HIR
Healthcare Informatics Research

Clinical Data Element Ontology for Unified Indexing and Retrieval of Data Elements across Multiple Metadata Registries

Senator Jeong, PhD[1],*, Hye Hyeon Kim, MS[2],*, Yu Rang Park, PhD[3], Ju Han Kim, MD, PhD[2,4]

[1]National Center for Medical Information and Knowledge, Korea National Institute of Health, Cheongju; [2]Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, Seoul; [3]Department of Biomedical Informatics, Asan Medical Center, Seoul; [4]Systems Biomedical Informatics National Core Research Center (SBI-NCRC), Seoul National University College of Medicine, Seoul, Korea

Objectives: Classification of data elements (DEs), which is used in clinical documents is challenging, even in across ISO/IEC 11179 compliant clinical metadata registries (MDRs) due to no existence of reliable standard for identifying DEs. We suggest the Clinical Data Element Ontology (CDEO) for unified indexing and retrieval of DEs across MDRs. **Methods:** The CDEO was developed through harmonization of existing clinical document models and empirical analysis of MDRs. For specific classification as using data element concept (DEC), The Simple Knowledge Organization System was chosen to represent and organize the DECs. Six basic requirements also were set that the CDEO must meet, including indexing target to be a DEC, organizing DECs using their semantic relationships. For evaluation of the CDEO, three indexers mapped 400 DECs to more than 1 CDEO term in order to determine whether the CDEO produces a consistent index to a given DEC. The level of agreement among the indexers was determined by calculating the intraclass correlation coefficient (ICC). **Results:** We developed CDEO with 578 concepts. Through two application use-case scenarios, usability of the CDEO is evaluated and it fully met all of the considered requirements. The ICC among the three indexers was estimated to be 0.59 (95% confidence interval, 0.52−0.66). **Conclusions:** The CDEO organizes DECs originating from different MDRs into a single unified conceptual structure. It enables highly selective search and retrieval of relevant DEs from multiple MDRs for clinical documentation and clinical research data aggregation.

Keywords: Data Sharing, Registries, Semantics, Information Storage and Retrieval, Ontology

Corresponding Author
Ju Han Kim, MD, PhD
Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 110-799, Korea. Tel: +82-2-740-8320, Fax: +82-2-747-8928, E-mail: juhan@snu.ac.kr
*These authors contributed equally to this work.

I. Introduction

The International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 11179 Metadata Registry (MDR) standard [1] provides a framework that enables the semantic interoperability of data originating from various sources with exact definitions of data elements (DEs). The utility of the MDR standard has been widely recognized, and an increasing number of healthcare stakeholders have been adopting it for the management of metadata for clinical trials and the aggregation of clinical

research data, sharing, reuse, and clinical documentation in Electronic Health Records [2-11].

An appropriate classification scheme is needed to enable searching relevant DEs registered within/across multiple MDRs. Although Part 2 of the MDR standard provides a conceptual basis for classifying DEs [12], it does not specify how to create a classification scheme, or how to designate a particular one. Currently there is no classification scheme in healthcare MDRs by which to organize the DEs into well-organized conceptual categories. The CDE browser (http://cdebrowser.nci.nih.gov/CDEBrowser/) in the Cancer Data Standards Registry and Repository (caDSR) [13] simply organizes DEs into 35 contexts, which represent the sources of the DEs themselves. As Nadkarni and Brandt [14] correctly pointed out, the CDE Browser lacks interconcept relationships and a synonyms search function, and it simply returns a long unorganized list of DEs and their contexts. The browse registry in the Australian Metadata Online Registry (METeOR, http://meteor.aihw.gov.au/) [15] provides an improved search interface, whereby the following six object class groups are displayed in the tree structure: Entity, Life event, Person/group of persons, Service episode, Service/Care event, and Service/Care provider. Properties are classified into 29 groups, such as Accommodation/living characteristics, Birth event, Client characteristic, Communication characteristics, and Crisis event. However, it is far from clear whether the classification scheme is semantically sound and comprehensive. The same problem exists in the Clinical and Histopathological Metadata Registry (CHMR) that was created by three of the present authors [16]. The CHMR currently lacks a means by which to effectively browse and identify DEs by semantic groups.

Existing reference terminologies, such as Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT) and Medical Subject Headings (MeSH), may be used as alternatives for indexing and retrieving DEs from MDRs. However, their functions and structures are not optimized for DEs. The content structure of SNOMED-CT, which was devised to support effective clinical recording of data across specialties and sites of care, is too complex for browsing and retrieving relevant DEs, and MeSH, which was devised for the indexing and retrieval of biomedical journal articles, lacks the terms and content structures required for consistent indexing and retrieval of DEs across multiple clinical MDRs. Locating all relevant DEs across multiple MDRs, at least in the clinical domain, is very difficult because no reliable standard exists for identifying DEs, and there are several different local classification schemes. This problem demands a solution for organizing clinical DEs according to a global concept system.

