



Modeling and Validating FHIR Profiles Using Semantic Web Shape Expressions (ShEx)

Harold R. Solbrig, MS¹, Eric Prud'hommeaux², Deepak K. Sharma, MS¹

Christopher G. ChuteMD, Dr. PH³, Guoqian Jiang, MD, PhD¹

¹ Department of Health Sciences Research, Mayo Clinic, Rochester, MN; ² W3C/MIT, Boston, MA;

³ Johns Hopkins University, Baltimore, MD

Abstract

Introduction: HL7 Fast Healthcare Interoperability Resources (FHIR) is an emerging standard for the electronic exchange of healthcare information. FHIR defines a collection of "resources" that "can easily be assembled into working systems." While a robust software ecosystem is emerging from FHIR community, the clinical community still faces a substantial challenge of transformation between existing databases and the FHIR models. The current tooling for defining FHIR profiles involves a publication process for producing machine-enforceable (XML) schemas, complicating the workflow for iterative design of FHIR profiles.

The Resource Description Framework (RDF) Shape Expressions language (ShEx), developed as an input to the W3C RDF Data Shapes working group, is a mechanism to formally describe RDF structures, providing the RDF analog to XML schema.

Objective: We present the ShEx-based tools for modeling and validating HL7 FHIR profiles.

Methods and Results: In this study, we used the core resources in both the the HL7 FHIR DSTU2 Ballot released on April 2, 2015 and the nightly builds. The core resources are classified into clinical, administrative, infra-structure and financial categories. Both the FHIR resources and the schemas that describe have equivalent XML and JSON representations.

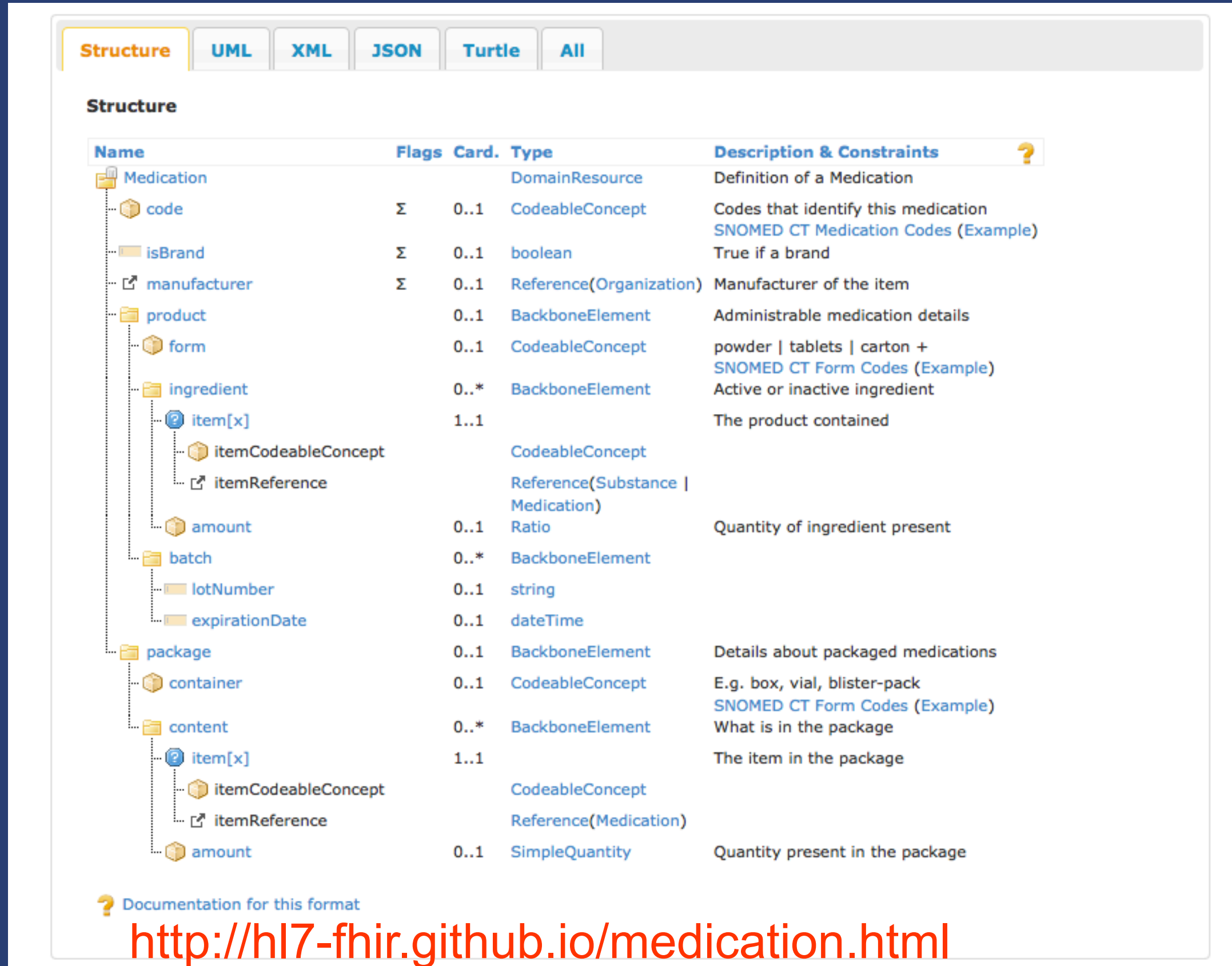
The FHIR schema to ShEx (FHIRSchema2ShEx)⁶ transformation tools were developed in collaboration with the HL7 ITS/W3C HCLS group, based on the their mappings between the JSON representation of the FHIR schemas and their ShEx equivalent. We implemented these transformation rules in JavaScript and integrated them with the FHIRData2RDF application.

