EU Network of Excellence Semantic Mining

Stefan Schulz

semantic interoperability and data mining in biomedicine
SEMANTIC MINING

HOME   Home

ACCESS MERMIG
ABOUT LOGIN

Information Society
Technologies

Semantic Mining

The aim of the Network of Excellence entitled Semantic Interoperability and Data Mining in Biomedicine (NoE 507505) is to establish Europe as the international scientific leader in medical and biomedical informatics. The long-term goal of the network will be the development of generic methods and tools supporting the critical tasks of the field; data mining, knowledge discovery, knowledge representation, abstraction and indexing of information, semantic-based information retrieval in a complex and high-dimensional information space, and knowledge-based adaptive systems for provision of decision support for dissemination of evidence based medicine. The NoE is a response to the strategic objectives addressed in the IST call 1, areas "eHealth" and "Semantic-based Knowledge Systems".

The general objective of the network is to bridge gaps in the European research infrastructure and to facilitate cross-fertilisation between scientific disciplines. Traditionally academic departments in the domain have their roots either in computer science, system engineering (including a variety of engineering disciplines) or in a medical or clinical context. The network is composed of partners from these scientific areas, all bringing their experience and in-depths knowledge together into a common framework. An important aspect of this is the merging of medical or clinical informatics and bioinformatics including the new fields of genomics and proteomics. Another bridging activity addressed is knowledge-transfer and co-operation between academia and organisations in the health and welfare sector, such as standardisation bodies and public and user-driven health care organisations.

The NoE is based on the partnership of 25 participants from 11 European countries with 110 identified researchers (25 female) and 31 associated PhD students (10 female).

SemanticMining is coordinated by Linköpings universitet, Sweden.
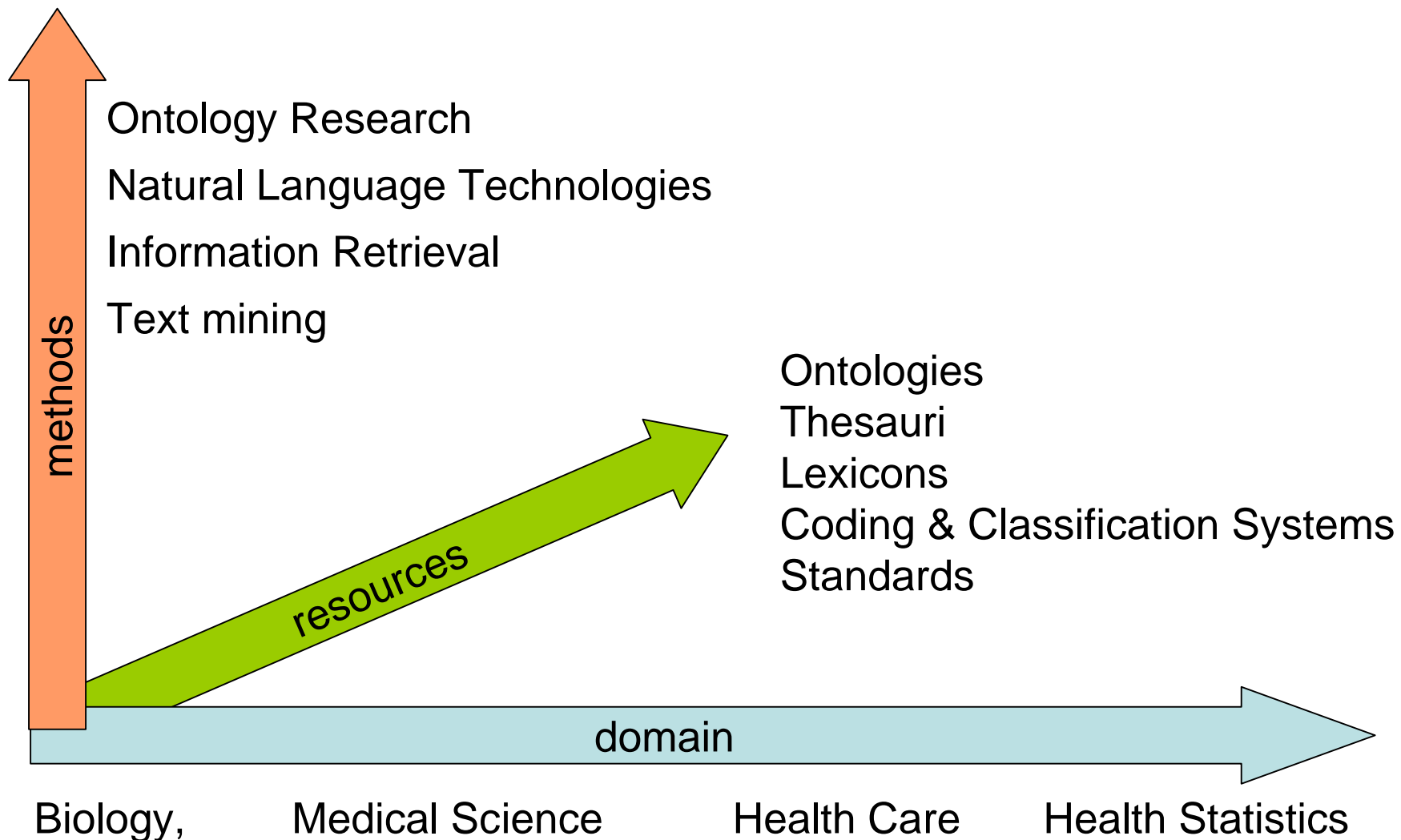
Print