The purpose of the present study was to develop a system, named the Clinical Data Element Ontology (CDEO), which can be applied to—but is not limited to—ISO/IEC 11179 compliant MDRs. As a global reference concept system, it was intended that the CDEO would enable the unified indexing of DEs registered in multiple MDRs.

1. Related Work
For consistent document naming in the Health Level 7 (HL7)/Logical Observation Identifiers Names and Codes (LOINC), Frazier et al. [17] proposed the Document Ontology (DO), which initially comprised seven primitive axes, but ultimately reduced them to five. Shapiro et al. [18] extended the *Clinical Category* axis of the DO (renamed as *Subject Matter Domain*) to organize and facilitate searching for clinical documents. Chen et al. [19] mapped local clinical document names to the DO and to specific LOINC codes. They identified some limitations in the DO with respect to coverage, granularity, and loss of meaning. Another studies [20, 21] produced an ontology-based definition management for the creation and maintenance of clinical document templates. While the focus of these previous studies was to developing ontologies for the classification of the clinical documents themselves, the present study of the CDEO focused on the indexing and retrieval of DEs, which are the building blocks of clinical documents. We believe that the CDEO is also useful for organizing clinical documents.

II. Methods

1. Setting the Basic Requirements
An ontology with a strong structure is required for effective unified indexing and efficient retrieval of DEs [22]. Therefore, the following basic requirements that must be met by the CDEO were set:

1) Index data element concept (DEC). This is based on a DEC being an abstract of one or more DEs in MDRs, and thus facilitates coherent grouping and selective searching for DEs.

2) Organize DECs into a poly-hierarchical structure, taking into consideration their conceptual relationships, such as broader, narrower, and related concepts.

3) Control the terms and language variants of a concept in order to enable consistent indexing and multilingual searches for DECs across different MDRs.

4) Enable each DEC to have a unique resource identifier (URI) so that it can be uniquely identified and referenced by different MDRs.

5) Be able to identify the two main components of a DEC,

object class and property.

6) Assist users to postcoordinate concepts to find relevant DECs.

The above requirements were used to evaluate the CDEO.

2. Data Representation of the CDEO using the Simple Knowledge Organization System

The Simple Knowledge Organization System (SKOS) [23] was chosen to represent and organize concepts in the CDEO. This decision was based on the understanding that SKOS is an official W3C standard and is highly regarded worldwide. Moreover, in the new version of ISO/IEC 11179-3 standard [24], SKOS supports the concept system that is used for

modeling a classification scheme (Figure 1).

SKOS is a Resource Description Framework (RDF) vocabulary for expressing the basic structure and content of concept systems, such as thesauri, classification schemes, subject headings, or terminologies within the framework of the Semantic Web. The key elements of the SKOS include the following:

1) Concept classes (*skos:ConceptScheme, skos:Concept,* and *skos:Collection*).

2) Labeling properties (*skos:prefLabel, skos:altLabel,* and *skos:hiddenLabel*).

3) Semantic relation properties (*skos:narrower, skos:broader,* and *skos:related*).
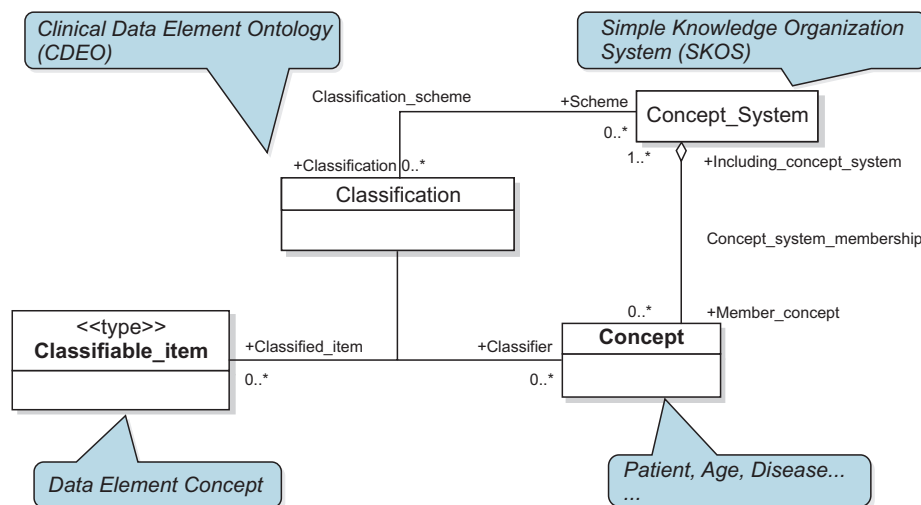

