

CALCULATING HEDIS MEASURES USING KNOWLEDGE REPRESENTATION TECHNIQUES ON HL7 RIM

Peter Hendler[†] Robert Piro[‡] Michael Rossman[†]

[†])Kaiser Permanente and the [‡])University of Oxford

”SNOMED CT: making health records make sense”
IHTSDO, 31 October, 2014



OVERVIEW



- 1 OVERVIEW
- 2 MOTIVATION
- 3 HEDIS MEASURES FOR DIABETIC CARE
- 4 OUR APPROACH
 - Data Structure and Encoding
 - OWL 2 RL / Datalog Ontology
- 5 RDFox
- 6 CONCLUSION

ONTOLOGY

- Logical theory
 - OWL 2 EL statements
 - OWL 2 RL statements/Datalog rules
- Encodes background information
 - clear
 - succinct
 - formal
- Popular in Bio-Chemistry and Medicine (SNOMED)

EXTEND ONTOLOGY APPLICATION TO HEALTH CARE

- HL7 RIM Healthcare data format
- Ontologies allow
 - enhanced querying
 - computer aided tool support
- Proof of concept: HEDIS Measures

HEDIS FOR DIABETIC CARE

- NCQA (USA) Performance Measures
- Entry requirements for Medicare
- Consumer comparison benchmark
- Measures: $\frac{\# \text{Eligible with eye exam}}{\# \text{Eligible}(=\text{Diabetic}, \geq 65\text{yo}, \text{etc})}$

HEDIS FOR DIABETIC CARE

- NCQA (USA) Performance Measures
- Entry requirements for Medicare
- Consumer comparison benchmark
- Measures: $\frac{\# \text{Eligible with eye exam}}{\# \text{Eligible}(=\text{Diabetic}, \geq 65\text{yo}, \text{etc})}$

HEDIS MEASURE COMPUTATION IS DIFFICULT

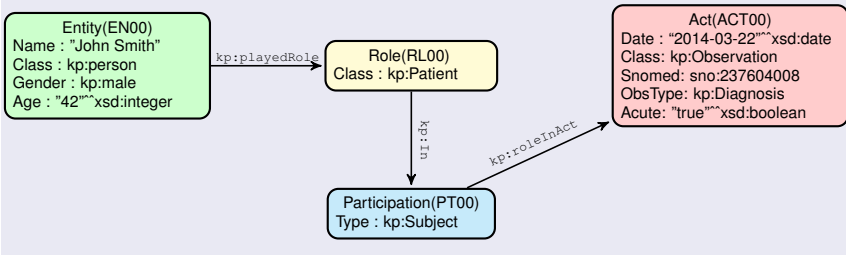
- Disparate data sources (historically grown)
- Ad-hoc schemas used to store data (meaning implicit)
- Involved programs for analytics software
 - mix data (re)formatting and measuring
 - difficult to maintain
 - require high expertise of IT-experts

APPROACH: SEPARATE DATA UNIFICATION FROM MEASURING

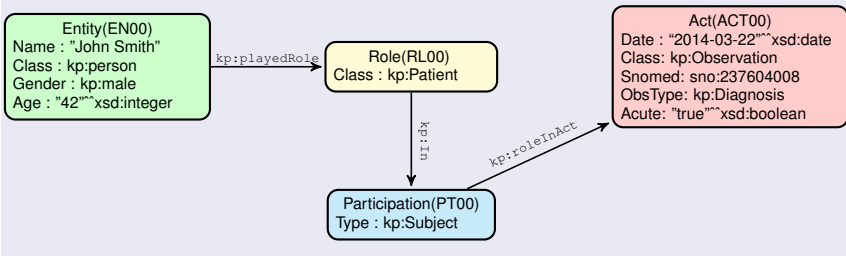
APPROACH: SEPARATE DATA UNIFICATION FROM MEASURING



DATA: HL7 RESOURCE INFORMATION MODEL (RIM)



DATA: HL7 RESOURCE INFORMATION MODEL (RIM)