We began by creating a library of ShEx schemas, one per core FHIR resource. The library is generated by a tool that extracts a FHIR ShEx schema name from the FHIR2ShEx transformation and produces a file that contains the target FHIR schema and its dependencies.

We also discuss our plans to integrate this process into the Open Refine platform (D2Refine) to provide a user-friendly interface to support RDF/FHIR data element harmonization and model transformation.

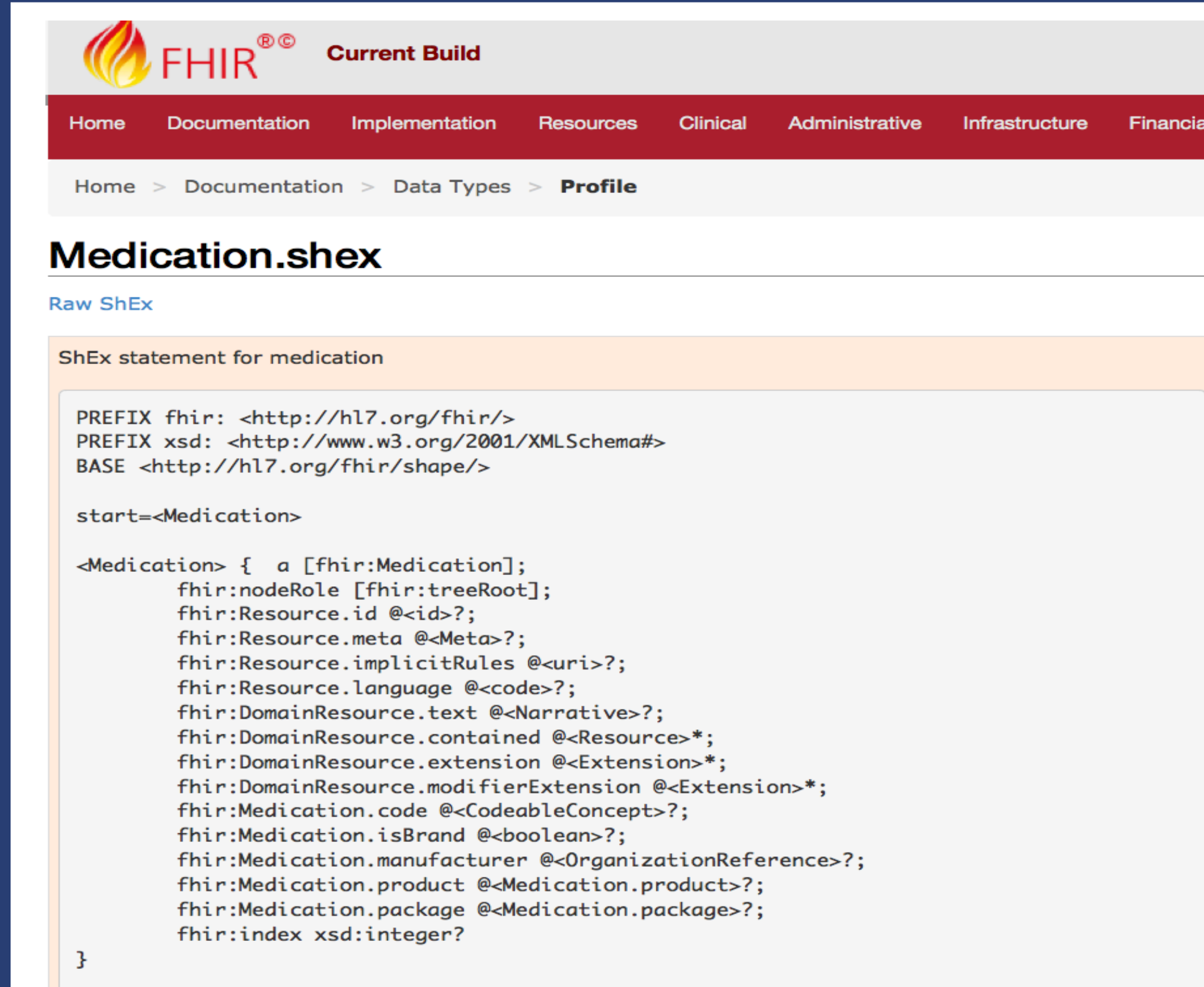
Conclusion: In this study, we described our ongoing efforts in developing tools that transforms HL7 FHIR core models into ShEx-based schema. We demonstrated that ShEx is a useful mechanism to model and test conformance with FHIR profiles.

