer:</b> Forty years of SNOMED: a Literature Review.</p> <p>11. <b>Ming Zhang, Jon Patrick, Donna Truran:</b> Deriving a SNOMED CT® Data Model.</p>
<p>Room: 'Grand Ball' Chair: Pieter Zanstra</p> <p><b>Mapping of Terminologies</b></p> <p>7. <b>Rahil Qamar, Alan Rector:</b> MoST: A System to Semantically Map Clinical Model Data.</p> <p>8. <b>Iulian Alecu, Cedric Bousquet, Marie-Christine Jaulent:</b> Extraction of SNOMED-CT® Associative Relations to Improve Grouping of the Related WHO-</p>	<p>Room: 'Amalienborg' Chair: Stefan Schulz</p> <p><b>Miscellaneous</b></p> <p>10. <b>Ronald Cornet, Nicolette de Keizer:</b> Forty years of SNOMED: a Literature Review.</p> <p>11. <b>Ming Zhang, Jon Patrick, Donna Truran:</b> Deriving a SNOMED CT® Data Model.</p>		

	<p>ART Adverse Drug Reactions Terms.</p> <p>9. <b>Geraldine Wade, S. Trent Rosenbloom:</b> Experiences: Mapping a Legacy Interface Terminology to SNOMED CT®.</p>	<p>12. <b>Erik Sundvall, Rahil Qamar, Mikael Nyström, Mattias Forss, Håkan Petersson, Hans Åhlfeldt, Alan Rector:</b> Integration of Tools for Binding Archetypes to SNOMED CT®.</p>
10.15-10.45	Coffee Break	
10.45-12.30	<p>Room: 'Grand Ball'. Chair: Gunnar Klein</p> <p><b>Organisational Issues</b></p> <p><b>SNOMED CT® International in a Global Perspective</b> Kevin Donnelly, SNOMED CT®</p> <p><b>The New Organisation for SNOMED CT®, Support and Development</b> Martin Severs, NHS UK</p> <p>Discussion</p>	
12.30-14.00	Lunch	
14.00-15.30	<p>Room: 'Grand Ball'. Chair: Arne Kverneland</p> <p><b>The Relation between WHO and International Classification</b> Tevfik Bedirhan Üstün, WHO</p> <p><b>From National Vision to International Perspectives and Back</b> Ulrich Andersen, Danish National Board of Health</p>	
15.30	End of Conference	

# Workshop on Foundations of Clinical Terminologies and Classifications

## Administrative information

**Lead contractor:** UKLFR (Freiburg University Hospital)  
**Responsible:** UKLFR  
**Participants:** UOM  
**Deliverable:** D41  
**Author:** Stefan Schulz (UKLFR)

## Objectives

The objective was to prepare an international workshop to foster research and cooperation addressing the semantic foundations of clinical terminologies. The choice of its location, Timișoara, Romania, underlined the organizers' concern to integrate Eastern European Medical Informatics researchers into the ongoing discussions on biomedical terminology systems.

In the call for papers we emphasized ongoing research involving specialists from different disciplines, such as Medicine, Computer Science, Philosophy, and Linguistics as well as the existence of different genres of biomedical terminology systems and competing approaches mainly centered on the question of whether to represent the world of reality or the world of language.

## Preparation

In order to reach a maximum of synergies we co-located the event with the 2006 Special Topic Conference 2006 "Integrating Biomedical Information. From eCell to ePatient", being prepared by the European Federation of Medical Informatics and the Romanian Association of Medical Informatics. FTC 2006 was therefore integrated into this major event (expecting several hundred participants), so that we did not have to separately handle registration and payment issues.

The scientific Program Committee was constituted by the following experts: Stefan Schulz (Co-Chair, Freiburg, Germany), Jeremy Rogers (Co-Chair, Manchester, UK), Hans Åhlfeldt (Linköping, Sweden), Olivier Bodenreider (Bethesda, U.S.), Anita Burgun (Rennes, France), Werner Ceusters (Saarbrücken, Germany), Christopher Chute (Rochester, U.S.), Ronald Cornet (Amsterdam, The Netherlands), James Cimino (New York, U.S.), Rolf Engelbrecht (Munich, Germany), Josef Ingenerf (Lübeck, Germany), Rüdiger Klar (Freiburg, Germany), Anand Kumar (Saarbrücken, Germany), Alan Rector (Manchester, U.K.), Assa Reichert (Tel Aviv, Israel), Barry Smith (Buffalo, U.S.), Albrecht Zaiss (Freiburg, Germany).

We received 13 submissions. Each of them was seen by at least three reviewers. Twelve contributions were accepted after modifications.

We invited two keynote speakers, Alan Rector (University of Manchester) and Olivier Bodenreider (National Library of Medicine), which are amongst the world top experts of biomedical terminologies.