Figure 1. Correspondence between the classification metamodel and the SKOS-based CDEO. Adapted from the classification metamodel region in ISO/IEC FDIS 11179-3:2012(E).
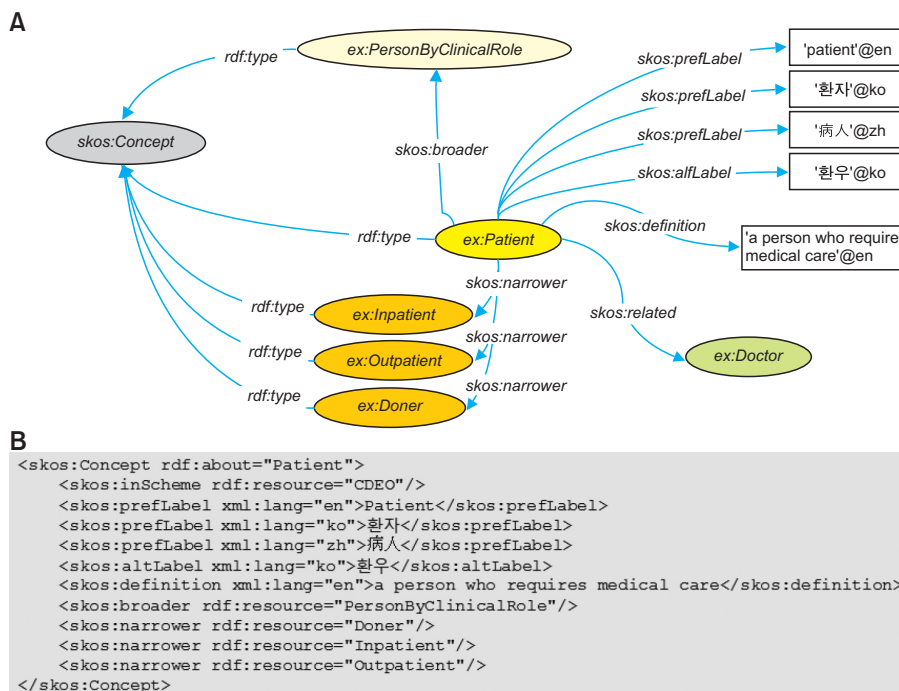


```
<skos:Concept rdf:about="Patient">
    <skos:inScheme rdf:resource="CDEO"/>
    <skos:prefLabel xml:lang="en">Patient</skos:prefLabel>
    <skos:prefLabel xml:lang="ko">환자</skos:prefLabel>
    <skos:prefLabel xml:lang="zh">病人</skos:prefLabel>
    <skos:altLabel xml:lang="ko">환우</skos:altLabel>
    <skos:definition xml:lang="en">a person who requires medical care</skos:definition>
    <skos:broader rdf:resource="PersonByClinicalRole"/>
    <skos:narrower rdf:resource="Doner"/>
    <skos:narrower rdf:resource="Inpatient"/>
    <skos:narrower rdf:resource="Outpatient"/>
</skos:Concept>
```

Figure 2. (A) Resource Description Framework (RDF) graph model and (B) serialization of concepts using Simple Knowledge Organization System (SKOS).

4) Documentation properties (*skos:note, skos:definition, skos:changeNote, skos:scopeNote, skos:editorialNote, skos:historyNote,* and *skos:example*).

Figure 2 illustrates how a concept is organized and represented with the SKOS model. The preferred label, the alternative label, and the definition for the concept '*Patient*' is represented in different languages: 'Patient' in English, '환자' in Korean, and '病人' in Chinese. Its broader (*Person by Clinical Role*), narrower (*Outpatient, Inpatient*), and related (*Doctor*) concepts are represented by the semantic relation properties.

3. Development Strategy

The CDEO was developed using both top-down and bottom-up approaches; that is, harmonization of existing clinical document models was complemented by empirical analysis of the MDRs.

1) Top-down approach

Several clinical document models provided a powerful reference point for the creation of the top concepts for the CDEO. Each model reflects its own perspective on clinical documentation. The Clinical Investigation Record (CIR) ontology reflects the investigator's clinical care delivery process, with the five information types: *Observation, Opinion, Instruction, Action,* and *Administrative event* [25]. The DO focuses on clinical document classification and naming with five axes: *Service, Kind of Narrative, Setting, Role,* and *Subject Matter Domain.* The HL7 Reference Information Model (RIM) represents a static and behavioral model of healthcare workflow with three foundation classes: *Acts, Entities,* and *Roles* [26]. These models are essentially prescriptive, which is somewhat inconsistent with the reality of clinical documentation. Although they provide suitable class levels, these models are barely sufficient to incorporate all concepts included in clinical DEs. Thus, we identified the overlap among the three models, and harmonized their views to define a manageable number of DEC groupings.