- … to bridge gaps in the European research infrastructure and to facilitate cross-fertilisation between disciplines …
  - Bioinformatics and medical informatics [11 partners]
  - Computer science (engineers, logicians) [3 partners]
  - Natural language technologies [3 partners]
  - Health care organisations, standardisation bodies [6 partners]
  - Philosophy [2 partners]
  - SMEs [2 partners]
- facilitate co-tutoring and exchange of students and researchers

Partners

- **Biomedical Engineering, Medical Informatics, Linköping University, Sweden**
- Computer Science, Linköping University, Sweden
- Committee Nomenclature, Properties and Units in Lab Medicine, Linköping University, Sweden
- Karolinska Institutet, Stockholm, Sweden
- Sahlgrenska University Hospital,Göteborg, Sweden
- Dept of Swedish, Göteborg University, Sweden
- Dept of Medical Informatics, Universitätsklinikum Freiburg, Germany
- Jena University Language and Information Engineering (JULIE),Friedrich-Schiller-Universität, Jena, Germany"
- IFOMIS, Saarland, Germany
- Institute of Informatics and Applied Mathematics,Christian-Albrechts-University of Kiel, Germany
- Division of Medical Informatics, Geneve University Hospital, Switzerland
- Dept of Computer Science, University of Manchester, UK
- Centre for Health Informatics and Multiprofessional Education, University College London, UK
- The Information Technology Research Institute, University of Brighton, UK
- Public Health and Medical Informatics Laboratory, Broussais University Hospital,Paris, France
- Institute of Cognitive Science, Laboratory for Applied Ontology , Italy
- European Bioinformatics Institute, UK
- National Institute for Strategic Health Research, Budapest, Hungary
- WHO Collaborating Centre for Classification of Diseases in the Nordic countries, Uppsala University, Sweden
- The National Board of Health and Welfare, Sweden
- National Research and Development Centre for Welfare and Health, Finland
- KITH AS, Norway
- National Board of Health, Denmark
- Merrall-Ross International Ltd, UK
- European Dynamics S.A., Greece

- Principles in ontology engineering
  - examples: FMA, GO, SNOMED CT
- Evaluation of SNOMED CT
  - strategies and experiences from evaluation and translation
- Concept systems in laboratory medicine
  - communication between bioinformatics, laboratory medicine and the EHR
- Multi-lingual medical dictionaries
  - English, German, French, Portuguese, Spanish, Swedish …
  - Semi-automated lexicon acquisition techniques
- Data/text mining in bioinformatics
  - NLP, IR applied in biomedicine
- The semantic-based electronic health record
  - contribution to standards, information models and concept systems
- What can ontologies do for health statistics?
  - information quality versus aggregation level
  - use of SNOMED CT as aggregation system

Ontology Engineering: Objectives, Activities

- Share understanding across communities
  - Philosophy, Logicians, SW Engineers, Linguists, Domain experts
- Coordinate future research efforts
- Coordinate input to standardisation activities
  - ISO, CEN, IEEE and HL7.
- Argue case for ontology-based biomedical vocabularies and coding systems
- Contribute to a consensus on a biomedical "upper ontology".
- Contribute to the convergence of biomedical ontologies
- SNOMED CT and description logics

Three lines of work:

- Subword lexicon: Links minimal, semantically atomic lexical units in 6 languages
(approx. 80,000 entries, 27,000 equivalence classes). Purpose: Cross-language text retrieval, semantic interface between medical dictionaries

- Semi automated lexical acquisition: guessing Spanish subwords out of Portuguese subwords, and Swedish out of German and English ones.

