



SemanticMining

NoE 507505

Semantic Interoperability and Data Mining in Biomedicine

D21.1

Report: Ontology Engineering

Report Version: 1.0

Report Preparation Date:

Classification: RE

Contract Start Date: 2004-01-01

Duration: 3 years

Project Co-ordinator: Hans Åhlfeldt

Department of Biomedical Engineering / Medical Informatics

S-581 83 Linköping University, Sweden

<hans.ahlfeldt@imt.liu.se>



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



Table of Contents

1	ADMINISTRATIVE INFORMATION	3
2	INTRODUCTION	3
3	OBJECTIVES OF THE WORK PACKAGE	3
4	RELEVANT QUALITY INDICATORS	4
4.1	Q1 Workshops and symposiums	4
4.2	Q7 Co-authoring of research papers, reports and educational materials	4
4.3	Q9 Jointly executed research programmes	4
5	ONTOLOGY ENGINEERING RESEARCH ACTIVITIES	5
5.1	Workshops and symposiums	5
5.1.1	Summer School	5
5.1.2	Special Interest Workshops	5
5.1.3	International Conference Participation	6
5.2	Co-authoring	7
5.3	Joint research programme	7
5.4	Exchange visits	8
6	ASSESSMENT	9
6.1	Against Quality Indicators	9
6.2	Against Work package Objectives	9
6.2.1	Share understanding across the three ontological disciplines	9
6.2.2	Coordinate future research efforts	9
6.2.3	Coordinate input into standardisation activities	9
6.2.4	To argue the case for ontology-based domain modelling	9
6.2.5	To contribute to a consensus on a bio-medical "upper ontology"	10
6.2.6	To contribute to the convergence of clinical and bio-ontologies	10

1 Administrative information

Lead contractor: UoM (University of Manchester)

Responsible: UOM, IFOMIS, CNR-ISTC, UKLFR, UNIFR

Participants: LiU(IDA), LiU(IMT), KI, CAU, INSERM, ITRI, SU, UGOT

Deliverable: D21.1

Author: Jeremy Rogers

2 Introduction

The SemanticMining network of excellence comprises a partnership of 25 participant organisations from 11 European countries. This deliverable presents a progress report for work package 21 “Research Activity Ontology Engineering”, covering the period from the start of the project in January 2004 until the end of November 2004.

The deliverable target audience is the Commission and other project participants. It is not intended for public distribution.

‘Ontologies’ are being widely promoted as a necessary central component in the next generation of data interoperability and integration projects, both in biomedicine and in other domains. The field of information systems ontology has therefore acquired new importance. However, as the description of the work package notes, the field is divided into three disjoint disciplines and research streams:

- underlying ontological theory and philosophy,
- formal logics for ontologies
- how to construct and maintain ontologies that are functionally useful.

3 Objectives of the work package

The work package objectives are to enable greater knowledge transfer, integration and interworking between its participants. Five specific objectives were identified in the project proposal:

- To share understanding across the three ontological disciplines, especially those issues unique to or especially important to the biomedical domain, and to coordinate future research efforts so as to achieve coherent divisions of labour and to avoid duplication of effort.
- To coordinate input into standardisation activities relevant to biomedical ontologies, including the emerging semantic web and its associated ontology authoring and delivery environments as well as established international medical informatics standards activities such as ISO, CEN, IEEE and HL7.
- To argue the case for developers of bioscience and clinical vocabularies and coding systems to adopt ontology-based domain modelling, and to develop migration pathways that make it practical for them to do so.
- To contribute to a consensus on a bio-medical "upper ontology".
- To contribute to the convergence of clinical and bio-ontologies

Further general objectives of the network are to encourage and enable dialogue between different workpackages within this network, where their interests overlap. WP21 has several such overlaps,

including active liaisons with the SNOMED Evaluation, EHR, SemanticWeb and text mining research activities.