2) Bottom-up approach

In searching for empirical evidence for grouping DECs (as at May 23, 2013), we analyzed the occurrence of the concepts in the object classes and the properties in the caDSR (n = 4,414 and n = 5,638, respectively) and the CHMR (n = 2,091 and n = 878, respectively). The most frequently occurring concepts in the object classes in the caDSR included '*procedure*,' '*patient*,' and '*therapy*,' while the most frequently occurring for property were '*name*,' '*duration*,' '*date*,' and '*outcome*.' Many concepts (960 in total) in the object classes also occurred in the properties, which indicate that they probably share a conceptual domain.

A DEC is the combination of an object class and a property. We thus simply adopted this structure to our ontology and added *Qualifier* to limit or specialize the meaning of the concept being represented. The distribution of concepts into the *Object Class* and *Property* is determined by the meaning dependency. For instance, the object class '*person*' is meaningful when it is used alone, whereas the property '*age*' is not. Properties should be associated with one or more object classes to convey the meaning of a DEC. The choice to make groups for *Object Class, Property*, and *Qualifier* appears to be appropriate for our purpose.

We aimed to ensure that the CDEO covers the comprehensive and detailed themes of DECs. Thus, after defining the top concepts, subconcepts were populated from the caDSR—the most comprehensive metadata repository in the biomedical domain— and from existing vocabularies such as SNOMED-CT and MeSH. When a new concept was introduced, for the sake of semantic soundness, a methodological question was asked: *is concept* 'A' *truly a subtype of concept* 'B'?

4. Indexing with the CDEO for Evaluation

The usefulness of the CDEO as a means of categorizing DECs according to some shared high-level characteristics was evaluated. We investigated whether the CDEO produces a 'coherent' index to a given DEC across different indexers who might have different perspectives. One medical record administrator and two PhD students at the Division of Biomedical Informatics indexed DECs using the CDEO without any guidance with respect to choosing terms. They were allowed to assign more than one CDEO term to a given DEC whenever they considered it appropriate. Four hundred DECs were randomly selected from the caDSR. The inter-indexer agreement level was determined using the intraclass correlation coefficient (ICC), which can be used to assess the rating reliability by comparing the variability of different ratings of the same subject to the total variation, and was computed using the Psych package in R 3.0. For each DEC, if one indexer had two shared terms with the other two indexers, who each shared three terms, he obtained a score of 0.67, and the other two indexers obtained a score of 1. This validation process made it possible to evaluate different interpretations of the same DEC as well as the consistency of DEC assignments.

III. Results

1. Structure of the CDEO

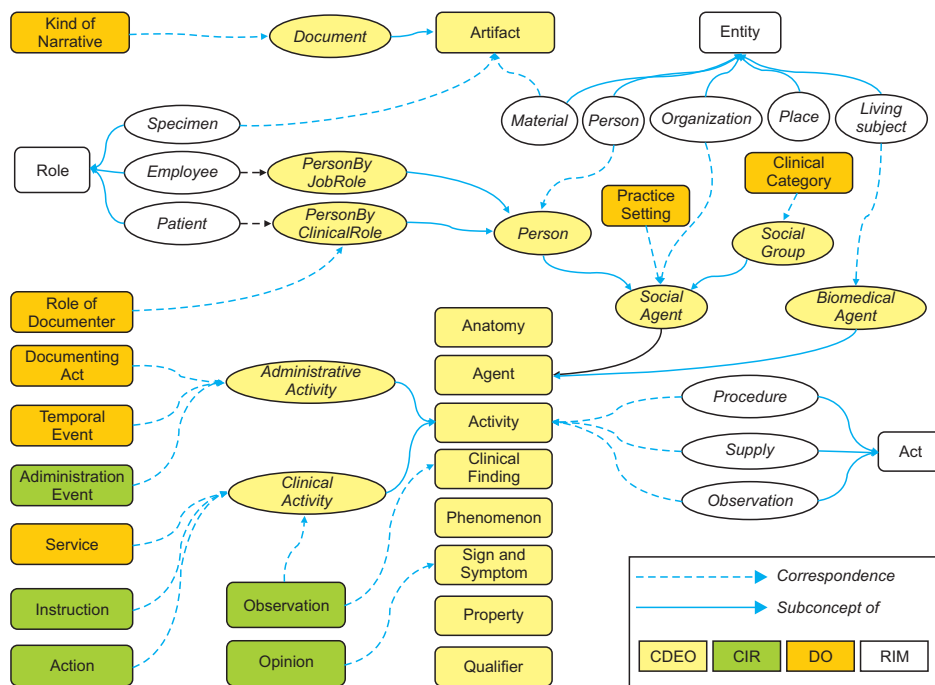The structure of the CDEO arises partly as a compromise

