

Semantic Enrichment of CDISC Operational Data Model

Sarah N. Lim Choi Keung, PhD¹, Lei Zhao, MSc¹ Jean-François Ethier, MD², Vasa Curcin, PhD³, Anita Burgun, MD, PhD², Brendan C. Delaney, MD³, Theodoros N. Arvanitis, DPhil¹

¹Institute of Digital Healthcare, WMG, University of Warwick, UK; ²INSERM UMR 1138, Université Paris Descartes - Sorbonne Paris Cité, France; ³Department of Primary Care and Public Health Sciences, King's College London, UK

Abstract

The semantic interoperability between electronic health records and clinical research systems is lacking. The CDISC Operational Data Model standard for the interchange of clinical research data has been extended to establish common semantics between the two domains, using archetypes and mappings to an ontology.

Introduction and Background

The increasing presence of electronic health records (EHR) at healthcare sites opens new opportunities to integrate the processes of clinical care and clinical research. Of particular interest is enabling clinical data exchange between the various systems, e.g. between an EHR system and a clinical trial management systems (CTMSs). CDISC Operational Data Model (ODM) is a vendor neutral, platform independent format which facilitates the archive and interchange of the metadata and data for clinical research. Despite the wide support of ODM in modern CTMSs, the absence of semantic definition of the underlying clinical data elements limits their interoperability with the EHR systems. Additionally, the sharing of research data forms in a meaningful way has been limited. This paper outlines the archetype approach to improved links between the two domains by semantically enriching CDISC ODM.

Methods

The semantic extension to ODM uses an openEHR archetype-based approach to semantic interoperability. This is a two-level modelling paradigm, adopted in the ISO 13606 standard. The higher level modelling defines a generic reference model to represent the basic information structure. The lower level applies archetypes which are constraint models to the reference model to define detailed clinical contents. By associating archetypes with ODM form elements such as *ItemGroupDef* and *ItemDef*, the ODM form is enhanced with semantic definitions of clinical data elements, where each *ItemGroupDef* represents a clinical data element instance (archetype) and each *ItemDef* represents an attribute of the data element. Each attribute, in turn refers to an ontology item. As a computable representation, the archetype enables to maintain the semantic meaning of clinical data elements within the ODM format, for the data collection and exchange of eCRFs. The archetype binding to an ontology ensures that the semantic link is preserved.

Results and Discussion

In the EU FP7 TRANSFoRm project, the archetype binding uses the Clinical Data Integration Model, an ontology of primary care clinical domain. TRANSFoRm added extensions to the ODM, including elements such as *CdimConcept*, which refers to specific CDIM ontology concepts, *transform:CdimBinding*, which links form items to clinical data elements. This semantic extension of the ODM enables linkage to EHR data, e.g. for eCRF pre-population. In addition, research data forms can be shared with additional semantic information about the form elements. In TRANSFoRm, the pre-population of eCRFs from EHR is currently being evaluated in a pan-European clinical trial examining the effectiveness of continuous vs. on-demand PPI prescriptions in treatment of GERD.

Acknowledgements

The TRANSFoRm project is partially funded by the European Commission under the 7th Framework Programme (Grant Agreement 247787).