The workshop was advertised in several mailing lists, and a website was set up at <http://www.imbi.uni-freiburg.de/medinf/fctc-2006/>

## Workshop

The workshop was held on April, 8<sup>th</sup>, in Timișoara, Romania. All speakers were present. Due to its integration into the STC conference, an exact number of participants could not be exactly determined, but we counted, on average, 25 attendees.

Here is the final program of the workshop:

<b>8:15 - 10:45</b>	<b>Session 1 [chair: Stefan Schulz]</b>  <b>Keynote Speech: Olivier Bodenreider</b> Biomedical terminology and beyond: Ontology and terminology services  <b>I. Alecu, C. Bousquet and M.-C. Jaulent</b> Mapping of the WHO-ART Terminology on Snomed CT to Improve Grouping of Related Adverse Drug Reactions  <b>A. Zaiss, S. Hanser, S. Schulz</b> Mapping of ICHI to CCAM Basic Coding System  <b>K. Giannangelo</b> Principles to Guide Maintenance of Classifications
<b>10:45 - 11:05</b>	<b>Coffee Break</b>
<b>11:05 - 12:35</b>	<b>Session 2 [chair: J.M. Rodrigues]</b>  <b>R. Cornet</b> Clinical Terminology in Practical Use for Recording and Researching Reasons for Admission in Intensive Care  <b>P. Kolesa, P. Preckova</b> Effective Creation of Czech Biomedical Ontologies  <b>K. Denecke, I. Kohlhof, J. Bernauer</b> Use Of Multiaxial Indexing for Information Extraction From Medical Texts  <b>H. R. Straub, M. Duelli, H. Mosimann, A. Ulrich, N. Frei</b> From Terminologies to Classifications - the Challenge of Information Reduction
<b>13:00 - 14:00</b>	<b>Lunch</b>
<b>14:00 - 16:00</b>	<b>Session 3 [chair: Olivier Bodenreider]</b>  <b>Keynote Speech: A. Rector</b> Terminology, EHR's and EHR models - what are they 'about'?  <b>J. Ingenerf, R. Linder</b> Ontological Principles Applied to Biomedical Vocabularies  <b>M. Samwald</b> Classes Versus Individuals: Fundamental Design Issues for Ontologies on the Biomedical Semantic Web
<b>16:00 - 16:20</b>	<b>Coffee Break</b>

**J.M. Rodrigues, B. Trombert Paviot, C. Martin, P. Vercherin**

From Classifications and Coding Systems to Ontology: the Role of a CEN Standard: the Categorical Structure

**R. Serban, A. ten Teije**

Formalization of Medical Guidelines Exploiting Medical Thesauri

After the workshop, a common dinner was sponsored by the companies ID (Berlin, Germany) and SemFinder (Kreuzlingen, Switzerland).

All regular workshop papers (not the two invited contributions) were published in the STC 2006 proceedings volume:

**Integrating Biomedical Information: From eCell to ePatient.** Proceedings of the European Federation for Medical Informatics Special Topic Conference 2006. Edited by: A. Reichert, G. Mihalaş, L. Stoicu-Tividar, S. Schulz and R. Engelbrecht. June 2006, 404 pp., softcover. ISBN: 1-58603-614-9, Price: US\$80 / €64 / £44.

## Follow-up

After the workshop all manuscripts and presentation slides were made accessible through the event's website (see above). The program committee selected the three best contributions to be included into a special issue of *Methods of Information in Medicine*. They are currently under review. The special issues is planned to be published in April.

## Assessment and Conclusions

By organizing an international workshop addressing one of the most central topics not only of the discipline of medical informatics but also of the EU Network of Excellence SemanticMining, we contributed to the exchange of ideas on an international level. All contributions were of good to excellent quality and the discussions of high level. The visibility of this event was guaranteed by the publication of the results via the Internet and by a proceedings volume. Due to the collocation with a bigger medical informatics event we could use an existing infrastructure and had not to worry about organizational issues such as user registration and accounting.

However, we must recognize that the interaction with Medical Informatics researchers from Eastern European countries remained far below our expectations. The main reason for this was the schedule of STC. The STC organizers had put our workshop in parallel with paper and poster sessions ROMEDINF (The national conference on medical informatics), sponsor exhibitions, and a workshop on the electronic health record, which prevented many conference participants to attend our presentations. It was also obvious that the more practice-oriented health record workshop attracted more interest amongst East European participants than our more theoretical and research-oriented terminology workshop, in spite of high quality papers and two internationally renowned keynote speakers.

