-SemanticHealthNet-  
A Semantic Infrastructure Towards Semantic Interoperability

Catalina Martínez-Costa¹, Stefan Schulz¹ and Dipak Kalra²

¹IMI, Medical University of Graz, Austria  
²CHIME, University College London, UK
Introduction - SemanticHealthNet (SHN) Project

• **3 years** EC NoE Project (2012-2015)

• **Global mission:**
  – Develop and test approaches to **improve semantic interoperability** of health data
  – Create a **virtual organization** to sustain semantic interoperability developments and their adoption across Europe

• **Workpackage 4 mission:**
  – Provide an intermediate semantic layer able to **deal with the unavoidable heterogeneity** which arises when clinical information is represented across or within the same medical domain
Workpackage 4 Basic Assumptions

• Plurality of Information Model approaches exists:
  – openEHR, EN ISO 13606, HL7 RIM, CIMI, SIAMM, etc.
  – Local schemas are still predominant
  – Information model like structures in existing terminology context model of SNOMED CT
  – Free text (out of scope in SHN)

• Plurality of representations within one specification

• WP4’s relation to Information Models:
  – Does not develop “yet another” information model
  – Maintains equidistance and neutrality
  – Looks at content and not at structure
The role of SNOMED CT in the Project

• Provide agreed formal definitions of medical concepts (i.e. act as medical domain ontology)

• Reference ontology for representing medical domain concepts (e.g. mappings from ICD-11, LOINC, etc. to SNOMED CT)

Reference terminology

Formal definition Fracture of femur concept:

```
Description: 'Fracture of femur (disorder)'
Equivalent To

'Fracture of lower limb (disorder)' and 'Role group (attribute)' some
  ('Associated morphology (attribute)'
    some 'Fracture (morphologic abnormality)') and
  ('Finding site (attribute)'
    some 'Bone structure of femur (body structure)'))
```

Logical reasoner

mapped-to

OWL
Medical records contain items of information that refer to his clinical situation (described using medical concepts).
Existing Terminology Binding Approaches

- **Guidelines specifications-based approaches**: they address the most common overlaps and provide modelling guidelines to resolve ambiguities (e.g. TermInfo, NHS openEHR work)

**HL7 TermInfo guideline rule example:**

```
“An Observation class instance in which the Observation.value is a SNOMED CT expression representing a [<<404684003 | clinical finding] or a [<<413350009 | finding with explicit context] SHALL NOT contain an Act.code which when interpreted with the Observation.value yields a meaning that is substantially different from the meaning implied if the Act.code was "ASSERTION".

- For example, an Act.code meaning "Past history" or "Family history" may substantially alter the interpretation of a [<<404684003 | clinical finding] and should not be used in this way. Instead the SNOMED CT context model should be used to capture these significant differences in meaning.”
```
Existing Terminology Binding Approaches (II)

• Define **clinical models** and constrain their elements and values to a **set of SNOMED CT values**.
  – **Simpler approaches**: EN ISO 13606, openEHR, etc.
  – More **sophisticated approaches** based on the definition of a set of general clinical models: CIMI, SIAMM

• **Simple approach example (EN ISO 13606)**:

  ![Diagram](image)

  Non semantic interoperable clinical models
Existing Terminology Binding Approaches (II)

- **Sophisticated approach example (CIMI):**

  ```
  ENTRY[at0000.1] matches { -- Observation
      link matches {LINK[at0.1] occurrences matches {0..*} -- Associated request}
      data matches {
          \rightarrow use_archetype CLUSTER [CIMI-CORE-CLUSTER.observable.v1] -- Observable
          use_archetype CLUSTER [CIMI-CORE-CLUSTER.finding.v1] -- Results
          use_archetype CLUSTER [CIMI-CORE-CLUSTER.observe_action.v1] -- Observe action
      }
  }
  
  CLUSTER[at0000] matches { -- Observable
      item matches {
          ELEMENT[at0001] occurrences matches {1} matches { -- Name
              value matches { TEXT matches {*}}}
          ITEM[at0002] occurrences matches {0..*} -- Reason
          ITEM[at0003] occurrences matches {0..*} -- Method
          ELEMENT[at0004] occurrences matches {0..1} matches { -- Status
              value matches { CODED_TEXT matches {*}}}
      }
  }
  
  CLUSTER[at0000] matches { -- Action
      item matches {
          ELEMENT[at0001] occurrences matches {1} matches { -- Action type
              value matches { CODED_TEXT matches {*}}}
          ITEM[at0009] occurrences matches {0..*} -- Reason
          ITEM[at0010] occurrences matches {0..*} -- Method
      }
  }
  ```

