Summary of the satellite conference on Ontologies
Eurorec 2004, Brussels 25 November, 2004

See http://www.ecor.uni-saarland.de/announce/eurorecontologyworkshop.html
for agenda speakers and individual presentations.

Background and rationale

1. THE MORNING SESSION

What: To discuss and highlight opportunities and challenges in integration of clinical and genetic data (information, knowledge) and in linking databases of "bioinformatics world" with "medical informatics world". In particular the role of ontologies such as Gene Ontology and Snomed CT and others.

Why: To better understand the importance and role of existing or future ontologies in integrating biomedical information, i.e. integrating information on different levels (molecule, cell, organ, person, population).

Note: European Commission, IST Programme has launched a call for proposals (call 4).
The ICT for health unit has issued a call on integration of biomedical information and its impact. See http://www.cordis.lu/ist/workprogramme/fp6_workprogramme.htm
(pg 28-29, English pdf version)

What is expected:

Ideas on opportunities and challenges for such integration, for example

- what are the main challenges in integration of clinical and genetic information ?
- what are the approaches to link the different ontologies ?
- is distributed environment of heterogeneous biomolecular and clinical information possible or do we need to build warehouses/biobanks and dump all data into one physical place with one format and standards ?
- what exist today as good example of interoperable biomedical information on which we could build ?

Please note that the term biomedical informatics will be used and should not be confused with bioinformatics. BIOMEDICAL INFORMATICS (BMI) - interdisciplinary area based on synergy between medical informatics, bioinformatics and possibly neuroinformatics. For a first roadmap for BMI (http://bioinfomed.isciii.es/Bioinfomed/The%20White%20Paper/results/White%20Paper.pdf)
See also varying other definitions at: http://www.biohealthmatics.com/knowcenter.aspx, http://ycmi.med.yale.edu/research/, http://cci.uchicago.edu/about.html#definitions

Examples of EU projects include INFOGENMED (5th FP), SEMANTIC MINING, BIOPATTERN INFOBIOMED (6th FP): http://www.cordis.lu/ist/directorate_c/ehealth/projectbooklet/projects.html

2. THE AFTERNOON SESSION

What: What is ontology and its role in semantic interoperability of health information systems (e.g. Electronic health records systems)

Why: Please note that all Member States on regional, national and cross border level struggle with this concept/requirement.
European Commission has issued Communication to Member States to agree on common approaches to interoperability by 2006.
See the pages 17 and 24 of our Communication COMM (2004) 356: http://europa.eu.int/information_society/doc/qualif/health/COM_2004_0356_F_EN_ACTE.pdf
Moreover, ICT for health unit is calling (see the above mentioned call for proposals) for a Roadmap on interoperability.

Ontology play crucial role in semantic interoperability which is the basic aspect of interoperability

3. PERSONAL OBSERVATIONS (TO BE COMPLETED)

3.1. *Werner Ceusters*

The word "ontology" is used by in different meanings. There is a need to further explain clearly in what sense(s) it is used and what it aims to represent.

"Ontology" has received lot of attention over the years both in medical informatics and bioinformatics from fairly high number of high quality Organisations.

People involved in ontology could be considered belonging to at least two categories: those who care about the design of "ontologies" itself and those who are more concerned about building tools for ontology development. The tool builders (amongst them the description logics people) work primarily on adequate (in terms of speed and decidability) reasoning facilities, thereby putting constraints on the models that may be used for representations, paying less attention to whether or not the knowledge expressed corresponds to reality.

The ontology developers fall again in at least two categories depending on whether or not they care about a theory describing how the entities in their ontology relates to the actual (real world) entities. For those who don't, it is either a matter of not being aware of the need for such a theory, or the belief that such a theory is not needed in the domain or for the purposes they are building the ontology.

What is common to all groups is the belief that a "knowledge representation" is a key component in making systems interoperable. What is further believed by a majority (but not those that stress the importance of having a theory describing how ontology relates to reality) is that developing and using ontologies is a matter of agreement inside the community.

That belief is however based on a too high emphasis on the terminological aspects of an "ontology": inside a community, there must indeed be agreement on what meanings are allowed to be assigned to terms (otherwise verbal communication would not be possible, and here we stress the adjective "verbal"), but the conditions agreed upon to use a specific term for entities in reality are not necessarily those conditions that explain why and in what way such entities are different from other entities.

My conclusion of the workshop is then:

- given the wide interest in "ontology", it should be put high on the agenda of the research and standardisation community, and as a consequence also on the priorities for CEC-funded projects

- given the confusion around the topic, there is a huge need for education

- given the broad interpretation of the term "ontology", one should come to an "ontology of 'ontologies'". It would help to have a document that describes the various things that the developers of them denote as "ontologies" and how they differ from each other. The purpose would not be that from a certain point on the term "ontology" should be reserved to one specific type of such things, but rather that anybody using the term would be forced to explain what he exactly means.