4 Relevant Quality Indicators

Annex I of the contract defines those quality indicators by which the consortium seeks to assess its progress. A subset of these indicators are listed below; these have been adopted in this report in order to structure the assessment of progress in work package 21:

4.1 Q1 Workshops and symposiums

Summer Schools

Number: 1 per year

Length 2 weeks

Special Interest Group Workshops

Number: at least 3 per year; length 3 days; Location: variable sites

Subjects covered: each workshop should cover one or more work packages

Participation in national and international conferences

The NoE should be represented with several NoE member-institutions presenting as a unit via workshops, information meetings, and the like in at least 3 national and 1 international conference per year

4.2 Q7 Co-authoring of research papers, reports and educational materials

Participants in all institutions at senior researcher, post doc and PhD student levels should be involved with participants from other institutions in the preparation of co-authored scientific papers (original research paper, reviews, conference proceedings etc.) and reports.

Baseline: assumed 0

Goal: all research institutions should have at least 1 and an average of 2 co-authored scientific paper per year

4.3 Q9 Jointly executed research programmes

Goal: To qualitatively follow the progress of the jointly executed research programme in relation to objectives and time frame set forth. Assessment will be based on task descriptions and deliverables for each work package.

Baseline: N/A

5 Summary of Activities in 2004

5.1 Workshops and symposiums

5.1.1 Summer School

The 2004 consortium summer school was held over 6 days of scientific session during July at Lake Balaton, Hungary. 80 network participants attended for all or part of the week. A separate consortium deliverable (D7.2) will provide an evaluation of the summer school as a whole.

A significant part of the 2004 summer school program was devoted to ontology research issues, including a one day tutorial on ontological principles and tools, a two day workshop on Ontology Engineering (see Deliverable D12) and a one day workshop on the Semantic Web (see Deliverable D11).

5.1.2 Special Interest Workshops

In addition to the WP12 Ontology Engineering workshop activity, which took place within the summer school, the following special interest ontology research workshops were organized and hosted within the network. Both were attended primarily by network members, but with special invited guests and a small number of delegates external to the network:

Title	Date	Location	Network Delegates	Comments
Gene Ontology Workshop	28/29 May	Leipzig, Germany	50	Delegates from IFOMIS, Dresden/Kiel, Manchester, Linkoping, Frieberg, Jena, and with guests from EBI (Mike Ashburner, Jane Lomax) and UC Berkeley (Suzi Lewis)
SNOMED and DL workshop	22/23 November	Saarbrücken, Germany	50	Content overlap with WP22 (SNOMED) Delegates from IFOMIS, Dresden/Kiel, Manchester, Karolinska, Linkoping, Frieberg, Jena, INSERM, and with guest from USA (Kent Spackman)

Future plans

The next workshop in the prestigious series of occasional IMIA WG2 workshops will be organised by the network in Rome, April 29-May 2 2005 under the heading: "Ontology and Biomedical Informatics"¹. The traditional format is a 'by invitation only' workshop gathering the best international expertise in clinical terminology design.

Proposals for further specialist interest ontology workshops are also included within the draft 2005 consortium workplan, covering topics including human and cognitive factors of ontology engineering, and boundary problems between information models and terminology systems (Joint with WP26)

¹ <http://ontology.buffalo.edu/wg6/>

5.1.3 International Conference Participation

Network sponsored participation in international conferences include:

Title	Date	Location	Network Delegates	Comments
DILS 2004	25-26 March	Leipzig, Germany	4	International Workshop on Data Integration in Life Sciences
Society of UK Indexers	3 rd April	Chester, UK	1	Presentation about Topic Maps and indexing
SOFG workshop	7/8 April	Edinburgh, UK	2	Meeting with researchers external to consortium, to discuss need for a common cross-species, developmental ontology of anatomy for bioinformatics.
EuroMISE	12-15 April	Prague, Czech Republic	4	International Joint Meeting of the European Centre for Medical Informatics, Statistics and Epidemiology
KR-MED workshop	1 June	Whistler, Canada	6	International conference; Papers presented by network members
KR Conference	2-5 June	Whistler, Canada	6	International conference; Papers presented by network members
DL Workshop	6-8 June	Whistler, Canada	4	International conference; Papers presented by network members
IJCAR 2004	2-8 July	Cork, Ireland	1	International conference; Tutorial presented by network member
WG2004	21-23 June	Bad Honnef, Germany	1	International conference on graph-theoretic concepts; Paper presented by network member
ISMB 2004	30 July	Glasgow UK	3	Annual international workshop on bioontologies, within international bioinformatics conference
Computerm 2004	August 3 rd	Geneva, Switzerland	2	International Workshop on Computational Terminology
ECAI 2004	22-27 August	Valencia, Spain	3	International Conference; Papers presented by 3 network members
COLING 2004	23-27 August	Geneva	4	20th International Conference on Computational Linguistics. Keynote speech and 5 papers presented by network members
MEDINFO 2004	7-11 September	San Francisco,	6	International conference; Papers and tutorials presented by network

		USA		members
WONTO 2004	28 September	Sao Luis, Brazil	1	WORKSHOP ON ONTOLOGIES AND THEIR APPLICATIONS colocated with Brazilian Symposium on Artificial Intelligence. Paper presented by network member
FOIS 2004	4-6 November	Turin, Italy	5	International Conference on Formal Ontology in Information Systems. Papers presented.
TERMINFO	17 November	Houston, USA	1	In collaboration with WP22 (SNOMED) Meeting with USA researchers external to consortium to discuss interface of SNOMED with HL7
Eurorec 2004	25-26 November	Brussels, Belgium	10	Satellite conference on ontology: interoperability of patient records and biomedical informatics
GMDS	26-30 November	Innsbruck, Austria	2	German-Austrian-Swiss joint meeting on Medical Informatics. Keynote speech by network member; papers presented.

5.2 Co-authoring

A special issue of Health Technologies and Informatics was published with joint contributions from network partners in IFOMIS and CNR:

Pisanelli, D (Ed). M.; Gangemi, A.; Catenacci, C.; Battaglia, M.; Grenon, P.; Smith, B.; Goldberg, L; Munn, K.; Papakin, I.; Kumar, A.; Ciccarese, P.; Piazza, M.; Stefanelli, M.; *Health Tech and Inf.* (102) Ontologies in Medicine; 2004 IOS Press, Amsterdam

The summer school and workshop series have been successful at bringing researchers together, and have provided opportunities for sharing of materials or understanding that would not otherwise have happened. A larger number of joint publications are anticipated during 2005.

5.3 Joint research programme

The summer school and workshop series have been particularly successful at bringing the network researchers together, and have provided opportunities for direct discussion and sharing of materials or understanding that would not otherwise have happened.

A common theme of research activity directed towards modelling anatomy was identified at the summer school, and work to explore these issues further in joint papers, and to take practical work forward (particularly with the Digital Anatomist) is ongoing, subject to resource constraints. Several IFOMIS researchers are working on a CEN standardization project in anatomy, including Gunnar Klein, Alan Rector and Anand Kumar.

A new pan-European virtual institute for ontology research (<http://www.econ.uni-saarland.de/>) was founded by IFOMIS in 2004. An interdisciplinary network of philosophically rigorous researchers in ontology, it will work towards the development of a common methodology for ontology design, implementation, evaluation and application. It is founded on the underlying principle that reference ontology construction should start from the theories of how to describe reality developed by philosophers over 2500 years of ontological research, and not with the intuitions of domain experts, or the constraints of software systems.

