Conflicts between terminology and EHR information models as obstacles to semantic interoperability: a scientific review

Louise Pape-Haugaard, Anne Randorff Rasmussen, Pia Britt Elberg, Stig Kjær Andersen

Medical Informatics Group, Department of Health Science and Technology, Aalborg University, Denmark

Abstract and Objective

A focus on sharing knowledge is essential for ubiquitous care. Several terminological approaches can be taken to meet the challenge, when terminology models are used to provide a bridge across a range of standardized EHRs information models. Developments in standardized information models as presented by CEN EN 13606 and HL7 have strived for an intrinsic extensibility based on an ad-hoc methodology. We will address the role of terminology systems and EHR information models in achieving semantic interoperability and highlights issues that arise when combining these two domains. A review of scientific literature on information models and terminology systems is conducted with a purpose to categorize research and on obstacles concerned with combining terminology systems and EHRs information models. The majority of the papers analyzed in our review highlight that issues exists, however only a minority contains extensive discussions on these issues. Findings show that ambiguity is key and one attempt to solve this issue is to use terminology binding.

Keywords:

Medical informatics, Semantics

Methods

Present study is conducted as a metaanalysis, covering both terminology and information models. The objective was to select peer-reviewed English-language papers where terminology and information models play a significant role. The selected papers were published between January 2002 and September 2009. This period were selected as it overlaps with the start of SNOMED CT in 2002, following the fusion between SNOMED RT and CTV3. The search strategy contained Boolean operators AND, OR and NOT which were used in combination with following keywords: 'information model', 'HL7', 'CEN', 'openEHR', 'Archetyp*', 'data model*', 'Data model*', 'terminology*', 'SNOMED CT' and 'decision support'.

Initially, papers were to be divided, following a reading og abstracts, into 4 groups with primary focus on: 1) terminology, 2) EHR information models and 3) both terminology and EHR information model. Abstracts from the third group were selected for further assessment. Discussion sections in included papers have been disseminated to determine how different approaches to deal with interfacing issues are handled, and only papers stating issues relevant for the second objective of current study is included.

Key features have been extracted through analysis supported by questions as '*which elements in the discussion section are related to the issues and overlaps across terminology systems and EHR information models?* The obstacles have been identified by extracting key issues in selected papers. In this case questions such as the following were asked: '*which main issue can be recognized across categories?'* and *'how are views on issues argued and documented?*

Results

The 16 included papers can be labeled to a meta-category being methodologically approach and further sub-labeled as belonging to one of the following subcategories, which categorizes the research.

Ontological:	Representation of concepts within a clinical application. Involves papers addressing the issue of ensuring a unique concept registration within a specific terminology system.
Architectural:	Facilitating semantic interoperability through primarily software architectural solutions.
Terminology binding:	Linking terminology systems and EHR information models by terminology binding. Characteristic for this solution is; binding occurs from the information model to the terminology system.

Obstacles to semantic interoperability

The substance of interest is the ambiguity happening when terminology systems and EHR information models are to interact and to be implemented to support semantic interoperability.

Multiple mappings: Observations described by an information model may relate to several concepts. Finding appropriate concepts to align with an information model is a very low-practical approach and time-demanding. For example, there exist at least 20 pre-coordinated concepts to describe the relatively simple finding of 'tobacco smoking behavior'.

Semantic ambiguity: Combining data models with terminology can lead to a conflict between the semantics of a given expression.

Conclusion

Findings show that ambiguity is key issue. One attempt to solve this issue is to use terminology binding. This is not straightforward; because the flexibility built into both HL7 and SNOMED CT means that there is normally more than one way to perform the binding. More studies concerning how to solve the key issue needs to be conducted.