FHIR Logical Model



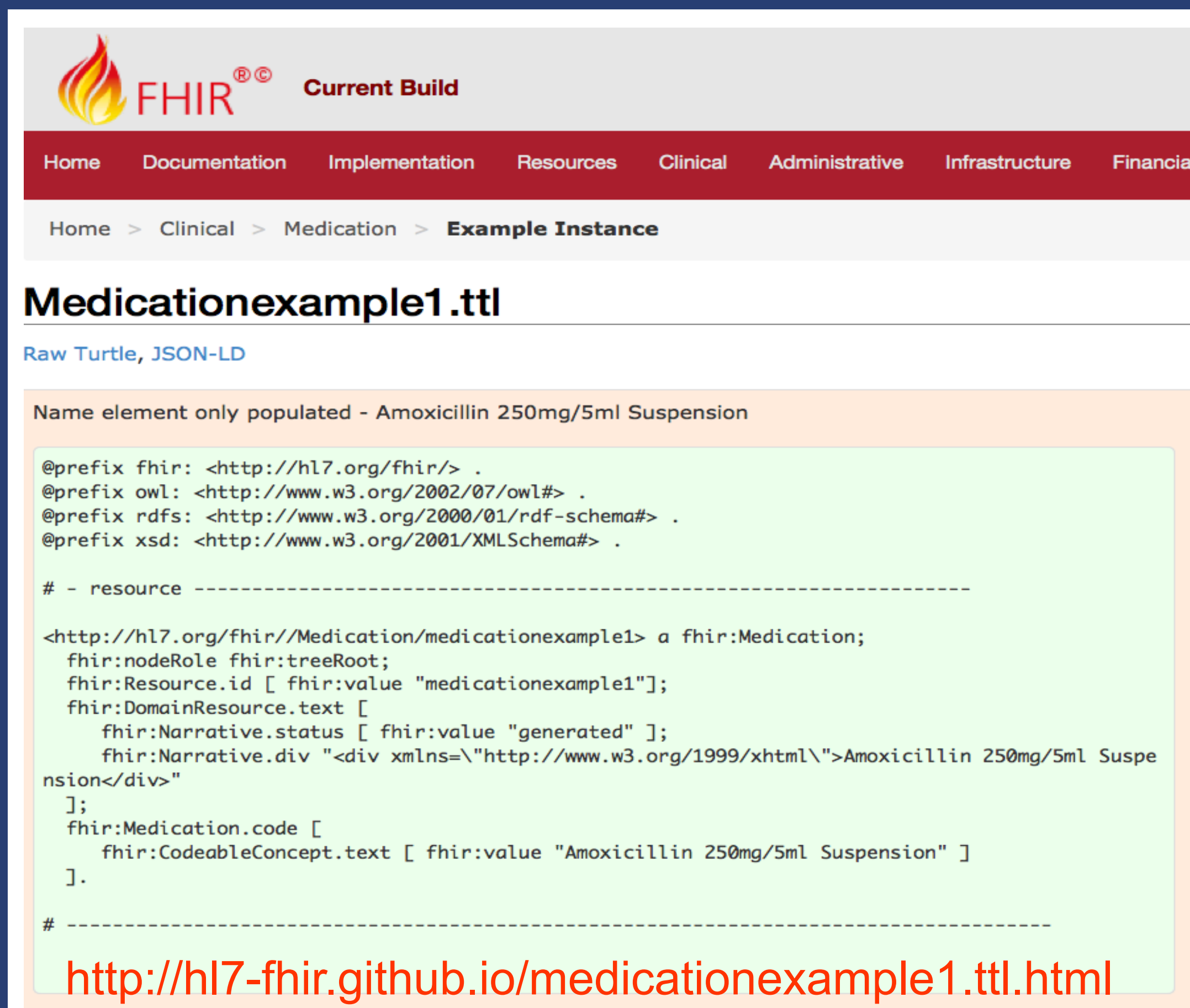
<http://hl7-fhir.github.io/medication.html>

