Expert Systems with Applications 41 (2014) 3005–3015

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

STEP: An ontology-based smart clinical document template editing and production system

Sejin Nam [a], Sungin Lee [a], James G. Boram Kim [a], Hong-Gee Kim [a,b,c,*]

[a] Biomedical Knowledge Engineering Laboratory, Seoul National University, Republic of Korea
[b] Dental Research Institute, Seoul National University, Republic of Korea
[c] Institute of Human-Environment Interface Biology, Seoul National University, Republic of Korea

ARTICLE INFO

Keywords:
Clinical document template
Electronic medical record
Ontology
Knowledge management system
Semantic web
Web Services

ABSTRACT

Clinical documents are complex in nature and reflective of the knowledge, structured or otherwise, of physicians. A clinical document template (CDT) is that in which this knowledge manifests itself in various relations that exist among clinical concepts or entities. In this work, we present (1) CDT ontology, (2) A web-based knowledge management system called STEP (Smart Template Editing and Production), (3) Web Services Interfaces to STEP, and (4) A GUI-based CDT editor that uses the Web Services. The CDT ontology explicitly specifies clinical document constituents called template description entities (TDE), and their inter-relations in the CDT. STEP stores CDTs and their components in accordance with the CDT ontology. Web Service interfaces were developed for search, retrieval and storage for CDTs and TDEs. STEP holds entities and relations in use, spanning across different functions such as admission and discharge. A GUI-based CDT editor was developed that uses the Web Services to access data stored in STEP.

1. Introduction

As more clinical documents have become electronic under the technical umbrella of electronic medical records (EMR) (Vawdrey, 2008), there has arisen necessity for reuse of, and facilitated access to, existing clinical document templates (CDT). Extant clinical data models, such as Detail Clinical Model (DCM) (Goossen, Goossen-Baremans, & van der Zel, 2010; Coyle, Mori, & Huff, 2003), Clinical Element Model (CEM) (Coyle, Heras, Oniki, & Huff, 2008), Clinical Contents Model (CCM) (Clinical Contents Model Retrieved, 2012), and openEHR Archetypes (Beale, 2003; Chen & et al., 2009; Garde & et al., 2009), offer conceptual tools to organize and represent clinical concepts and inter-relations between them in structural, standardized, and reusable ways. Briefly, CEM uses the Abstract Instance Model to represent clinical concepts, Abstract Constraint Model to constraint values related to concepts, and the Semantic Links to define relations between concepts. CCM uses EQV (Entity, Qualifier/Modifier, Value) structures to represent concepts, and supports concepts specified in Korean, which can be further mapped to standard terminologies such as SNOMED-CT (Systematized Nomenclature of Medicine – Clinical Terms) (SNOMED-CT Retrieved, 2012). openEHR uses reusable structures called Archetypes, and a formal language, ADL (Archetype Definition Language), to describe the archetypes. The CDT ontology set forth in this paper is to capture and represent the current structure of CDTs in use, and better facilitates CDT reusability and creation. That is, the ontology formally represents the structural semantics at the CDT level, not at clinical concept levels as some previous works have focused on (Kim, Ha, Lee, & Kim, 2005).

Ontology has been used extensively in representing biomedical knowledge. It offers semantics to data representations in data models, such as archetypes and CEM. By transforming ADL to OWL (Lezcano, Sicilia, & Rodríguez-Solano, 2011; Martínez-Costa & et al., 2009), comparison, selection, classification, and consistency checking for archetypes can be done efficiently, and the relations between objects in archetype object model (AOM) become explicit to accommodate reasoning. The abstract meta-representation defined in ontology (Tao & et al., 2011) also makes possible semantic processing of clinical data. The CDT ontology has a different focus from these studies had in that its main focus is to capture CDT production activities of domain experts, and re-use them in a semantically consistent way in a given organization.

Production of a CDT typically includes selection and disposition of clinical concepts or elements, which may not be always conformant to standardized clinical models. A CDT is the artifacts of professional decision-making processes, which mainly consist of two above activities: concept selection and concept disposition. These production activities are the core knowledge assets we take and use in CDT creation. An ontology, CDT ontology, to represent these assets are defined in OWL2 (W3C, 2012),

* Corresponding author at: Biomedical Knowledge Engineering Laboratory, Seoul National University, Republic of Korea. Tel.: +82 116627307
E-mail address: hgkim@snu.ac.kr (H.-G. Kim).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.10.031

and it represents both structural and semantics-based clinical knowledge embedded at the level of CDTs. For example, it may not be evident that two clinical concepts, Hypertension and Tuberculosis, are clinically inter-related when viewed from a clinical conceptual level. In the context of an admission note, however, they may belong together under the heading of the "Past Illness". Only in a clinical document do these concepts appear significantly related. The purpose of the ontology is to capture such knowledge. The core modeling constructs of the ontology – template description entities (TDEs) and TDE relations – are similar to the clinical concepts and relations employed in other clinical data models. The focus, however, is on capturing the clinical purpose and intention (i.e., clinical knowledge) in the document template, in such a way that CDT production activities across diverse functions in clinical organizations are facilitated by the use of such knowledge.