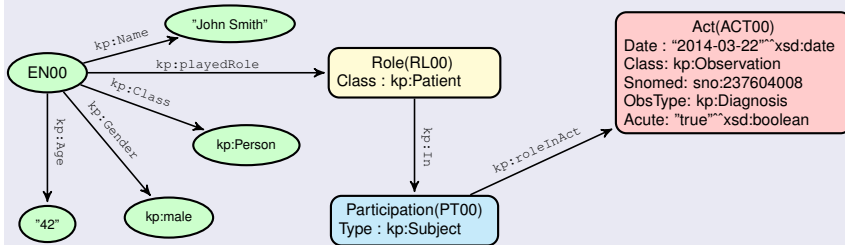


ENCODING IN RDF-TRIPLES

EN00 kp:Age "42"^^ xsd:integer .	RL00 kp:In	PT00 .
EN00 kp:playedRole _:RL00 .	PT00 kp:roleInAct	ACT00 .
RL00 kp:Class kp:Patient .	ACT00 kp:SNOMED	sno:237604008 .

APPROACH: SEPARATE DATA UNIFICATION FROM MEASURING

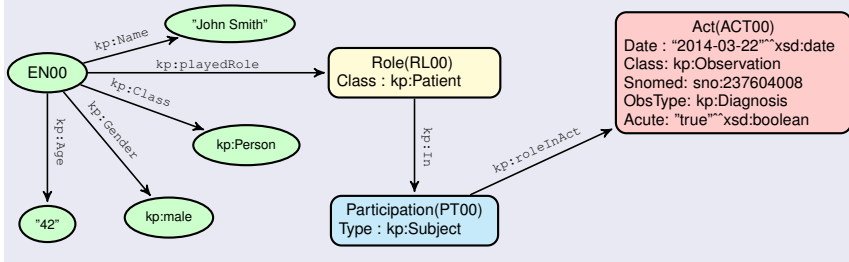
DATA: HL7 RESOURCE INFORMATION MODEL (RIM)



ENCODING IN RDF-TRIPLES

EN00 kp:Age "42"^^ xsd:integer .	RL00 kp:In PT00 .
EN00 kp:playedRole _:RL00 .	PT00 kp:roleInAct ACT00 .
RL00 kp:Class kp:Patient .	ACT00 kp:SNOMED sno:237604008 .

DATA: HL7 RESOURCE INFORMATION MODEL (RIM)



- HL7 RIM standard data model in healthcare. Not ad-hoc
- Structures data in Entity-Role-Participation-Act categories.
- Directly translatable into RDF-triples
 - Universal data format used in the Semantic Web
 - Represents data as subject predicate object
 - Human readable. Third party processible.

DATALOG RULE

$$H \leftarrow B_1, \dots, B_n$$

$$\langle ?x, \text{kp:hasDiagnosis}, ?y \rangle \leftarrow \langle ?x, \text{kp:playedRole}, ?r \rangle, \langle ?r, \text{kp:In}, ?p \rangle$$

$$\langle ?p, \text{kp:roleInAct}, ?a \rangle, \langle ?a, \text{kp:SNOMED}, \text{sno:237604008} \rangle$$

RULE EVALUATION

$$\langle \text{EN00}, \text{kp:hasDiagnosis}, \text{ACT00} \rangle \leftarrow \langle \text{EN00}, \text{kp:playedRole}, \text{RL00} \rangle, \langle \text{PT00}, \text{kp:roleInAct}, \text{ACT00} \rangle$$

$$\langle \text{RL00}, \text{kp:In}, \text{PT00} \rangle, \langle \text{ACT00}, \text{kp:SNOMED}, \text{sno:237604008} \rangle$$

ENCODING IN RDF-TRIPLES

EN00 kp:Age "42"^^	xsd:integer .	RL00 kp:In	PT00 .
EN00 kp:playedRole	_:RL00 .	PT00 kp:roleInAct	ACT00 .
RL00 kp:Class	kp:Patient .	ACT00 kp:SNOMED	sno:237604008 .

DATALOG RULE

$$H \leftarrow B_1, \dots, B_n$$

$$\langle ?x, \text{kp:hasDiagnosis}, ?y \rangle \leftarrow \langle ?x, \text{kp:playedRole}, ?r \rangle, \langle ?r, \text{kp:In}, ?p \rangle \\ \langle ?p, \text{kp:roleInAct}, ?a \rangle, \langle ?a, \text{kp:SNOMED}, \text{sno:237604008} \rangle$$