Logical Model in ShEx



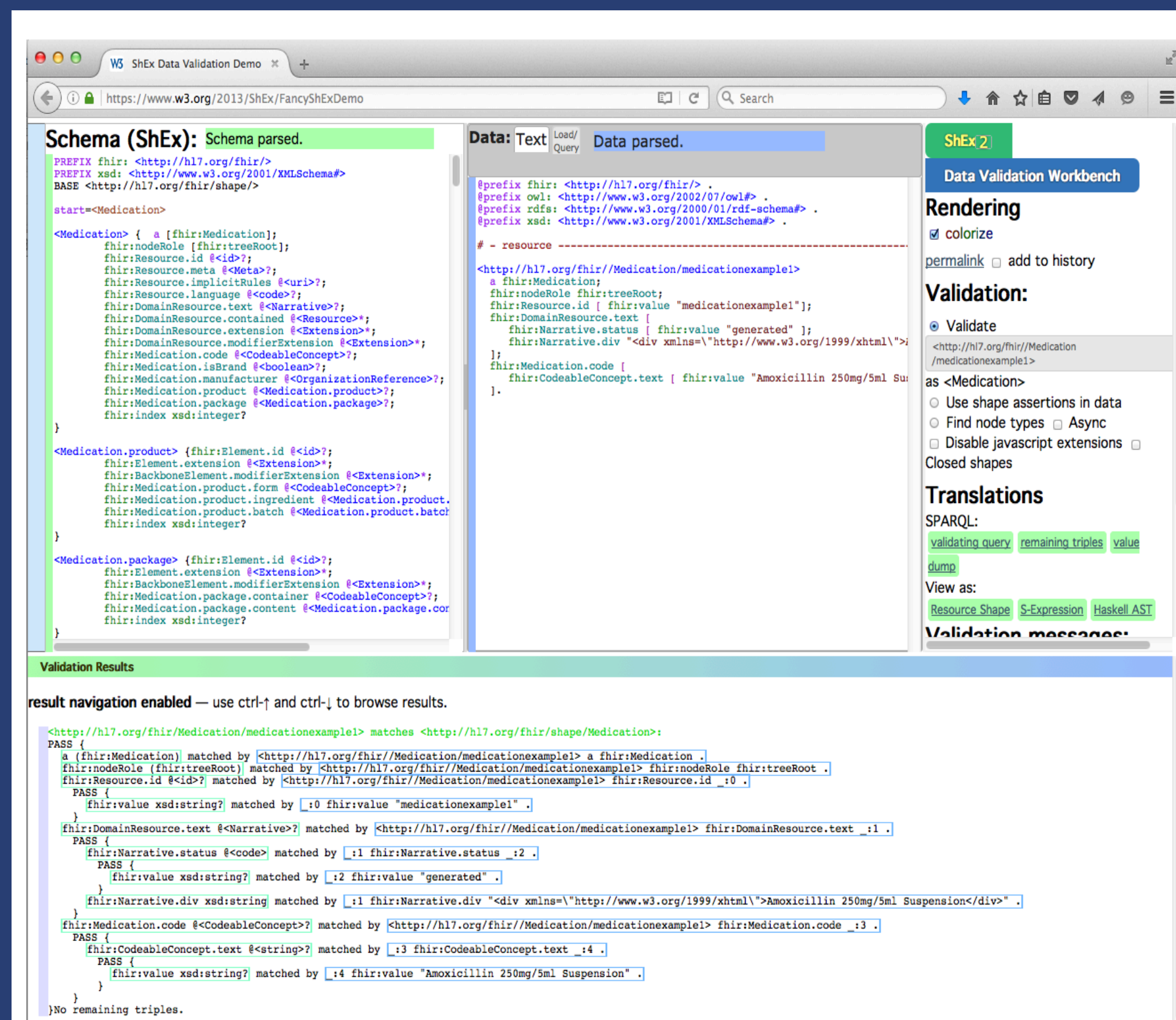
<http://hl7-fhir.github.io/medication.shex.html>