- Common Lexicon Interchange Format
Based on the MULTEXT morpho-syntactic description. Facilitates the re-use of lexical resources

- Shared platform for multilingual lexicons and corpora

☑ ALL ☑ German ☑ English ☑ Portuguese ☑ Spanish ☑ French ☑ Swedish

Order 1: Lexeme ▾ Order 2: Lexeme ▾ fra Search

Order 1: Lexeme ▾ Order 2: Lexeme ▾ bruch Search

Up

○	fractur	Stem	null	
○	fracture	Stem	null	
○	fradiomycin	Stem	dot	
○	fradiomyzin	Stem	fg	
○	fradizin	Stem	phil	
○	fradizin	Stem	phil	
○	fraeg	Stem	schulz	
○	fraes	Stem	schulz	
○	fraess	Stem	schulz	
○	fraeu	Stem	schulz	

Down

○	bruch	Stem	paulo	
○	bruch	Stem	phil	
○	bruch	Stem	null	
○	bruchbaender	Stem	phil	
○	bruchband	Stem	phil	

Eq Class **997** **for indexing**
Edit/View Properties

○	fraktur
○	frakt
○	fratur
○	fractur
○	fract
○	fracture
○	_frakt
○	_frakture
○	_fraktur
○	_fractur
○	_fractales

Eq Class **997**
sense_of
  ○ break
  ○ _break
sense_of
  ○ bruech
sense_of
  ○ abbruch
  ○ abbrech
sense_of
  ◉ bruch
sense_of
  ○ spreng

- Semantic Interoperability
  - Normalized vocabulary (Gene Ontology, MeSH…)
  - Online integration tool:
    http://www.ebi.ac.uk/Rebholz-srv/whatizit/form.jsp
- Information Retrieval and Extraction
  - Gene and Proteins Names, Drugs and Chemicals…
  - Gene Ontology categories:
    - Protein Functions
    - Cellular Components
    - Biological Processes
- Knowledge coupling
  - From Text (MEDLINE; EHR) to Uni-Prot
    → Standard: via Sequence Retrieval System
    → New Tools for New IR modalities

WHATIZIT

-WHATIZIT-

Whatiz the text you want to highlight?
You may copy&paste text from a full paper available in HTML or PDF, for
example.

swissprot+go – Swissprot and Go Terms ▾   Submit

abbreviation – Abbreviation Tagging
swissprot+go – Swissprot and Go Terms
kinetics – Enzyme Kinetics
compounds – Chemical compounds with links to ChEBI
mgi – Mouse Phenotypes
snp – Mutations
locuslink+go – LocusLink and Go Terms
ppi – Protein Interactions
drugs+swissprot – Drug and Protein Names

nal states have become
signaling, cell growth, and **cell**
protein quality control, afford
conformational flexibility to proteins and serve to integrate stress-signaling events that influence
aging and a range of diseases including cancer, cystic fibrosis, amyloidoses, and
neurodegenerative diseases . We describe here characteristics of celastrol, a quinone methide
triterpene and an active component from Chinese herbal medicine identified in a screen of
bioactive small molecules that activates the human heat shock response . From a
structure/function examination, the celastrol structure is remarkably specific and activates **heat
shock transcription factor 1** (**HSF1**) with kinetics similar to those of heat stress, as determined by
the induction of **HSF1** **DNA binding**, **hyperphosphorylation** of **HSF1**, and expression of chaperone
genes . Celastrol can activate heat shock gene **transcription** synergistically with other stresses
and exhibits cytoprotection against subsequent exposures to other forms of lethal cell stress .
These results suggest that celastrols exhibit promise as a new class of pharmacologically active
regulators of the heat shock response .

- Rebholz Group
- Scripted Access
- Download
- Help
- Remove Menu

2004

- Workshop on the Gene Ontology
- Workshops on Ontology, Health Statistics, Semantic Web
- Workshop on NLP for Biomedical Applications
- WHO-FIC meeting on Classifications in Health Care
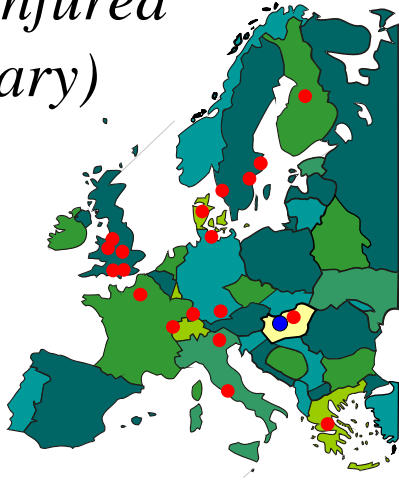- Description Logics and SNOMED CT
- Workshop on EHR

2005

- Workshop on Mereotopolgy
- Symposium on Semantic Mining in Biomedicine, EBI,
- Ontology and Biomedical Informatics, in cooperation with IMIA WG6
- Workshop on Human issues in handling large scale ontologies
- Workshops on the Boundary problem between Information and Terminology models
- Workshop on the Semantic Web
- Workshpo on concept systems in laboratory medicine
- Workshop on Text mining

Balatonfured (Hungary)

www.semanticmining.org