Figure 3. The organization of the Clinical Data Element Ontology (CDEO) was derived from existing clinical document models. CIR: Clinical Investigation Record, DO: Document Ontology, RIM: Reference Information Model.

between existing clinical document models. Figure 3 shows that five of the nine top concepts in the object class correspond to those in the three clinical document models. *Agent* corresponds to the subtypes of *Entity* in the RIM and to *Practice Setting* and *Clinical Category* in the DO via *Social Agent* and *Social Group*. Activity corresponds to Act in the RIM and to *Documenting Act* (i.e., report) and *Temporal Event* (i.e., discharge) in the initial version of the DO; *Clinical Finding* corresponds to *Observation* and *Sign and Symptom* corresponds to *Opinion* in the CIR; and *Artifact*—via *Document*—corresponds to *Kind of Narrative* in the DO. *Anatomy* and *Phenomenon* were introduced because they include many subconcepts that appear frequently in MDRs.

The last two top concepts are *Property* and *Qualifier* (Table 1). *Property* includes *Geospatial* (i.e., *Location*), *Identification* (i.e., *Name*), *Qualitative* (i.e., *Color*), *Quantitative* (i.e., *Frequency*), *Temporal* (i.e., *Date*), and *Unclassified* (i.e., *Impact*) properties. The *Qualifier* specializes or limits the meaning of concepts. CDEO classifies the *Qualifier* into five groups: *Ordinal* (i.e., *First*), *Quantitative* (i.e., *Maximum*), *Sentential* (i.e., *Unique*), *Spatial* (i.e., *Lower*), and *Temporal* (i.e., *Early*).

2. SKOS Representation of the CDEO

Implementation of the CDEO with SKOS resulted in 578 concepts (as at October 3, 2014), including 9 top concepts. The prefix 'cdeo' was declared for the dereferenceable URI (http://www.snubi.org/software/cdeo/), where an MDR application can identify cdeo.rdf. Table 2 lists the namespace prefixes and their URIs used in the CDEO.

3. Use-Case Scenarios

In this section we present two application use-case scenarios to demonstrate the utility of the CDEO. The first scenario is shown in Figure 4A. An indexer loads RDF data from the CDEO and the metadata objects arising from an MDR. It then displays the CDEO concepts in an appropriate style (i.e., a tree view). Given a DEC, the user identifies appropriate CDEO concepts to map. For each object class and property, the concept source (CDEO) and identifier (the URI of the CDEO concept) are in turn inserted back into the MDR (see Table 3). When a user queries against the MDR, the matched metadata objects are presented to the user.

The second scenario is illustrated in Figure 4B. The CDEO is intended for use as a global reference concept system for MDRs, since existed ontology is limited for instance, MeSH is used for biomedical research articles; thus, it does not contain any indexed DECs. Instead, the CDEO is imported into the Index Ontology (IDO), an ontology that simply comprises one class (*ido:DataElementConcept*) and two properties (*ido:hasObjectClass and ido:hasProperty*). Given a DEC originating from one of the listed MDRs, the mapping is conducted and the mapping results are stored in the IDO in RDF format. In a search session, the user navigates concepts in a CDEO-supported browser and coordinates the terms—'*Person*' for object class and '*Age*' for property—to identify the correct DEC: '*Patient age*.' Meanwhile, the browser's backend fetches the URIs arising from the IDO to the structured query language (SQL) query string for MDRs to present the user with the matched metadata objects.

Table 1. Top concepts and major subconcepts in the CDEO

Top concept	Definition	Major subconcept
Agent	Someone or something that affects or changes a situation.	Biomedical agent (chemicals, organism, or person), geospatial agent (city or country), or social agent (organization, person, or social group).
Activity	Any specific behavior including an organic process that takes place in the body.	Administrative activity, clinical activity, communication activity, legal activity, living activity, or physical activity.
Anatomy	Alternative names for the bodily structure of humans, animals, and other living organisms.	Animal structure, body regions, cardiovascular system, or digestive system.
Artifact	A man-made object and something observed in a scientific investigation or experiment that is not naturally present but occurs as a result of the preparative or investigative procedure.	Device, information artifact (information or metadata), document (form, note, or report), specimen, or vehicle.
Clinical finding	Clinical information discovered as the result of medical diagnosis, inquiry, or investigation.	Disease, evaluation finding, laboratory finding, radiology finding, or evaluation finding.
Sign and symptom	Objective and subjective physical or mental feature that is considered indicative of a condition of disease or the existence of an undesirable situation.	Complaint or pain.
Phenomenon	A fact or situation that is observed to exist or happen, especially one whose cause or explanation is questioned.	Death, adverse event, metastasis, or disease.
Property	Any feature that humans naturally use to distinguish one object from another.	Geospatial property, identification property, qualitative property, quantitative property, temporal property, or unclassified property.
Qualifier	Atomic term that could be used to limit or specialize the meaning of a concept being represented.	Ordinal qualifier, quantitative qualifier, sentential qualifier, spatial qualifier, or temporal qualifier.