# Exploitation and dissemination

## Administrative information

Lead contractor: Merrall-Ross International Ltd

Participants: DIM, Board

Author of report: Janine Ross

## Dissemination through the web

The MERMIG platform is used as public website available at [www.semanticmining.org](http://www.semanticmining.org), and as internal network communication platform and repository. The web site is regularly updated and populated with new material. During the last year material from the major workshops and conferences has been made available. Site performance and statistics are monitored.

A significant role of this NoE is to provide educational material based on both the workshops and the research. It is hoped that this educational material available through the web site will be useful, not only to students associated with SemanticMining, but also for the general public, and other interested parties. An important way of sharing research results is also through scientific publication. During the last two years, there has been a significant increase in co-authored research papers within the network.

All NoE documentation relating to administration, assembly meetings and conferences has been uploaded to the restricted access site. The Semantic Mining public website is regularly updated and populated with new content. The content of the web pages is mainly general NoE information, public deliverables, as well as an overview of the activities of the NoE. In addition there is a page listing forthcoming international conferences, and pages providing some basic educational material.

To ensure a sustainable web platform after the ending of the project, a new web platform based on the Wiki-technology, hosted by an academic partner, is under development.

## Exploitation

Discussions were held during 2006, between a major international publishing company and the WP20 lead contractor with respect to the possible exploitation of the multilingual medical dictionary. This link was made via the lead contractor of WP9. Discussions are still ongoing, and no agreement has yet been reached. Regarding the Morphosaurus sub-word dictionary, a subset of the WP20 multilingual dictionary, exploitation contracts were closed with a German medical library and a major German publisher of online content.

Discussions have also been held with publishers over the use of SNOMED CT for information retrieval from medical journals. Several European medical publishers e.g. Elsevier, Royal Pharmaceutical Society of Great Britain (for the British National Formulary), are now starting to utilize and incorporate SNOMED CT into their online medical resources. This will help to increase the pan-European use of SNOMED CT and European access to medical information. In addition, there is likely to be a need by medical publishing houses for consultancy work by members of WP22, and a preliminary approach was made to one partner of this NoE for such help in 2006.

Description of results from the various SemanticMining work packages has been sent to ICT-vendors through available e-mailing lists. A series of contacts with industry for exploitation

of results and know-how has been triggered by the joint work program of SemanticMining. Main areas of interest for exploitation are language technology as worked on in WP20, WP24 and WP27, and the EHR-related work ongoing in WP21, WP22, and WP26. Specific discussions concern the uptake of open source components Protégé OWL and openEHR modules for archetype generation and terminology binding.

During the remaining phase of the project, “public-friendly” material will be produced where results from the SemanticMining project will be presented based on the rich list of deliverables. Target groups for distribution will be ICT-vendors, public health care organisations and regional health care networks through available lists with contact information provided by EFMI, the European Commission, and national networks.

One important way of disseminating the scientific results of the Network partners has been the interaction with relevant standards organisations for health information. This is co-ordinated by WP8 but include participation of most of the research work packages.

Among the many activities that the NoE has contributed to are:

- The establishment of the eHealth Standardization Co-ordination Group which in co-operation with WHO and ITU now includes CEN from Europe, ISO, IEEE, HL7, DICOM and OASIS. A web site was established ([www.ehcs.org](http://www.ehcs.org)) with information on all major eHealth standards and activities.
- The further work on finalising the EN 13606 Health Informatics - Electronic Health Record Communication series (in co-operation with WP26) now also being balloted as an ISO standard and in close co-operation with HL7. As a special subtopic of this, a joint CEN-HL7 project on an Archetype Framework Standard in five parts was started with NoE partners in the lead.
- The finalisation of the EN 12967 Health Informatics - Service Architecture (HISA) standard.
- The work on the EN 1614 model for representing a Structure for nomenclature, classification, and coding of properties in clinical laboratory sciences. After intensive discussions at SemanticMining meetings with WP25, a new version was established. During 2006, EN 1614 has been finalised and approved as European standard: “Health Informatics — Representation of dedicated kinds of property in laboratory medicine”. The standard provides a metrology and terminology framework for Laboratory Medicine developed within the NoE and the Committee on Nomenclature Properties and Unit of the IFCC and IUPAC. The new EN 1614 is now being used as input to the LOINC–C-NPU SNOMED CT mapping discussion.
- Work on the new CEN standard for a Categorial structure for system of concepts for human anatomy took a completely new start during 2005 after extensive interactions with the NoE ontology experts. During 2006 the standard was sent out for enquiry, receiving valuable comments from, among others, NoE participants.
- A CEN Technical Specification for medical knowledge resource metadata descriptions has been developed, Clinical knowledge resources - Metadata (MetaKnow).
- Guiding standardization in CEN and ISO in the field of terminology and concept systems on the relation between the world of concepts and the real world described by ontologies (paper by Klein and Smith).