Quality of Life through Quality of Information
J. Mantas et al. (Eds.)
IOS Press, 2012

doi:10.3233/978-1-61499-101-4-677

Ontology-based Reusable Clinical Document Template Production System

Sejin NAM, Sungin LEE, James G. Boram KIM[1], and Hong-Gee KIM
Biomedical Knowledge Engineering Laboratory, Seoul National University

Abstract. Clinical documents embody professional clinical knowledge. This paper shows an effective clinical document template (CDT) production system that uses a clinical description entity (CDE) model, a CDE ontology, and a knowledge management system called STEP that manages ontology-based clinical description entities. The ontology represents CDEs and their inter-relations, and the STEP system stores and manages CDE ontology-based information regarding CDTs. The system also provides Web Services interfaces for search and reasoning over clinical entities. The system was populated with entities and relations extracted from 35 CDTs that were used in admission, discharge, and progress reports, as well as those used in nursing and operation functions. A clinical document template editor is shown that uses STEP.

Keywords. Electronic Medical Record, Clinical Document Template, Clinical Description Entity, Ontology, Knowledge Base

Introduction

EMR (Electronic Medical Record) systems have spurred increases in the use of electronic documents both in numbers and types [1], which provides more importance in reuse and availability of existing clinical document templates (CDTs).

Research in clinical documents mainly focuses on data modeling such as Clinical Contents Model [2], Clinical Element Model [3][4], and openEHR Archetypes [5][6]. From the perspectives of generating CDTs, core modeling constructs are similar to building blocks of which CDTs are made.

Data modeling defines and classifies clinical concepts of interest in ways that facilitate reuse, whereas CDT production is close to clinical knowledge editing and assembling that reflects the purpose and use intended for the document. And the main processes involved in CDT production are: choice of relevant clinical concepts, and their disposition inside the document. Our study takes these processes as knowledge assets, useful in clinical document template generation. In order to capture these assets, a Clinical Description Entity (CDE) model and a CDE ontology in OWL [7] were defined. A knowledge base was implemented that stores user choice of concepts and their placements in CDTs. The CDE model defines clinical entities (concepts) and their relations found in clinical document templates. The ontology is used as a representation means for knowledge storage purpose. A knowledge management system called STEP

[1] Corresponding Author: James G. Boram Kim, jgkim@bike.re.kr

was implemented that stores ontology-based knowledge artifacts gleaned from CDT generation events. SPARQL [8]-based searches are available that provide access to knowledge assets stored in STEP.

1. Methods

1.1. Overview

Figure 1 shows a snippet of an admission note from which our target CDEs and relations are extracted.

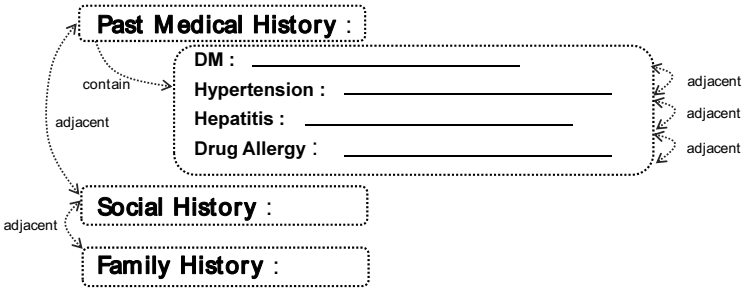

Figure 1. CDE and Relations

Entities found are “Past Medical History”, “DM”, “Hypertension”, “Hepatitis”, “Drug Allergy”, “Social History”, and “Family History”. These entities reflect the template creator’s perceptions on the relevance and importance of them in the context of an admission note. The relation *contain* used in the CDE “Past Medical History” shows that CDEs “DM”, “Hypertension”, “Hepatitis”, and “Drug Allergy” are of importance in the context of admission in relevance to past medical history.

“Social History” and “Family History” CDEs are linked with an *adjacent* relation, which signifies primarily dispositional “nextness” (both upwards and downwards) between two CDEs in the CDE hierarchy. The relation could also be interpreted as semantic closeness as well. In the above case, the CDEs all represent the history of the patient. Another relation, *usedAt* shown at Figure 2, links each template with both *contain* and *adjacent* relations and CDEs used in it.