FHIR Data Objects



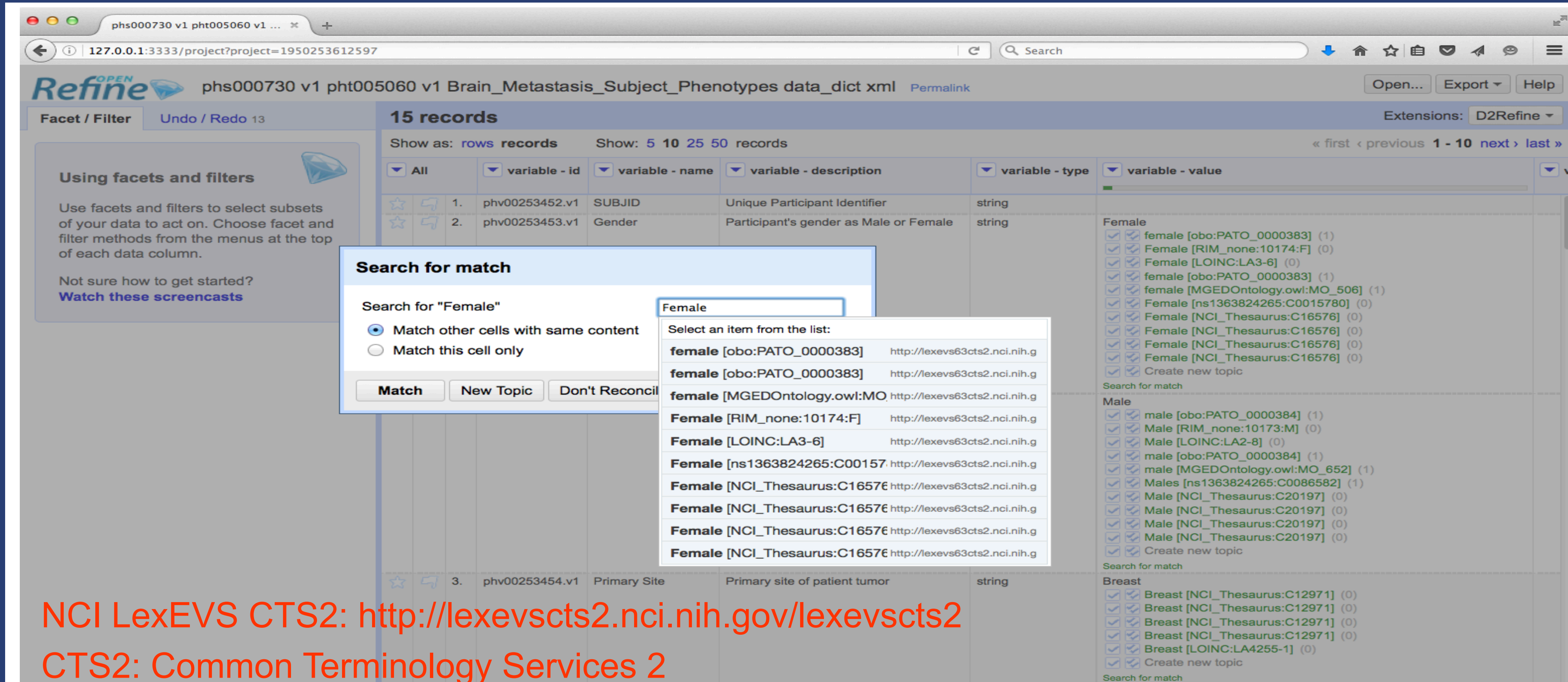
<http://hl7-fhir.github.io/medicationexample1.ttl.html>

ShEx-based Validation Tool



<https://www.w3.org/2013/ShEx/FancyShExDemo>

D2Refine CTS2-based Reconciliation Plugin



NCI LexEVS CTS2: <http://lexevscts2.nci.nih.gov/lexevscts2>
CTS2: Common Terminology Services 2