CDEO: Clinical Data Element Ontology.

Table 2. Namespace prefixes and their URIs used in the CDEO

Prefix	Namespace URI
cdeo	http://www.snubi.org/software/cdeo/
dc	http://purl.org/dc/elements/1.1/
skos	http://www.w3.org/2004/02/skos/core#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
owl	http://www.w3.org/2002/07/owl#

URI: unique resource identifier, CDEO: Clinical Data Element Ontology.

4. Evaluation Relative to the Requirements

To demonstrate the utility of the CDEO, a competency question should be answered: Does the CDEO ever satisfy all of the six requirements considered at the beginning of the study? The answer is definitely yes. The CDEO inherits the merits of the SKOS. It uniquely identifies DECs through the namespace URI. It also controls the labels and language variations of the object class and property through labeling properties and language encoding. Thus, a search for '환자' returns all DECs including 'patient' with their source URIs. The CDEO also provides a poly-hierarchical organization of object classes, properties, and qualifiers, and indexing for each object class and property in the DEC. This enables users to coordinate terms for more precise mapping on the semantic structure of the CDEO.

5. Evaluation Result with Classification Agreement among Indexers

Four hundred DECs were mapped to the CDEO terms, and the ICC among the three indexers was estimated to be 0.59 (95% confidence interval, 0.52–0.66), which is a moderate level of inter-indexer agreement, indicating that the CDEO is need necessary to provide users semantic information or
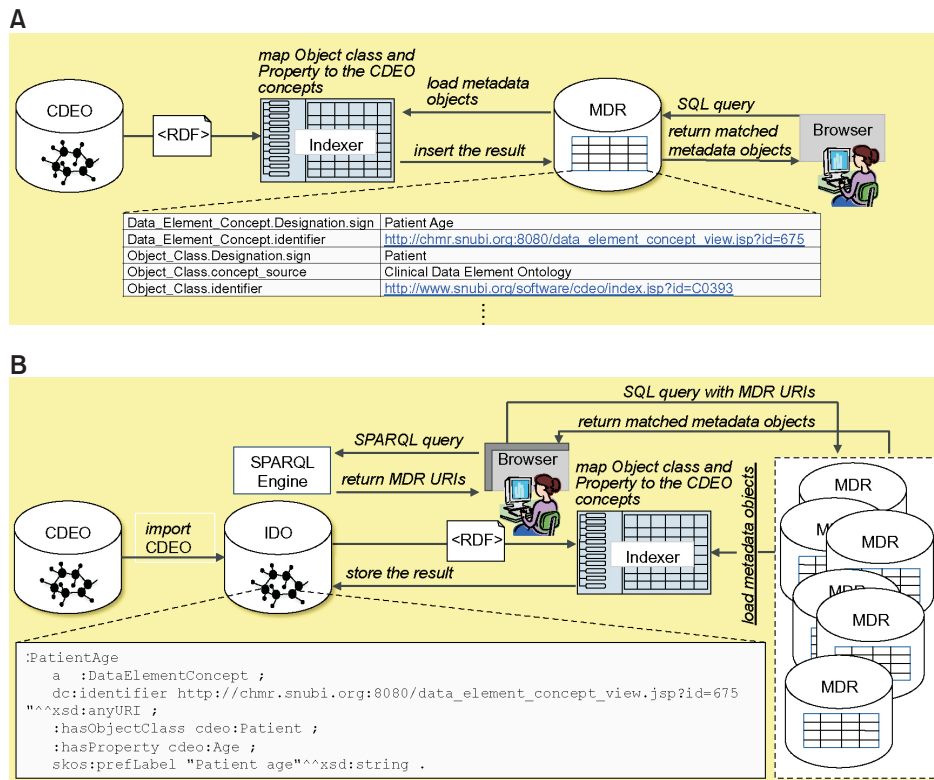
Figure 4. Two use-case scenarios of the Clinical Data Element Ontology (CDEO). (A) User queries against the single metadata registry (MDR), (B) user queries against the multiple MDR by using Index Ontology (IDO). RDF: Resource Description Framework, URI: unique resource identifier.

definition for each CDEO term in order to reduce vogue for indexing or categorizing DECs.

IV. Discussion

This study has developed a novel global reference concept system, called the CDEO, for unified indexing and effective locating of DEs registered across multiple clinical MDRs. This ontology, which was implemented with SKOS, met all of the considered requirements, and its utility was demonstrated through the application of use-case scenarios, which yielded a moderate level of inter-indexer agreement.

