A Semantic Framework for Quality Data Model to Support EHR-driven Phenotype Authoring And Execution

Guoqian Jiang, MD, PhD[1], Harold R. Solbrig[1], Richard Kiefer[1], Luke V. Rasmussen[2]; Huan Mo, MD, MS[3], Peter Speltz[3], William K. Thompson, PhD[2], Joshua C. Denny, MD, MS[3], Christopher G. Chute, MD, Dr. PH[1], Jyotishman Pathak, PhD[1]
[1]Mayo Clinic College of Medicine, Rochester, MN; [2]Northwestern University, Chicago, IL; [3]Vanderbilt University, Nashville, TN

Abstract
In this study, we describe our efforts in developing a semantic framework for representing the Quality Data Model (QDM) to support phenotype authoring and execution. We discuss the modeling challenges and potentials of the framework that could not only provide a semantic meta-data repository and data elements services, but also enable a standard-based mechanism that could incorporate data elements from external clinical data and information models, such as HL7 FHIR.

Introduction
Quality Data Model (QDM) is an information model developed by the National Quality Forum (NQF) for representing electronic health record (EHR)-based quality "eMeasures". Previous studies (1-2) have demonstrated that QDM with HL7 Health Quality Measures Format (HQMF) contains core elements that make it a promising candidate for representing EHR-driven phenotyping algorithms for clinical research. However, currently the QDM specification is available only as PDF file(s), which is not optimal for broad use and machine consumption. We consider that a standards-based, semantically annotated, machine-readable rendering of the QDM would be critical in support of development of phenotype authoring and execution applications.

System Architecture
We developed a semantic framework that comprises three layers: 1) a semantic data element repository layer; 2) a semantic services layer; and 3) a phenotype application layer. Figure 1 shows the system architecture of the framework. In the repository layer, we leverage both W3C standards (3), such as Resource Description Framework (RDF) and Web Ontology Language (OWL), and a meta-data standard known as ISO 11179 (4). Specifically, we describe the QDM reference model, data model elements and logic elements using the ISO 11179 standard and represent them using RDF and OWL. We developed a standard interface between the data repository and the Clinical Information Modeling Initiative (CIMI) (5) archetype modeling language (AML)/archetype definition Language (ADL), which enabled the system to load the data elements from externally developed clinical data and information models (such as HL7 FHIR (6), CIMI (5), Intermountain Healthcare CEM (7) and OpenEHR (8)). In the semantic service layer, we developed Semantic Web services on the top of our meta-data repository using Linked Data API (9). In the application layer, the phenotype authoring and execution applications are designed to consume the data element services.

Prototype Implementation
We first manually developed a QDM schema in OWL representing the QDM reference model with the notions of QDM Data Model Element (including Category, Datatype, Attribute, Valueset, etc.) and QDM Logic Element (including Logic Operator, Function, Comparison Operator, Temporal Operator, etc). The schema is designed as a natural extension of the ISO 11179 standard. We converted the data elements specified in the QDM model version 4.1 from a NQF specification document into an Excel spreadsheet. We then wrote a program to populate the schema with the model element data as QDM schema instances. In total, 18 instances of Category, 76 instances of Datatype, 528 instances of Attribute and 53 instances of Logic Element were populated. Each QDM element is also asserted as an instance of ISO 11179 Data Element. Using the ISO 11179 standard, we were able to identify a number of underspecified areas in QDM. For example, the meta-data "data element type" is not specified in QDM, which may potentially cause arbitrary interpretation of a data element when it is used.

Conclusion
Our framework provides a standards-based semantic infrastructure in enabling QDM data element services to support phenotype authoring and execution. We are in the process of implementing a RDF store-based repository and establishing Semantic Web RESTful Services using Linked Data API on the top of the repository. In future work, we plan to develop a standard interface mechanism with HL7 and CIMI-compliant clinical models.

Acknowledgements
This work has been supported in part by funding from PhEMA (R01 GM105688), eMERGE (U01 HG006379, U01 HG006378 and U01 HG006388), and caCDE-QA (1U01CA180940-01A1).

References
1. Li D, Endle CM, Murthy S, Stancl C, Suesse D, Sottara D, Huff SM, Chute CG, Pathak J.Modeling and executing electronic health records driven phenotyping algorithms using the NQF Quality Data Model and JBoss® Drools Engine. AMIA Annu Symp Proc. 2012;2012:532-41. Epub 2012 Nov 3.
2. Thompson WK, Rasmussen LV, Pacheco JA, Peissig PL, Denny JC, Kho AN, Miller A, Pathak J. An evaluation of the NQF Quality Data Model for representing Electronic Health Record driven phenotyping algorithms.AMIA Annu Symp Proc. 2012;2012:911-20. Epub 2012 Nov 3.
3. W3C Standards. http://www.w3.org/standards/; Last visited at September 23, 2014.
4. ISO/IEC 11179, Information Technology -- Metadata registries (MDR). http://metadata-standards.org/11179/; Last visited at September 23, 2014.
5. Clinical Information Modeling Initiative. http://informatics.mayo.edu/CIMI/index.php/Main_Page; Last visited at September 23, 2014.
6. HL7 FHIR. http://www.hl7.org/implement/standards/fhir/; Last visited at September 23, 2014.
7. CEM Browser. http://www.clinicalelement.com/#/; Last visited at September 23, 2014.
8. openEHR. http://www.openehr.org/; Last visited at September 23, 2014.
9. Linked Data API. https://code.google.com/p/linked-data-api/; Last visited at September 23, 2014.

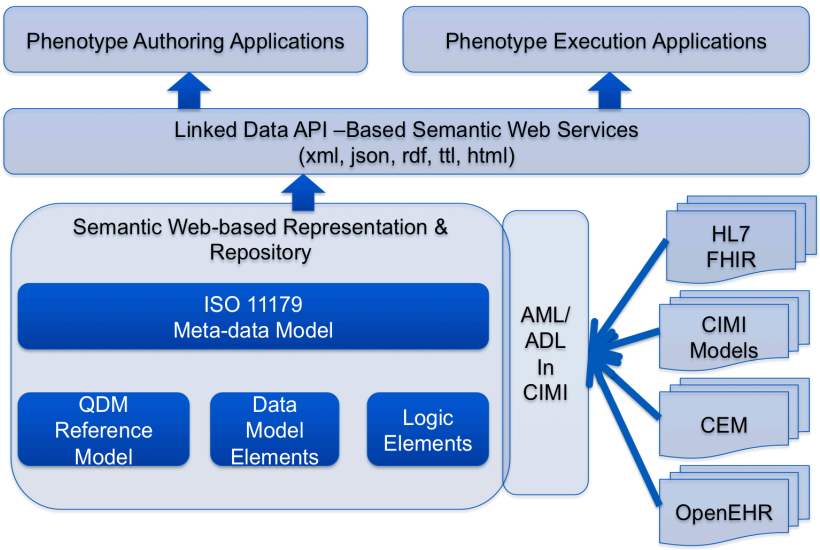

Figure 1. System architecture.