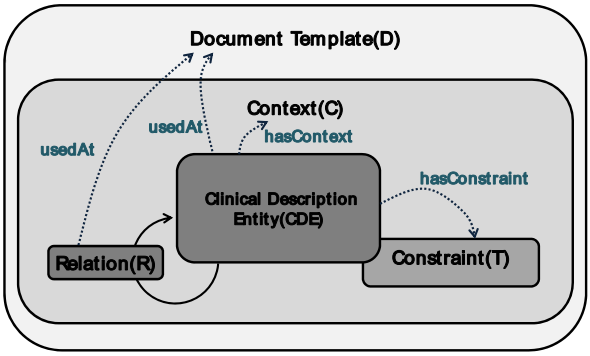

Figure 2. Conceptual Diagram of the CDE Model

A CDE model was created to capture clinical entities (concepts) and relations resident in CDTs, and subsequently an ontology in OWL was defined, which is based on the model. A knowledge management system called STEP was implemented to store and provide Web-based search services over the ontology-based knowledge assets gleaned from CDTs, as shown at Figure 3.

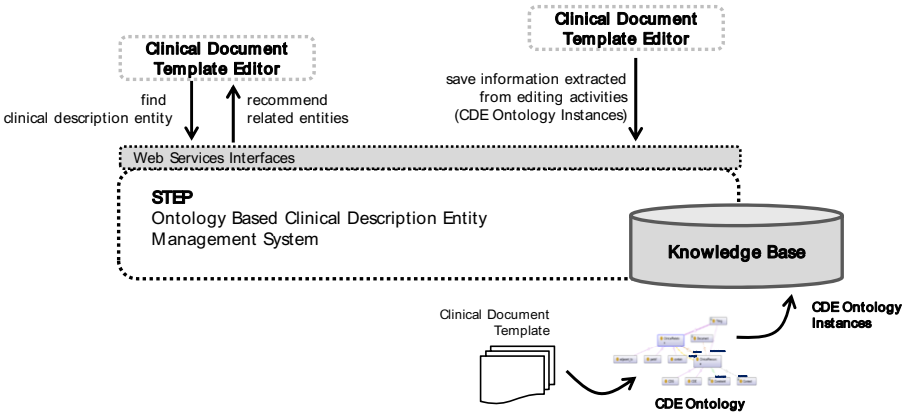

Figure 3. STEP: CDE Ontology-based Knowledge Management System

1.2. CDE Model: Clinical Description Entity Model

Five core constructs are defined in the model: *Document Template*, *Clinical Description Entity*, *Context*, *Relation*, and *Constraint*. Each CDT is a set of CDEs, each of which has relations with other CDEs in a certain context. Three relations – *contain*, *adjacent*, and *usedAt* – represent inclusion, dispositional nextness, and usage of CDEs and relations in a CDT, respectively. Two other relations *hasContext* and *hasConstraint* are defined as well. A *Constraint* defines a range of values a CDE can possess or specifies generic constraints between CDEs. (Figure 2)

1.3. Building Knowledge Base based on the CDE Model

In order to populate our CDE ontology-based knowledge base with instances, we had hand-picked 35 CDTs, most frequently used in fourteen departments, spanning across

admission, discharge, nursing, operation, and progress reports in a medium-sized hospital (500+ beds). A team of three experts participated in instance collection process with the following selection rules:

- Clinical Description Entity: All entities with a key/value pair (e.g. patient name or age) or entities having other entities as their values (e.g. General Appearance entity having Mental Status and Looking Appearance as values)
- *contain* Relation: An entity contains one or more entities.
- *adjacent* Relation: An entity is physically located at the same level and next to another entity in both upward and downward directions.
- *usedAt* Relation: All entities and relations found in a CDT have a *usedAt* relation.

2. Results

2.1. Knowledge Management System: STEP[2]

STEP is a Web-based knowledge management system that offers 1) a CDE ontology-driven knowledge base, and 2) SOAP (Simple Object Access Protocol)-based Web Services for access to its core functionalities by external systems. A sample of Web Services, 11 in total, is shown at **Table 1**.

Table 1. Sample Web Services

Web Service	Description
getCDEListByContext	Retrieve CDEs based on a user-specified context.
findRelatedCDEs	Find CDEs with similar names as a user-specified CDE. Supports cross-language (English and Korean) search.
getCDEContent	Retrieve detailed information of a user-specified CDE.
saveNewCDE	Store a new CDE into STEP.

We gathered in total 967 CDEs with an average of 78 CDEs from each CDT. There were more *adjacent* relations found (3,142) than *contain* relations (1,412), and 150 CDEs (15.5% of the total) used both relations. An average of 6.7 *adjacent* relations was used in 932 (96.4%) CDEs. A CDE within the knowledge base is semantically interlinked with other CDEs or CDTs by relations.