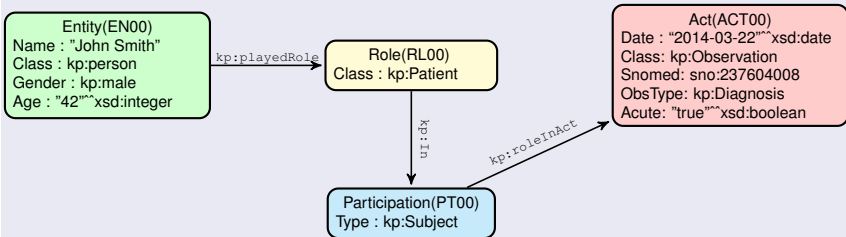
RULE EVALUATION

$$\langle \text{EN00}, \text{kp:hasDiagnosis}, \text{ACT00} \rangle \leftarrow \langle \text{EN00}, \text{kp:playedRole}, \text{RL00} \rangle, \langle \text{PT00}, \text{kp:roleInAct}, \text{ACT00} \rangle \\ \langle \text{RL00}, \text{kp:In}, \text{PT00} \rangle, \langle \text{ACT00}, \text{kp:SNOMED}, \text{sno:237604008} \rangle$$

ENCODING IN RDF-TRIPLES

EN00 kp:Age "42"^^	xsd:integer .	RL00 kp:In	PT00 .
EN00 kp:playedRole	_:RL00 .	PT00 kp:roleInAct	ACT00 .
RL00 kp:Class	kp:Patient .	ACT00 kp:SNOMED	sno:237604008 .
		EN00 kp:hasDiagnosis	ACT00 .

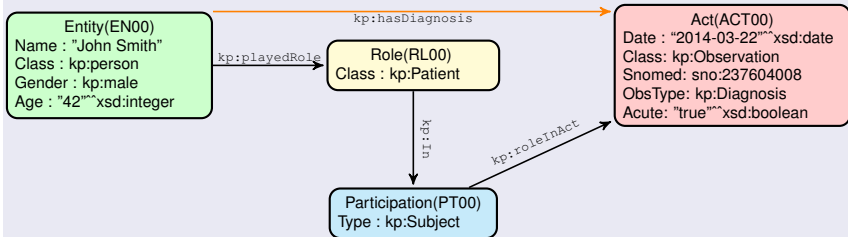
LABELLING



$$\langle ?e, kp:hasDiagnosis, ?a \rangle \leftarrow \langle ?e, kp:playedRole, ?r \rangle, \langle ?r, kp:In, ?p \rangle$$

$$\langle ?p, kp:roleInAct, ?a \rangle.$$

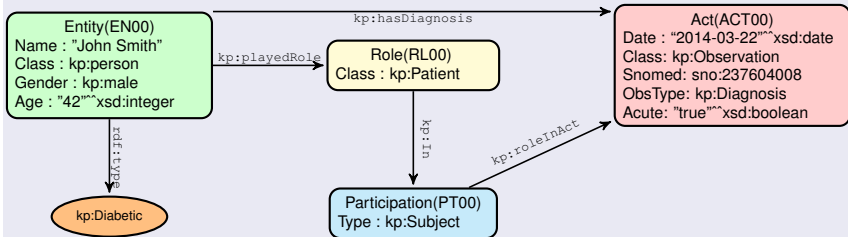
LABELLING



$$\langle ?e, kp:hasDiagnosis, ?a \rangle \leftarrow \langle ?e, kp:playedRole, ?r \rangle, \langle ?r, kp:In, ?p \rangle$$

$$\langle ?p, kp:roleInAct, ?a \rangle.$$

LABELLING

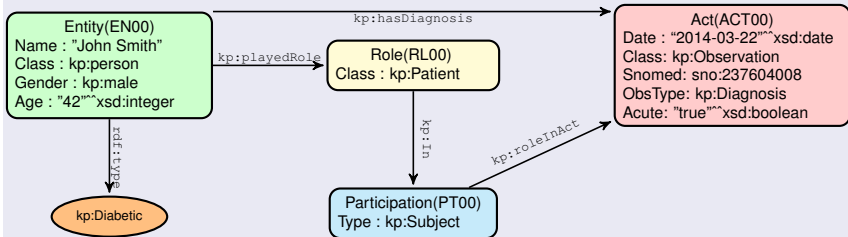


$$\langle ?e, kp:hasDiagnosis, ?a \rangle \leftarrow \langle ?e, kp:playedRole, ?r \rangle, \langle ?r, kp:In, ?p \rangle$$

$$\langle ?p, kp:roleInAct, ?a \rangle.$$

$$\langle ?e, rdf:type, kp:Diabetic \rangle \leftarrow \langle ?e, kp:hasDiagnosis, ?a \rangle, \langle ?a, kp:SNOMED, sno:237604008 \rangle$$

