aneurIST
Integrated biomedical informatics for the management of cerebral aneurysms

Medinfo 2007

Design of an Ontology on Cerebral Aneurysms: Representing the Conceptual Space of the @neurIST Project

Susanne Hanser[a], Martin Boeker[a], Kai Kumpf[b] and Stefan Schulz[a]

a *Department of Medical Informatics, University Hospital, Freiburg University, Germany*
b *Fraunhofer Institute for Algorithms and Scientific Computing SCAI, FhG St. Augustin, Germany*

Introduction: @neurIST

- The aim of the European project @neurIST (2006-2009)[1] is the provision of an integrated information structure related to intracranial aneurysms and subarachnoid hemorrhage. The expected benefit for clinicians and scientists includes the support of diagnosis and treatment planning, particularly with regard to aneurysm rupture risk, and an easier access to knowledge in the field, provided by a set of elaborate software tools and platforms.

Disease ontology

- A controlled vocabulary (e.g. an ontology) intends
  - To define a common vocabulary: using textual definitions, hierarchies, relations
  - To make domain assumptions explicit
  - Agreement on meaning

- Task according to DoW: “An ontology binding together the different levels of description of the disease intracranial aneurysm (e.g. from a clinicians or a geneticist’s view) and different sources of information (e.g. literature, clinical databases, image databases).”
  - Intended application contexts: text mining, semantic data mediation

„All relevant entity types …“

- Conceptual space of the project:
  - Clinical Medicine
  - Molecular Biology
  - Epidemiology
  - Simulation
  - Disease
  - Risk factor

- Classification of entities like…
  - Patient
  - Hypertensive Disease
  - Severe Headache
  - Sudden Onset
  - Intracranial Aneurysm
  - Aneurysm Rupture
  - Middle cerebral artery
  - Excessive Coffee Consumption
  - High Risk
  - Mortality
  - Shear stress

Materials and Methods

- Tools
  - Protégé ontology editor
  - OWL DL, Reasoner
- Reference ontologies and terminologies
  - Upper Level: DOLCE
  - Anatomical entity types: Foundational model of anatomy (FMA)
  - Links to Unified Medical Language System (UMLS), SWISSProt, Entrezgene
- Sources of relevant entity types (concepts)
  - clinical databases and information models (CRIM)
  - literature (high frequency domain specific terms extracted from abstracts)
  - domain experts
  - the UMLS metathesaurus and
  - open source databases (e.g. in the molecular biology field)

