keyCRF: Using Semantic Metadata Registries to Populate an eCRF with EHR Data

Gokce B. Laleci Erturkmen[1], Landen Bain[2], Anil Sinaci[1]

[1] SRDC Ltd., Ankara, Turkey

[2] CIDISC, USA

{gokce, anil}@srdc.com.tr , lbain@cdisc.org

Abstract. The goal of the keyCRF project is the creation of a semantically annotated electronic Case Report Form (eCRF) that can enable the pre-population of the eCRF from data elements in an EHR summary document through the use of semantically linked Common Data Element definitions across care and research domains
Keywords: semantic metadata registry, re-use of EHR data, eCRF

1   Introduction

A major barrier to repurposing clinical data of electronic health records (EHRs) for clinical research studies (clinical trial design, execution and observational studies) is that information systems in both domains – patient care and clinical research – use different data models and terminology systems. Different data representation standards, and Common Data Element (CDE) models are being used to facilitate seamless data exchange between disparate systems [1].

The Clinical Data Interchange Standards Consortium (CDISC) provides common dataset definitions in (a) Study Data Tabulation Model (SDTM) [2] for enabling the submission of the result data sets of regulated clinical research studies to the FDA and in (b) Clinical Data Acquisition Standards Harmonization (CDASH) [3] for integrating SDTM data requirements into the Case Report Forms. On the care site previously Health Information Technology Standards Panel (HITSP) has previously defined the C154: Data Dictionary Component [4] as a library of data elements. HITSP C32 [5] which describes the HL7/ASTM Continuity of Care Document (CCD) content for the purpose of health information exchange, marks the elements in CCD document with the corresponding HITSP C154 data elements to establish common understanding of the meaning of the CCD elements. Later as a part of Meaningful Use Stage 2, Consolidated CDA (C-CDA) templates are provided for enabling exchange of patient clinical data[6]. The Transitions of Care Initiative (ToC) maintains the S&I Clinical Element Data Dictionary (CEDD) [77] as a repository of data elements in support of meaningful use and improvement in the quality of care. As the data exchange formats and common data elements provided by these two domains are different, it is not automatically possible to pre-fill an electronic case report form annotated with CDISC SDTM and CDASH variables by re-using the medical history of a patient available in a C-CDA document.

The keyCRF project [8] initiated as a part of FDA PhUSE Semantic Technology workgroup, aims to facilitate this, through a metadata registry that maintains the semantic links between the common data elements used in research and care domains and through the use of IHE Data Element Exchange (DEX) [9], and IHE Retrieve Form for Data Capture (RFD)[10] profiles collaboratively.

2 Methods and Expected Results

One of the core activities of keyCRF project is to identify a sample set of CDEs in research sites that are often used to annotate eCRF forms, and the corresponding CDEs at clinical care site, and semantically linking them through a semantically enabled metadata registry implemented in conformance to ISO/IEC 11179 standard. We will be using the semantic MDR implementation provided by SALUS project [1] that enables mapping of CDEs managed by different domains through skos terms such as skos:exactMatch and skos:closeMatch. The activities being carried out can be summarized as:

- Examine the sample eCRF form provided by CDISC and identify the CDEs at research sites from the common CDASH and SDTM annotations of eCRF forms, such as "DM.Sex.Char" to indicate gender code in demographics domain.
- Examine CEDD repository, to find out the corresponding CDEs to the selected research CDEs, for example "PatientInformation.PatientAdministrativeGender.CE" to "DM.Sex.Char""
- Represent all these CDEs in a 11179 supporting MDR, by also semantically linking them with skos terms
- Define extraction specifications of the selected CEDD CDEs from C-CDA as XPATHS

By making use of these definitions available from a semantic MDR that also supports IHE DEX as a standard means to retrieve metadata of CDEs, we demonstrate that it is possible to pre-fill an electronic case report form by re-using the medical history available as follows: A research forms designer becomes able to build a case report form for a particular research study by referring to an on-line metadata registry of research data elements, and selects the desired data elements from a set of research friendly elements such as CDASH. He then retrieves the metadata defined by the metadata registry into an annotated case report form through the use of IHE DEX profile. The metadata includes the exact specification, using XPath, to find the corresponding data element in the C-CDA. The semantic MDR creates the metadata by checking the semantic links of CDASH data elements to CEDD data elements, which already have mappings to C-CDA documents. Using the XPath statements, the research system creates an extraction specification for all elements to be extracted from the C-CDA. The demonstration will employ the well-known mechanism of IHE RFD to define the necessary transactions between the EHR and the research system. The extraction specification could then be used with IHE RFD to pre-populate the case report form.

This prototype demonstration will show industry the value of the semantic approach to address the challenge of secondary use of EHR for research purposes.

References

1. Sinaci A.A., Laleci Erturkmen G.B, A federated semantic metadata registry framework for enabling interoperability across clinical research and care domains. J Biomed Inform. 2013 Oct;46(5):784-94
2. CDISC. Study Data Tabulation Model (SDTM), http://www.cdisc.org/sdtm
3. CDISC. Clinical Data Acquisition Standards Harmonization (CDASH),http://www.cdisc.org/cdash
4. HITSP. C 154: HITSP Data Dictionary, http://www.hitsp.org/ConstructSet_Details.aspx?&PrefixAlpha=4&PrefixNumeric=154
5. HITSP. C 32: HITSP Summary Documents Using HL7 Continuity of Care Document (CCD) Component, http://www.hitsp.org/ConstructSet_Details.aspx?&PrefixAlpha=4&PrefixNumeric=32
6. HL7 Implementation Guide for CDA® Release 2: IHE Health Story Consolidation, Release 1.1 - US Realm, http://www.hl7.org/implement/standards/product_brief.cfm?product_id=258

7. S&I Framework. S&I Clinical Element Data Dictionary (CEDD) WG, http://wiki.siframework.org/S%26I+Clinical+Element+Data+Dictionary+WG

8. EHR Enabled Research, http://www.phusewiki.org/wiki/index.php?title=EHR_Enabled_Research

9. IHE Data Exchange (DEX) Profile, http://www.ihe.net/Technical_Framework/upload/IHE_QRPH_Suppl_DEX_Rev1-0_PC_2013-06-03.pdf

10. IHE Retrieve Form for Data Capture Profile, http://www.ihe.net/Technical_Framework/upload/IHE_ITI_Suppl_RFD_Rev2-2_TI_2011-08-19.pdf