[2] STEP is accessible at: http://step.bike.re.kr/

2.2. Clinical Document Template Editor based on STEP

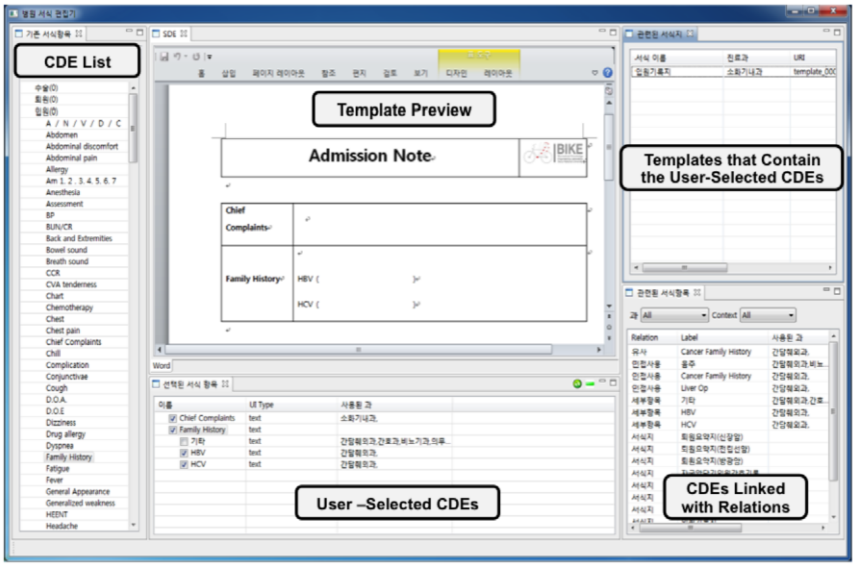

Figure 4. STEP-based Clinical Document Template Editor

Figure 4 is a screen shot of a working standalone editor that uses the Web Services to create clinical document templates. In order to create a template, for example, the user clicks and drags a CDE (multiple CDEs can be chosen) from the CDE List to User-Selected CDEs panel. The CDE is automatically shown in the Template Preview panel. At the same time, search functions are executed that retrieve and display all the templates that contain the chosen CDE, as well as CDEs that are linked with the chosen CDE by *adjacent* or *contain* relations. These retrieved CDEs can be further refined by a context. When a new template is created, it is saved back to the knowledge base.

3. Discussion

STEP houses 12,845 instances now, and we are working to integrate STEP with a CDT production system in a hospital. And it is of paramount importance to ensure semantic interoperability of medical terms used in CDEs with international standard terminology systems. To this end, the current editor provides an interim means to map terms in CDEs to SNOMED CT [9] and ICD-10 [10] terms. We are currently working to integrate STEP with LexCare Suite [11], a comprehensive terminology server that acts as a mapping intermediary between STEP and standard terminology systems.

The primary goal of STEP is to maximize the use of current knowledge assets embodied in the clinical document templates, regardless of their being compliant to standard models such as Clinical Element Model or openEHR Archetypes. As such, the STEP model generic to the extent that it can accommodate both standard and non-standard clinical document structures and terminologies.

STEP is a work in progress, though complete with Web Services and an editor. Evaluation of STEP, such as its usability reflected by easy-of-use and speed in which templates can be created, is necessary to ensure its adaptation in local hospitals.

References

[1] Chae YM, Yoo KB, Kim ES, Chae H. The Adoption of Electronic Medical Records and Decision Support Systems in Korea. Healthc Inform Res. 2100 Sep; 17(03):172-177
[2] www.clinicalcontentsmodel.org (last accessed 25 Dec 2011)
[3] Huff SM, Rocha RA, Coyle JF, Narus SP. Integrating Detailed Clinical Models Into Application Development Tools. Proc 11th Medinfo. 2004; 107(Pt 2):1058-1062
[4] Coyle J, Heras Y, Oniki T, Huff S. Clinical Element Model Introduction & Data Types Reference Manual. 2008
[5] Beale T. Archetypes: Constraint-based Domain Models for Future-proof Information Systems. Proc 11th OOPSLA Workshop Behavioral Semantics. 2002; 16-32
[6] Chen R, Klein GO, Sundvall E, Karlsson D, Åhlfeldt H. Archetype-based conversion of EHR content models: pilot experience with a regional EHR system. BMC Med Inform Decis Mak. 2009; 9(33)
[7] OWL Web Ontology Language Overview. W3C; 2004
[8] SPARQL Query Language for RDF. W3C; 2008
[9] SNOMED Clinical Terms. IHTSDO; 2012
[10] International Classification of Diseases (ICD), Tenth Revision. WHO; 1994
[11] Lee S, Song SJ, Koh SJ, Lee SK, Kim H. National Medical Terminology Server in Korea. Proc 1st Int'l Conf SUComS. 2010; 541-544