Several implications of the CDEO are worth noting. As a global reference concept system, the CDEO organizes DECs, which often arise from different MDRs, into a single conceptual structure. It provides contextual knowledge about the DECs while simultaneously considering the semantic relationships between them. Thus, the CDEO promotes the unified indexing of DECs registered across different MDRs.

In terms of the findability of DEs, the user interface may utilize the structure of the CDEO for highly selective searching and retrieval of relevant DEs. The multilingual representation of DECs also enables users to browse and search DEs in different languages.

We expect that the CDEO will also be useful for clinical documentation. It groups DECs according to semantic structure, and lists all DEs bound to them. Thus, it facilitates the

selection of appropriate DEs for the generation of a specific clinical document form.

The CDEO is also a useful tool for clinical trials and research that depend upon the efficient collection of well-defined and appointed DEs. The definition of data exchange protocols, data quality management, and statistical data analysis models—which are necessary but very time-consuming for the development of DE collections—sustain these DEs and their definitions, value lists, and plausibility checks. The semantic structure of the CDEO reduces the burden of identifying and maintaining relevant DEs, and hence it facilitates the aggregation and synthesis of data arising from different sources.

The present study was subject to some ideological limitations. We do not claim that the CDEO includes an exhaustive list of concepts. However, this ontology is not designed for use in deep indexing; rather, it organizes similarly faceted DECs together. Moreover, postcoordination of broader concepts would cover concepts that are not yet included in the ontology. It could be argued that primitive axes are essentially counterintuitive, yet the model may well contain a flaw. However, the model proposed herein is not the answer, but rather *one of the possible* answers.

Concerning the browsing interface, seemingly counterintuitive deployment can be complemented by a keyword search, which is a quick way to provide users with assistance in locating potentially useful DEs. We would simply claim

Table 3. Mapping table of an object class and a property to their counterparts in the CDEO

Attribute in CDEO	Values and related URI
Data_Element.Designation.sign	Person age value
Data_Element.identifier	http://chmr.snubi.org:8080/data_element_view.jsp?id=20983
Data_Element_Concept.Designation.sign	Patient age
Data_Element_Concept.identifier	http://chmr.snubi.org:8080/data_element_concept_view.jsp?id=675
Object_Class.Designation.sign	Patient
Object_Class.concept_source	CDEO
Object_Class.identifier	http://www.snubi.org/software/cdeo/index.jsp?id=C0393
Object_Class.primaryFlag	Yes
Property.Designation.sign	Age
Property.concept_source	CDEO
Property.identifier	http://www.snubi.org/software/cdeo/index.jsp?id=C0124
Property.primaryFlag	Yes
Concept_System.Designation.sign	CDEO
Concept_System.identifier	http://www.snubi.org/software/cdeo/

CDEO: Clinical Data Element Ontology, URI: unique resource identifier.

that the CDEO is a reference concept system for identifying DECs that can be used to increase the organization and findability of DEs. An unavoidable issue is that within the same reference concept system, indexers may interpret the same DE differently, as indicated by the level of indexer agreement found in this study; this called the indexer effect [11]. This index inconsistency could be minimized by providing a CDEO indexing guide.

The implications and limitations of the CDEO have been identified. Further work includes polishing and refining the CDEO axes so that they provide more granular and enhanced coverage. The addition of language-encoding (i.e., @fr) to the labeling properties and lexical variants is one way of enhancing the utility of this model. The implementation of applications for (semiautomatic) indexing and semantic searches is another possible step. The initial version of CDEO includes only a small number of concepts, leaving room for this ontology to be extended.

We hope that the present work will be a starting point for professional vetting, discussion, and endorsement of the CDEO. We also expect the CDEO to promote the adoption of MDR standards in the clinical domain.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

References

1. International Organization for Standardization. Information technology: Metadata registries (MDR): Part 1. Framework. Geneva, Switzerland: International Organization for Standardization; 2004. ISO/IEC 11179-1:2004(E).

2. Radford MJ, Arnold JM, Bennett SJ, Cinquegrani MP, Cleland JG, Havranek EP, et al. ACC/AHA key data elements and definitions for measuring the clinical management and outcomes of patients with chronic heart failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Heart Failure Clinical Data Standards): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Failure Society of

America. Circulation 2005;112(12):1888-916.

3. Stover PJ, Harlan WR, Hammond JA, Hendershot T, Hamilton CM. PhenX: a toolkit for interdisciplinary genetics research. Curr Opin Lipidol 2010;21(2):136-40.

4. CDISC Inc. CDISC Shared Health and Research Electronic Library pilot report (version 1.0). Austin (TX): CDISC Inc.; 2010.