- given the so many claims by various speakers how useful and good their "ontologies" are, one should come to assessment criteria to validate such claims.

3.2 *Barry Smith*

a) Education: The Gene Ontology and Open Biological Ontologies Consortium has recognised the need for systematic education across the biomedical community in the principles governing the construction of good ontologies if everyone builds ontologies using rigorous principles and using the same top-level categories and relations, then the products of their work will be to that degree automatically interoperable. I was asked to organise a series of intense ontology training workshops directed to biologists in the US. Something similar should be worked out also in Europe.

b) Standards: Most of the existing international standards (e.g. in terminology) were created at a time when the requirements for good ontologies and good controlled vocabularies were not yet clear. Eeffort should be made to educate people in the need for more up-to-date standards

c) Accountability: work on data and information integration -- both with and without ontologies -- should be supported by real methods for achieving accountability. One proposal: there are huge bodies of data existing in hospitals which have for a variety of reasons not been harvested. Develop competitions to find the best methodologies for harvesting this data with real gold standards and real measures of success governing i) applications of the results to clinical care, public health, etc., ii) integration with genomics-based data to develop personalized care, iii) integration with third-parties, e.g. drug companies, to develop methods

for extracting new kinds of practically useful information about side-effects or drug-drug interactions.

The important thing -- and this should be part of any roadmap -- is to develop real measures of success; it is not enough that work should merely be done (on ontologies or anything else); one needs to know in advance what will count as successful work.

3.3. *Jean Marie Rodrigues*

At first there is a need to agree on the definition of the field “ontology” and on the main issues to be addressed .

1 Definition

The communication from the commission on eHealth-making healthcare better for European citizens :An action plan for a European e-Health Area (Sec(2004)539) has proposed a initial definition : “ An ontology defines the terms used to describe and represent an area of knowledge ,and are used by people, data bases and applications that need to share domain information ”.

This definition can be considered as an interesting starting point which must be clarify at least on 2 points :

Terms must be replaced by *entities* for instance and people who at large cannot be directly concerned by an ontology by *to support healthcare professionals and people*

Another important clarification is to precise the goal to share domain information by *to share domain knowledge* which is in favour of an open source method of coordination.

2 Main issues

The main issue concerning ontologies is their role in the interoperability between any type of health records from classification and coding systems to fine grained clinical terminologies supporting decision making or alerts.

The issues between designers and builders and between designers aiming at the essential verity or at harvesting the huge bodies of data existing in the domain must not prevent the final users to benefit from the knowledge and the tools to share information.

This priority must be emphasized and reinforced by the increasing quantity of knowledge supplied by Genomic which is in favour of a mixed approached of ontology within the bioinformatics community.

As a conclusion I would recommend a more user based definitions of ontologies within the health domain : “Formal knowledge representation with editing ,reasoning and nlp software

tools as OPEN SOURCE to share acquisition, validation ,use between countries, centres and languages"

Concerning the roadmap on interoperability of health records it is important to have a macro view and to proceed on the road with micro approaches

The macro view

The interoperability of health records shares with other fields common problems and generic standards are to be used but the semantic interoperability needs to address some specific needs of the health domain due to the balance between the ambiguities of verbal and term communications between different professional languages and different goals and the personal risks for the patients .

Since more than 10 years CEN ,ISO and different international users groups like HL7 have developed health domain standards in 3 Directions

1 Architecture

2 Archetypes/templates

3 Terminology

4 In parallel but with some contacts ontology works have produced some partial but important outputs made possible by technology advanced developments

All these 4 lines are essential for the semantic interoperability and must proceed together .

For these reasons the European commission must support an important pedagogic program for the member states decision makers who are under estimating the complexity of the issues .

On the other hand due to the aims of eHealth action plan the funding of research and development in ontology must be put as an high priority as well as the integrations of the different standards lines within real world implementations

The micro approaches

The users must stay the main topic of semantic interoperability .For these reasons micro tests must be supported for they can be on one hand very useful but as well they are the best practical way to train people and to convince then that Information Society developments are able to address new challenges which will ease their life rather than making it more complex as they are often checking

Some examples using ontology did exist for drugs in UK or surgical procedures in France.

The presentation from WHO during the ontology conference in Brussels has showed the future potential developments of microtests on a broad international basis .

It is essential for the commission to support such initiatives with a rigourous evaluation process

3.4 Asuman Dogac

Ontologies are expected to play an important role in semantic interoperability of health information systems. Semantic interoperability is the ability for information shared by systems to be understood at the level of formally defined domain concepts so that the information is computer processable by the receiving system [1]. Through ontology languages like, Web Ontology Language (OWL), the domain concepts can be formally defined.

Ontologies can play two major roles in the healthcare informatics: one is to provide a source of shared and precisely defined terms which can be used to dynamically discover and share domain artefacts. The other is to reason about the domain concepts.