- Simple approaches do not address the boundary problem and produce non-interoperable clinical models
- Sophisticated approaches forbid the use of certain SNOMED CT hierarchies (e.g. context model)
- In both the decision of the elements to include in a clinical model is mainly a modeler decision not guided by any formal constraint
• **Ontology design content patterns**: small fragments of an ontology for modelling a specific use case.

• First introduced by Gangemi, Blomqvist and Sandkuhl in 2005.

• They were devised to **guide and standardize** the way ontologies are developed.

• Intended to **help non-expert** ontology users.

• They **package best practice into reusable blocks** of ontology functionality, to be adapted and specialized by those users in their individual ontology development use cases.
SHN Semantic Patterns

- They are based on the **SemanticHealthNet ontological framework**
SHN Semantic Patterns

• They are language-independent and should be encoded in a high order representation language.

• We have represented them as:
  – A set of RDF (Subject-Predicate-Object) triples enhanced by cardinality constraint:

    \[
    \text{shn:InformationItem} \text{ describes situation } \text{shn:ClinicalSituation} \\
    \text{shn:InformationItem} \text{ results from process } \text{shn:ClinicalProcess}
    \]

  – OWL 2 DL:
    • RDF Subject and Object transform into OWL classes
    • RDF Predicate transforms into OWL DL expression:

\[
\text{shn:InformationItem} \text{ and shn:isAboutSituation only shn:ClinicalSituation} \\
\text{and shn:isOutcomeOf some shn:ClinicalProcess}
\]
They can be **specialised and composed** by following similar principles to object oriented languages.
OUR HYPOTHESIS:
• A limited set of top-level semantic patterns that can be specialized and composed is sufficient to represent a great variety of clinical information.

7 top-level patterns extracted from the SHN Heart Failure Summary

<table>
<thead>
<tr>
<th>Top-level pattern</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBSERVATION RESULT PATTERN (OB_PT)</td>
<td>record of body weight, height, etc.</td>
</tr>
<tr>
<td>INFORMATION CLINICAL SITUATION (I_CS_PT)</td>
<td>Cancer diagnosis, breathlessness symptom, etc.</td>
</tr>
<tr>
<td>PLAN CLINICAL PROCESS (P_CP_PT)</td>
<td>request to administer some drug, etc.</td>
</tr>
<tr>
<td>CLINICAL PROCESS (CP_PT)</td>
<td>assessment, history taking, etc.</td>
</tr>
<tr>
<td>CLINICAL SITUATION (CS_PT)</td>
<td>heart attack, diabetes, cancer, etc.</td>
</tr>
<tr>
<td>PAST HISTORY CLINICAL SITUATION (PH_CS_PT)</td>
<td>past history of heart failure, past history of cancer, etc.</td>
</tr>
<tr>
<td>FAMILY HISTORY CLINICAL SITUATION (FH_CS_PT)</td>
<td>family history of diabetes, family history of high blood pressure, etc.</td>
</tr>
</tbody>
</table>
Pattern Specialisation / Composition
Semantic Framework

Homogeneous access point to clinical information

Layer 5: Application
- Research
- Public health
- Quality measurement
- Primary care

Layer 4: Virtual Homogeneous Data

Layer 3: Semantic Mediator
- Common ontology + SNOMED CT
- DL reasoning

Layer 2: Semantic Mapping
- COP
- Content ontology patterns (COPs)