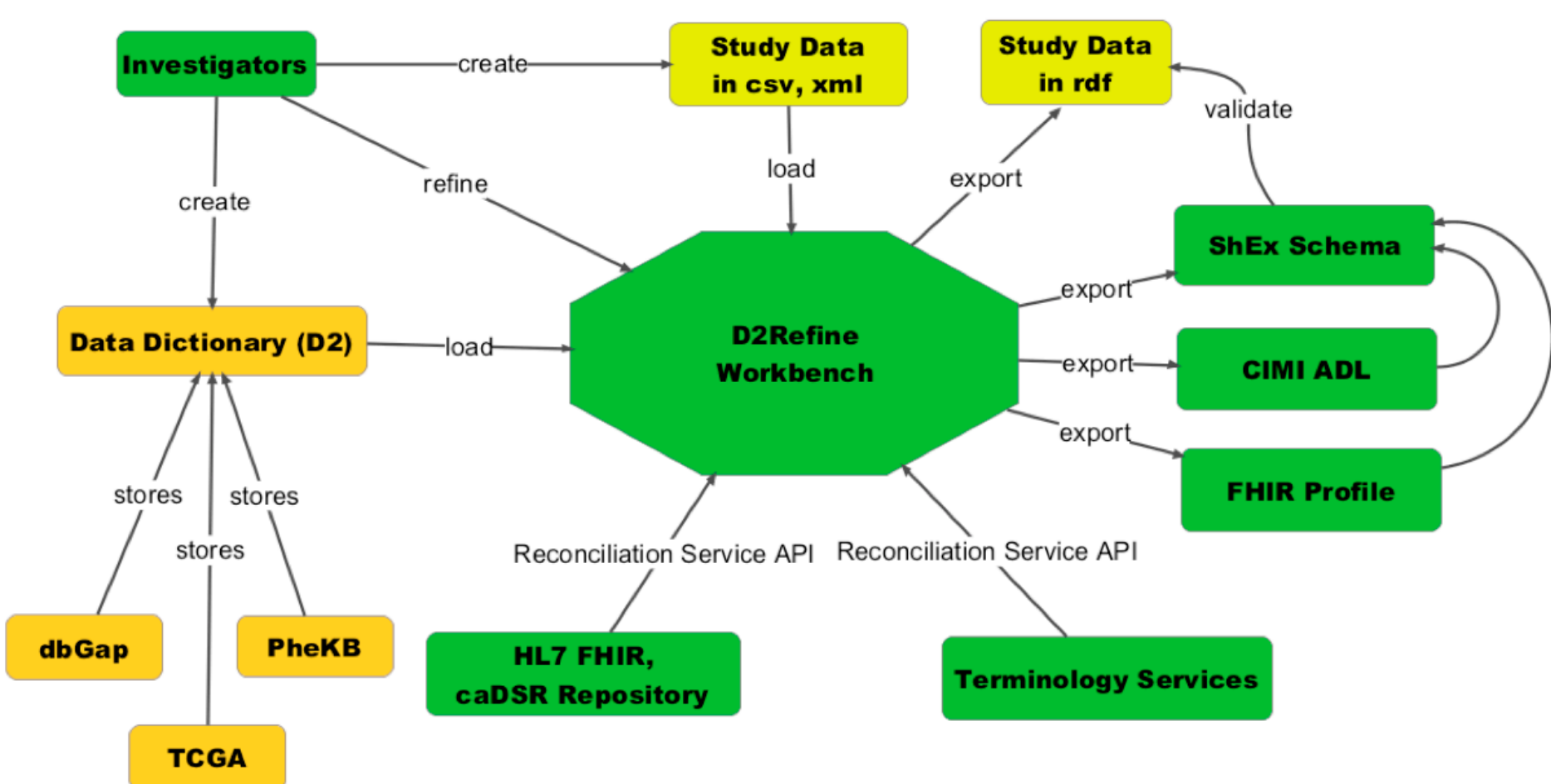
Discussion

1. There is an emerging need to represent the cancer phenotype models using standard FHIR modeling mechanism, which can be validated using our ShEx tools.
2. D2Refine serves as an integrated metadata harmonization and validation workbench with a user-friendly interface, which can capture the feedback from the cancer research investigators while automate the data quality check and validation.

References

1. Project Wiki;
 - <http://informatics.mayo.edu/caCDE-QA/index.php>
2. GitHub Code Repositories
 - <https://github.com/caCDE-QA/ShEx>
 - <https://github.com/caCDE-QA/D2Refine>

D2Refine Framework



Acknowledgements

This work has been supported in part by funding from a NCI U01 Project known as caCDE-QA (U01 CA180940) .