Within the scope of the IST-2002-002103-Artemis project [2], we have been using ontologies for the semantic interoperability of healthcare information systems. The brief descriptions of these efforts are presented in the following:

1. AMEF (Artemis Message Exchange Framework) [3] is developed to provide the exchange of meaningful clinical information among healthcare institutes through semantic mediation. The framework involves first providing the mapping of source ontology into target message ontology with the help of a mapping tool which produces a mapping definition. This mapping definition is then used to automatically transform the source ontology message instances into target message instances. Through a prototype implementation, we demonstrate how to mediate between HL7 Version 2 and HL7 Version 3 messages. However, the framework proposed is generic enough to mediate between any incompatible healthcare standards that are currently in use.
   The semantic mediation between HL7 Version 2 and HL7 Version 3 messages in AMEF Framework is realized in two phases:
   - *Message Ontology Mapping Process*: In the first phase, the message ontologies of two healthcare institutes are mapped one another. Assume that healthcare institute A is HL7 Version 2 compliant and healthcare institute B is HL7 Version 3 compliant. The message ontologies of these institutes are mapped one into other by using an ontology mapping tool. For this purpose we have developed an OWL ontology mapping tool, namely, OWLmt. With the help of a GUI, OWLmt allows defining semantic mappings between structurally different but semantically overlapping OWL ontologies, and produces a “Mapping Definition". Since message ontologies for HL7 messages do not exist yet, we use the HL7 Version 2 and Version 3 XML Schemas (XSDs) to generate OWL ontologies. This process, called “Conceptual Normalization" [6] produces a “Normalization map" describing how a specific message XSD is transformed into the corresponding OWL schema. Note that Version 3 message schemas are generated through HL7 Hierarchical Message Defnition (HMD) process. The “Mapping Definitions" and the “Normalization map" produced in the first phase are used during the second phase to automatically transform the message instances one into another.
   - *Message Instance Mapping*: In the second phase, first the XML message instances of healthcare institute A are transformed into OWL instances by using the “Data Normalization" engine. Note that if the message is in EDI format, it is first converted to XML. Then by using the “Mapping definition"s, OWLsource (healthcare institute A) messages instances are transformed into the OWL target (healthcare institute B) message instances. Finally the OWL messages are converted to XML again through the “Data Normalization" engine.

   A proof of concept implementation of the system is available at http://www.srdc.metu.edu.tr/~artemis/owlmt/.
2. Within the scope of the Artemis project, we also use ebXML Registry semantic constructs for annotating, storing, discovering and retrieving archetypes [4]. For semantic annotation of archetypes, we developed an example archetype metadata ontology and developed the techniques to access archetype semantics through ebXML query facilities.

3. We also use ontologies to semantically annotate Web services in the healthcare domain [5]. An essential element in defining the semantic of Web services is the domain knowledge. Medical informatics is one of the few domains to have considerable domain knowledge exposed through standards. These standards offer significant value in terms of expressing the semantic of Web services in the healthcare domain. The Artemis project exploits ontologies based on the domain knowledge exposed by the healthcare information standards through standard bodies like HL7, CEN TC251 and GEHR. Artemis Web service architecture does not propose globally agreed ontologies; rather healthcare institutes reconcile their semantic differences through a mediator component.

[1] ISO/TS Health Informatics - Requirements for an electronic health record architecture, Technical Specification, International Organization for Standardization (ISO), Geneva, Switzerland, 2004.
[2] Artemis project, http://www.srdc.metu.edu.tr/webpage/projects/artemis/
[3] Veli Bicer, Gokce Laleci, Asuman Dogac, Yildiray Kabak, "Artemis Message Exchange Framework: Semantic Interoperability of Exchanged Messages in the Healthcare Domain", submitted for publication, http://www.srdc.metu.edu.tr/webpage/projects/artemis/publications/_MappingHL7v3-v2.pdf
[4] Asuman Dogac, Gokce Laleci, Yildiray Kabak, Seda Unal, Thomas Beale, Sam Heard, Peter Elkin, Farrukh Najmi, Carl Mattocks, David Webber, "Exploiting ebXML Registry Semantic Constructs for Handling Archetype Metadata in Healthcare Informatics", accepted for publication, International Journal of Metadata, Semantics and Ontologies (Inderscience),
http://www.srdc.metu.edu.tr/webpage/projects/artemis/publications/IJMSO.pdf
[5] Dogac, A., Laleci, G., Kirbas S., Kabak Y., Sinir S., Yildiz A. Gurcan, Y., "Artemis: Deploying Semantically Enriched Web Services in the Healthcare Domain", Information Systems Journal (Elsevier), accepted for publication, http://www.srdc.metu.edu.tr/webpage/projects/artemis/ publications/
[6]Harmonise, IST200029329, Tourism Harmonisation Network, Deliverable 3.2 Semantic mapping and Reconciliation Engine subsystems.