5. caBIG Strategic Planning Workspace. The Cancer Biomedical Informatics Grid (caBIG): infrastructure and applications for a worldwide research community. Stud Health Technol Inform 2007;129(Pt 1):330-4.

6. Loring DW, Lowenstein DH, Barbaro NM, Fureman BE, Odenkirchen J, Jacobs MP, et al. Common data elements in epilepsy research: development and implementation of the NINDS epilepsy CDE project. Epilepsia 2011;52(6):1186-91.

7. Warzel DB, Andonaydis C, McCurry B, Chilukuri R, Ishmukhamedov S, Covitz P. Common data element (CDE) management and deployment in clinical trials. AMIA Annu Symp Proc 2003;2003:1048.

8. Tobias J, Chilukuri R, Komatsoulis GA, Mohanty S, Sioutos N, Warzel DB, et al. The CAP cancer protocols: a case study of caCORE based data standards implementation to integrate with the Cancer Biomedical Informatics Grid. BMC Med Inform Decis Mak 2006;6:25.

9. Pathak J, Wang J, Kashyap S, Basford M, Li R, Masys DR, et al. Mapping clinical phenotype data elements to standardized metadata repositories and controlled terminologies: the eMERGE Network experience. J Am Med Inform Assoc 2011;18(4):376-86.

10. Papatheodorou I, Crichton C, Morris L, Maccallum P; METABRIC Group, Davies J, et al. A metadata approach for clinical data management in translational genomics studies in breast cancer. BMC Med Genomics 2009;2:66.

11. Miller AC, Odenkirchen J, Duhaime AC, Hicks R. Common data elements for research on traumatic brain injury: pediatric considerations. J Neurotrauma 2012;29(4):634-8.

12. International Organization for Standardization. Information technology: Metadata registries (MDR): Part 2. Classification. Geneva, Switzerland: International Organization for Standardization; 2005. ISO/IEC 11179-2:2005(E).

13. National Cancer Institute. CDE browser [Internet]. Washington (DC): US National Institutes of Health; 2011 [cited at 2014 Oct 1]. Available from: http://cde-browser.nci.nih.gov/CDEBrowser/.

14. Nadkarni PM, Brandt CA. The Common Data Elements for cancer research: remarks on functions and structure. Methods Inf Med 2006;45(6):594-601.

15. Australian Institute of Health and Welfare. METeOR Browse registry [Internet]. Canberra: Australian Institute of Health and Welfare; 2010 [cited at 2014 Oct 1]. Available from: http://meteor.aihw.gov.au/content/index.phtml/itemId/237704.

16. Park YR, Kim JH. Metadata registry and management system based on ISO 11179 for Cancer Clinical Trials Information System. AMIA Annu Symp Proc 2006;2006:1056.

17. Frazier P, Rossi-Mori A, Dolin RH, Alschuler L, Huff SM. The creation of an ontology of clinical document names. Stud Health Technol Inform 2001;84(Pt 1):94-8.

18. Shapiro JS, Bakken S, Hyun S, Melton GB, Schlegel C, Johnson SB. Document ontology: supporting narrative documents in electronic health records. AMIA Annu Symp Proc 2005;2005:684-8.

19. Chen ES, Melton GB, Engelstad ME, Sarkar IN. Standardizing Clinical Document Names Using the HL7/LOINC Document Ontology and LOINC Codes. AMIA Annu Symp Proc 2010;2010:101-5.

20. Kashyap V, Morales A, Hongsermeier T, Li Q. Definitions management: a semantics-based approach for clinical documentation in healthcare delivery. Proceedings of the 4th International Semantic Web Conference (ISWC 2005); 2005 Nov 6-10; Galway, Ireland. p. 887-901.

21. Kim MK, Hwang SH, Kim HG, Kang YK, Choi HC, Kim DS. Classification and conceptualization of clinical documents using formal concept analysis. J Korean Soc Med Inform 2006;12(1):31-43.

22. Song TM, Park HA, Jin DL. Development of health information search engine based on metadata and ontology. Healthc Inform Res 2014;20(2):88-98.

23. W3C. SKOS Simple Knowledge Organization System Reference [Internet]. [place unknown]: World Wide Web Consortium; 2009 [cited at 2014 Oct 1]. Available from: http://www.w3.org/TR/skos-reference/.

24. International Organization for Standardization. Information technology: Metadata registries (MDR): Part 3. Registry metamodel and basic attributes. Geneva, Switzerland: International Organization for Standardization; 2013. ISO/IEC 11179-3:2013(E).

25. Beale T, Heard S. An ontology-based model of clinical information. Stud Health Technol Inform 2007;129(Pt 1):760-4.

26. Health Level Seven International. HL7 version 3: Reference Information Model [Internet]. Ann Arbor (MI): Health Level Seven International; 2004 [cited at 2014 Oct 1]. Available from: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=77.