5.4 Exchange visits

Another indicator of network effort expended towards future joint research might be the number of exchange visits between network members. In practice the workshop series and summer school have provided sufficient opportunity for many informal bilateral and group discussions, and to date there have been very few prolonged exchange visits by individuals:

Barry Smith (IFOMIS) visited Freiburg (UKLFR and UNIFR) Jan 13th

Pierre Grenon (IFOMIS) visited Robert Baud and Patrick Ruch (Geneva) April 14th

Pierre Grenon (IFOMIS) visited ISTC-CNR (Trento) April 14th

Pierre Grenon & Ingvar Johansson (IFOMIS) visited ISTC-CNR (Trento) April 22-23

Sebastien Brandt (Kiel) visited IFOMIS October 4th

Thomas Bittner (IFOMIS) visited UKLFR October 19-20

Boeker, Marko & Schulz (UKLFR) visited IFOMIS November 22-24

Work package assessment

5.5 Assessment against quality indicators

The network has met its year one target for number of specialist workshops organised and attended, including three plenary workshops attended by delegates from all network partners participating in WP21. Two of the workshops were specifically relevant to another work package, WP22.

The network has exceeded its year one target for participation in international conferences.

Co-authoring activity resulting in new material actually in print during Year One has been small. It is, however, to be expected that researchers who have previously known relatively little of each other's spheres of expertise should require time to become familiar with each other before opportunities for joint publication become apparent. The evidence suggests that the familiarisation process has been successfully completed in 2004: we are aware of several joint papers either in preparation, completed and submitted, or accepted for publication in 2005.

The healthy attendance at the three plenary workshops by delegates from all network partners, as well as by a small number of invited people external to the network, confirms that the workshop series is covering common research interests, particularly around the role of formal logics in general ontology construction, and the role of philosophical rigour in constructing specific ontologies (SNOMED, Gene Ontology). These workshops are serving a valuable function in facilitating the sharing, and critiquing, of ideas between the network participants. The planned 2005 workshop series provides evidence that this new dialogue will expand and continue.

5.6 Assessment against work package objectives

5.6.1 Share understanding across the three ontological disciplines

The workshop series and summer schools have clearly been very successful in starting a new dialogue between the three ontological disciplines identified in the work package description. Evidence for this success is provided by the good attendance at the workshops, and the proposed workshop series for 2005.

5.6.2 Coordinate future research efforts

The November Eurorec Satellite conference provides a first example of a coordinated network approach to identifying future collaborative research funding opportunities, particularly under Call 4 of Framework VI, as well to the preparation of specific proposals under this and other calls. At the time of writing, the issues and opportunities raised at that event are scheduled for a plenary discussion by the network assembly in early December 2004.

5.6.3 Coordinate input into standardisation activities

The network continues to support sending selected delegates to both CEN and HL7 meetings, and in addition specialists to the TERMINFO and equivalent UK NHS meetings to explore specifically the interface between clinical terminology/ontology and information models.

5.6.4 To argue the case for ontology-based domain modelling

The network has hosted successful workshops also attended by the leaders of three of the most significant biomedical ontologies – Mike Ashburner from the Gene Ontology, Cornelius Rosse from the Digital Anatomist and Kent Spackman from SNOMED CT. These workshops therefore constitute

arguing the case at the highest possible level. The attendance of these senior figures may be taken as an indicator of the esteem with which the network is viewed.

5.6.5 To contribute to a consensus on a bio-medical "upper ontology"

The network has been successful in identifying and engaging some of the key stakeholders in biomedical ontology, from both inside and outside the network, and is assisting them in adopting new ontological approaches to further development of their work, and in disseminating the results of this new activity. This is an important and necessary step in achieving true consensus.

However, the stakeholder engagement remains still only partial: significant sceptics exist, expressing doubts for the cognitive or computational scaling characteristics of formal ontology and logic based approaches, as well as reservations with regard to the cost effectiveness of the approach assuming scaling problems can be managed.

Future activities will focus on extending the engagement, communicating the ontological choices to stakeholders, and facilitating a consensus forming process.

5.6.6 To contribute to the convergence of clinical and bio-ontologies

A consensus on a common 'upper ontology' is one step towards convergence of clinical and bio-ontologies. Other issues remain to be addressed, including concerns such as potentially incompatible approaches to the intellectual property of developed ontologies. Fulfilling this convergence objective, therefore, remains necessarily more distant at this time.