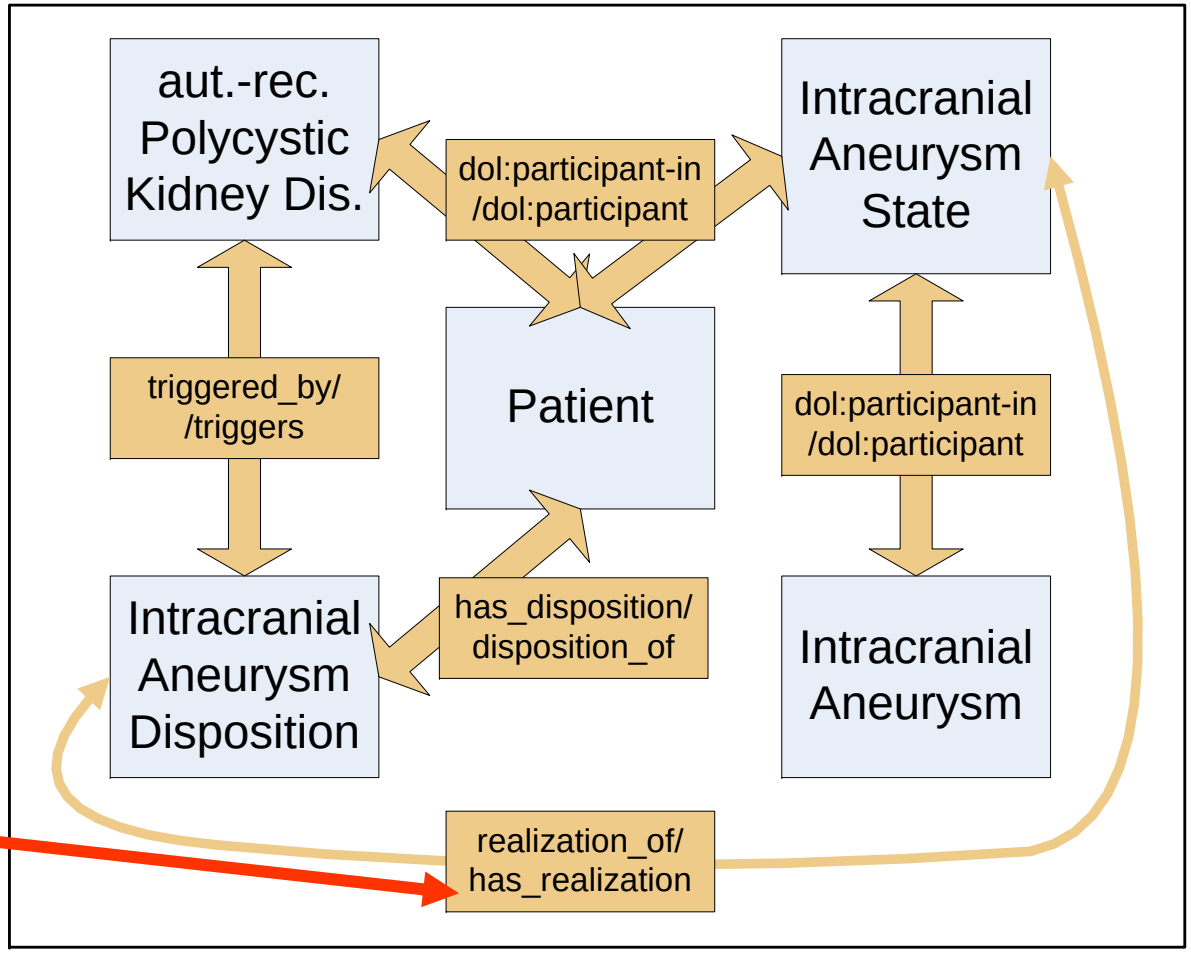
Ontology Design

- Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [2] as the basic onto-logical framework (DOLCE-Lite)
- Top-Level ”Particular” subdivided in terms of:
  - Endurants (including object- and substance-like entities, e.g. “Internal Carotid Artery”, “Blood”)
  - Perdurants (event- and state-like entities, e.g. “Angioplasty”, “Hypertensive Disease”)
  - Qualitys (individual attributes, e.g. “Blood Pressure”, “Sex”) and
  - Abstracts (mainly conceptual “regions” for structuring attributes – “values”, e.g. “Male” as a possible value for the quality “Sex”)

Results

- Coverage
  - 2000 entity types, which mostly belong to the realm of clinical medicine, throughout classified according to DOLCE top level categories
  - 86 properties (relations, e.g. part-of, participates-in)
  - Mapping to Unified Medical Language System in 51%
- Model of the conceptual space
  - see next slide for an exemplary model of a possible clinical situation in the project:
  - a “patient”, for example, participates in a disease (“Intracranial aneurysm state” which is the pathologic state of bearing an aneurysm), an “intracranial aneurysm (which is an anatomical abnormality) participates, too. The patient suffers from polycystic kidney disease; a fact which according to epidemiological knowledge leads to an increased probability for aneurysm development (i.e “triggers” the disposition to develop an aneurysm).
- Usability supported by:
  - “Particular in Context”: additional hierarchy for the representation of entity types in certain application-related contexts of human knowledge, including the possibility to extract ontology subtrees according to user’s needs, for example for text mining purposes
  - synonym database provides lexical information
  - Web Browser provides easy access to ontology content

Model - Example for a view on Intracranial Aneurysm: @neurIST ontology view on the disease



•Associated Probability

The @neurIST Ontology

Taxonomy

- __Arbitrary_Sum
- Non_Physical_Endurant
- Physical_Endurant
  - Amount_of_Matter
  - Physical_Object
    - Anatomical_Space
      - Subarachnoid_Space
      - Subdural_Space
    - Manufactured_Object
    - Material_Anatomical_Structure
      - Anatomical_Abnormality
      - Biological_Macromolecule
        - Protein
        - Structural_Gene
      - Embryonic_Structure
      - Fully_Formed_Anatomical_Structure
        - Cardinal_Body_Part
        - Cell
        - Cell_Part
        - Organ
        - Organ_Part
        - Organ_System
        - Organ_System_Subdivision
    - Organism
    - Population_Group

Cell | Change Entity | Search Entity

UMLS Identifier (CID)

C0007634

Definition

"The basic structural and functional unit of all organisms. Includes the plasma membrane and any external encapsulating structures such as the cell wall and cell envelope" [GO:curators] (Gene Ontology). "Anatomical structure that consists of cytoplasm surrounded by a plasma membrane, with or without the cell nucleus; together with other cells and intercellular matrix, it constitutes tissues. Examples: lymphocyte, fibroblast, erythrocyte, neuron" (Digital Anatomist)

Restrictions [to be done]

Lexical Information

Pref. Term	Synonym	Language	Source
1	Zelle	de	UMLS
1	cell	en	UMLS
0	normal cell	en	UMLS
1	cellula	la	UMLS
1	celula	sp	UMLS

Version Info

Different views on entities

- Representation of entity types in certain application-related contexts of human knowledge:
  - Clinical
  - Epidemiological
  - Biomolecular
  - Simulation
  - Risk Factor

- Classification of “Tobacco Smoking”?
  - “Tobacco Smoking” is a consumption habit.
  - What was discovered later:
    - “Tobacco Smoking” has a certain influence on the state of the human body
    - “Tobacco Smoking” is associated with certain pathological states (diseases)
    - “Tobacco Smoking” causes certain pathological states (diseases)

- “Tobacco Smoking” is a risk factor for aneurysm rupture” <=> View on a consumption habit as “risk factor” – in epidemiological and clinical context!

Conclusion

- The @neurIST ontology provides a structured representation of the conceptual space using DOLCE:
  - the use of the DOLCE upper level categories facilitated a classification of the @neurIST entity types in the difficult field of clinical terminology with its often vaguely defined vocabulary
  - substantial coverage of the @neurIST conceptual space is already achieved
- The ontology is currently used for text mining purposes
  - application for semantic mediation discussed

Acknowledgements

- This work was generated in the framework of the @neurIST Integrated Project, which is co-financed by the European Commission through the contract no. IST-027703.(http://www.aneurist.org).

References

- Frangi A. @neurIST. Integrated biomedical informatics for the management of cerebral aneurysms. http://www.aneurist.org. Last access 31/05/2007.
- Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A., Schneider, L.: The WonderWeb Library of Foundational Ontologies. WonderWeb Deliverable 17 http://wonderweb.semanticweb.org/index.shtml; http://www.loa-cnr.it/DOLCE.html Last access 01/05/2007.

Address for correspondence

- Susanne Hanser, Department of Medical Informatics, University Hospital, Stefan-Meier-Str. 26, D-79104 Freiburg, Germany,
e-mail: susanne.hanser@uniklinik-freiburg.de