LABELLING



$$\langle ?e, kp:hasDiagnosis, ?a \rangle \leftarrow \langle ?e, kp:playedRole, ?r \rangle, \langle ?r, kp:In, ?p \rangle$$

$$\langle ?p, kp:roleInAct, ?a \rangle.$$

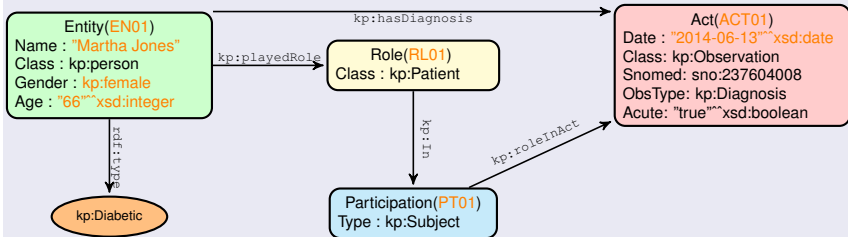
$$\langle ?e, rdf:type, kp:Diabetic \rangle \leftarrow \langle ?e, kp:hasDiagnosis, ?a \rangle, \langle ?a, kp:SNOMED, sno:237604008 \rangle$$

$$\langle ?e, kp:eligible, ?y \rangle \leftarrow \langle ?e, rdf:type, kp:Diabetic \rangle, \langle ?e, kp:hasDiagnosis, ?a \rangle,$$

$$\langle ?a, kp:date, ?date \rangle, BIND(YEAR(?date) as ?y)$$

$$\langle ?e, kp:age, ?age \rangle, FILTER(?age \leq 65).$$

LABELLING



$$\langle ?e, kp:hasDiagnosis, ?a \rangle \leftarrow \langle ?e, kp:playedRole, ?r \rangle, \langle ?r, kp:In, ?p \rangle$$

$$\langle ?p, kp:roleInAct, ?a \rangle.$$

$$\langle ?e, rdf:type, kp:Diabetic \rangle \leftarrow \langle ?e, kp:hasDiagnosis, ?a \rangle, \langle ?a, kp:SNOMED, sno:237604008 \rangle$$

$$\langle ?e, kp:eligible, ?y \rangle \leftarrow \langle ?e, rdf:type, kp:Diabetic \rangle, \langle ?e, kp:hasDiagnosis, ?a \rangle,$$

$$\langle ?a, kp:date, ?date \rangle, BIND(YEAR(?date) as ?y)$$

$$\langle ?e, kp:age, ?age \rangle, FILTER(?age \leq 65).$$



RDFox DESCRIPTION

- Highly optimised in-memory RDF triple store — fast querying
- Parallel Datalog Reasoner — low contention
- Scales exceptionally well — fine grained parallelisation

RDFox DESCRIPTION

- Highly optimised in-memory RDF triple store — fast querying
- Parallel Datalog Reasoner — low contention
- Scales exceptionally well — fine grained parallelisation

RDF ALLOWS TO COMPUTE HEDIS MEASURES

- For millions of patients with <256GB RAM
- In <5min with 8 cores (16 cores hyper-threading)

RDFox DESCRIPTION

- Highly optimised in-memory RDF triple store — fast querying
- Parallel Datalog Reasoner — low contention
- Scales exceptionally well — fine grained parallelisation

RDF ALLOWS TO COMPUTE HEDIS MEASURES

- For millions of patients with <256GB RAM
- In <5min with 8 cores (16 cores hyper-threading)

RDFox SUPPORTS

- Multi-threading — exploits modern hardware
- RDF — wide-spread standard data format
- Datalog — covers a range of OWL fragments
- SPARQL 1.1 — standard versatile query language

OUR APPROACH

- Separate format unification from Measuring
- Format: HL7 RIM data model encoded in RDF Triples
- Measuring: Datalog ontology

OUTLOOK

- Full implementation of HEDIS Diabetes Care Measurement
- Developing interfaces for query composition/data browsing
- Use SNOMED ontology with patient data