Layer 1: Structured Heterogeneous Data
- HL7 CDA
- openEHR
- ISO 13606
- LOCAL SCHEMA

SEMANTIC INFRASTRUCTURE

Description logics formal ontology encompassing clinical and informational entities

Semantic patterns as frame-like user-friendly representation

Non-interoperable, possibly isosemantic clinical information
SHN Semantic Patterns role

• Allow the consistent use of SNOMED CT within EHR clinical models (i.e. **address the boundary problem**)

• Enable **semantic interoperability** across heterogeneously structured clinical models within or not the same EHR specification
  – Enable **querying** over data heterogeneously structured and encoded

• Provide **advanced clinical information exploitation** capabilities

• **Guide** the **development** of new clinical models

• Detect **semantic inconsistencies** across existing clinical models
Semantic Interoperability Example

Clinical study about endocrine diseases

• Patients with history of some endocrine disease
• Patients with history of diabetes mellitus
• Patients with history of mild diabetes mellitus
• Patients with history of gestational diabetes
Semantic Interoperability Example (II)

General questionnaire – ISO 13606 Representation

ENTRY[at0000] matches { -- Question group
items matches {
CLUSTER[at0001] matches { -- Question group
items matches {
CLUSTER[at0002] matches { -- Question
items matches {
ELEMENT[at0003] matches { -- Answer
value matches {
BL matches {True, False}

SNOMED CT Terminology Binding

'History of clinical finding in subject (situation)' and 'Role group (attribute)' some ( 'Associated finding (attribute)' some 'Diabetes mellitus (disorder)' and 'Finding context (attribute)' some 'Known present (qualifier value)' and 'Temporal context (attribute)' some 'In the past (qualifier value)' and 'Subject relationship context (attribute)' some 'Subject of record (person)'

84100007 [history taking]

History - clinical situation pattern
Semantic Interoperability Example (III)

Gynaecologist form – ISO 13606 Representation

ENTRY[at0000] matches { -- Past history
items matches {
  ELEMENT[at0001] matches { -- Condition
value matches {
    CODED_TEXT matches {(*) }
  }
  }
  CLUSTER[at0002] matches { -- Details
items matches {
  ELEMENT[at0001] matches { -- Severity
value matches {
    CODED_TEXT matches {(*) } }}
  }
}

SNOMED CT Terminology Binding

terminology binding

'Pregnancy observable (observable entity)'

'Diabetes mellitus (disorder)'

'Mild (qualifier value)'

shn:InformationItem

'shns:ClinicalSituation'

'shns:ClinicalProcess'

'shns:InformationAttribute'

'sct:InThePast'

'sct:FindingContextValue'

btl:MaterialObject

btl:MentalObject OR btl:InmaterialObject

btl:TemporalRegion

'occurs at' [0..]

'happens at' [0..]

'follows' [0..]
“Non-technical” challenges:

- Get more evidence that clinical models information can be sufficiently represented by semantic patterns
- Get more evidence that a limited number of top-level patterns is sufficient to derive more specific patterns by specialisation / composition mechanisms
- Engage clinical modelers community to help in testing the content coverage of the patterns

Technical challenges:

- Provide different representations for different tasks
  - Closer to user representation
  - Logical representation (OWL DL)
  - RDF representation (RDF Shapes, SPIN)
  - UML-like representation
  - Etc.
- Solve performance issues related with the use of OWL DL
- Grow libraries of patterns
- Provide tools that facilitate clinical modelers engagement
Conclusion

- We need to be able to use terminologies consistently within EHR information models to achieve semantic interoperability.

- We need methods that allow their consistent use independently of the particular EHR representation (we need to focus on the content and not on the structure!)

- Semantic patterns allow setting the focus on the content (information meaning).

- Semantic patterns were motivated by our experiences of representing semantically clinical information.

- For getting more evidence of their usefulness we need the engagement of the community in order to see if this is something else than theoretical research.
Thanks for your attention

Questions?

Comments?