Fig. 1 is a typology of models that shows how the ontology fits and interacts with extant clinical data/terminology models. Reference terminologies are usually bound to specific domains such as disease or medication. Clinical data models, developed independently of clinical domains, use these terminologies to organize the structures of clinical concepts. TDEs in CDTs are practically agnostic to this chasm. That is, any concept, or composite concepts made of multiple concepts, can be used (imported or mapped) in TDEs. In essence, the ontology was made to enhance the use of existing clinical data models and reference terminologies than to replace them.

All the activities captured from CDT production become CDT ontology-based knowledge assets, which are managed in STEP, a Web-based knowledge management system for CDTs. SPARQL (W3C, 2013) – based queries were developed that provide access to STEP. A GUI-based CDT editor based on our approach has been developed as well.

The outline for the rest of the paper is as follows: Section 2 delineates detailed expositions on STEP, including the CDT ontology, and the knowledge base populated with TDEs. Section 3 presents a CDT editor that uses the STEP Web Services. Finally, Section 4 presents concluding comments and future work.

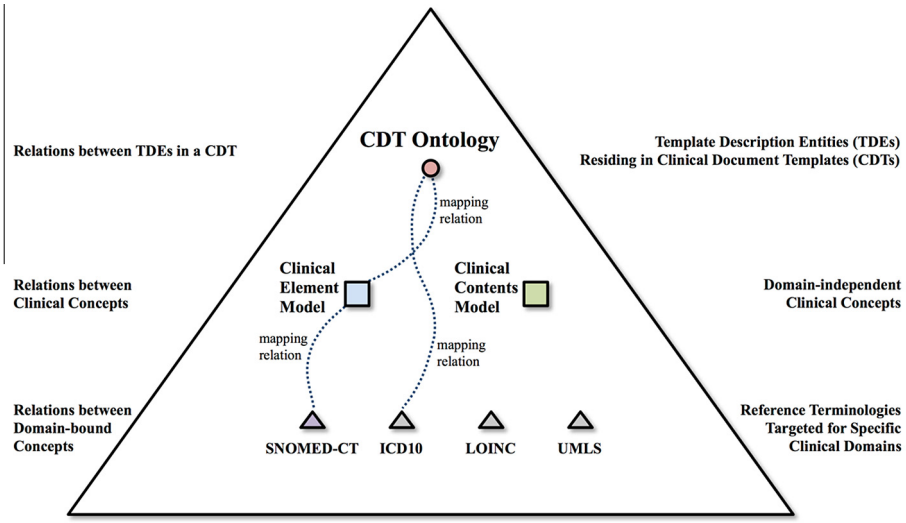

Fig. 1. Model typology.

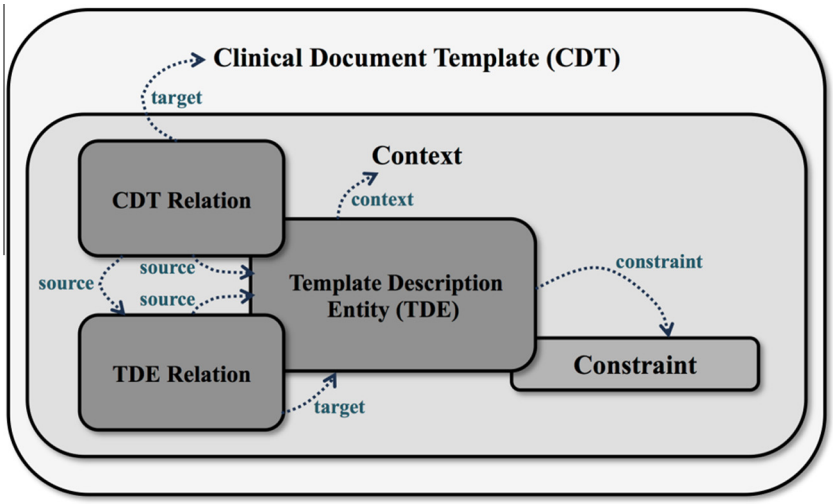

Fig. 2. Conceptual diagram of the CDT ontology.