A Semantic Approach to International Terminology Artifacts Harmonisation

Jean-Marie Rodrigues, Stefan Schulz, Alan Rector, Kent Spackman, Jane Millar, James Campbell, Bedirhan Üstün, Christopher Chute, Harold Solbrig, Vincenzo Della Mea, Kristina Brand Persson.

rodrigues@univ-st-etienne.fr

26 september 2014

EUON Amsterdam 24-25 September 2014

Plan

- SNOMED CT and ICD11
- Artifacts model
- Common Semantic
- Circulatory system pilot
- Conclusion

SNOMED CT and ICD11

- **SNOMED CT**
- standardized health terminology that provides clinical terms for health records.
  SNOMED CT “concepts” or entities
  provide standardized meaning by logic descriptions

- **ICD 11**
- delivering summary high level aggregations
- a multi-layered disease classification system.
- Foundation Component: Common Semantic plus queries
- Purpose-specific linearizations (mortality, morbidity, primary care) with jointly exhaustive and mutually exclusive classes

Terminology Artifacts Model

- Information entities
- Situations
- Conditions

Situations vs. Conditions

- SNOMED CT ICD11 :

artifact	Interpretation
SNOMED CT findings / disorders	Situations
Content of role groups in finding / disorder hierarchy	Conditions
ICD 11	Situations Information entities

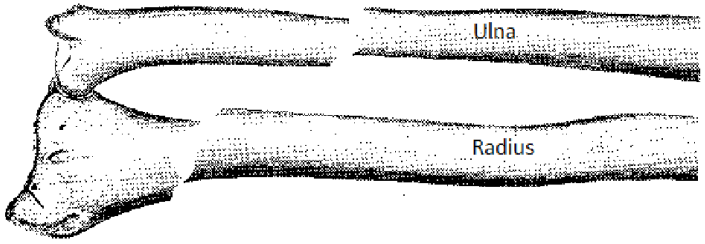
  - Relevance for JAG CO?

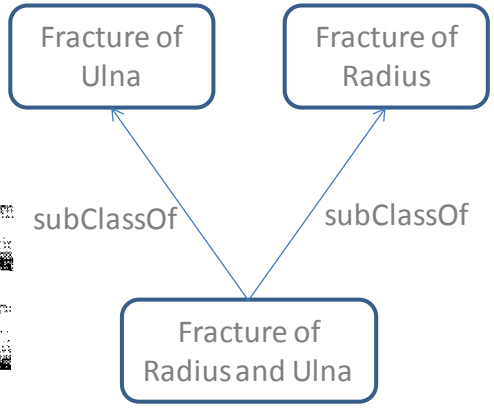
Example ICD 11 Information entities

ICD-11:1E10 *Rabies*

- ICD-11:1E10.1 *Suspected rabies*
- ICD-11:1E10.*2* *Probable rabies*
- ICD-11:1E10.3 *Confirmed rabies*

Example of situations: Fracture of Radius and Ulna Is this True?



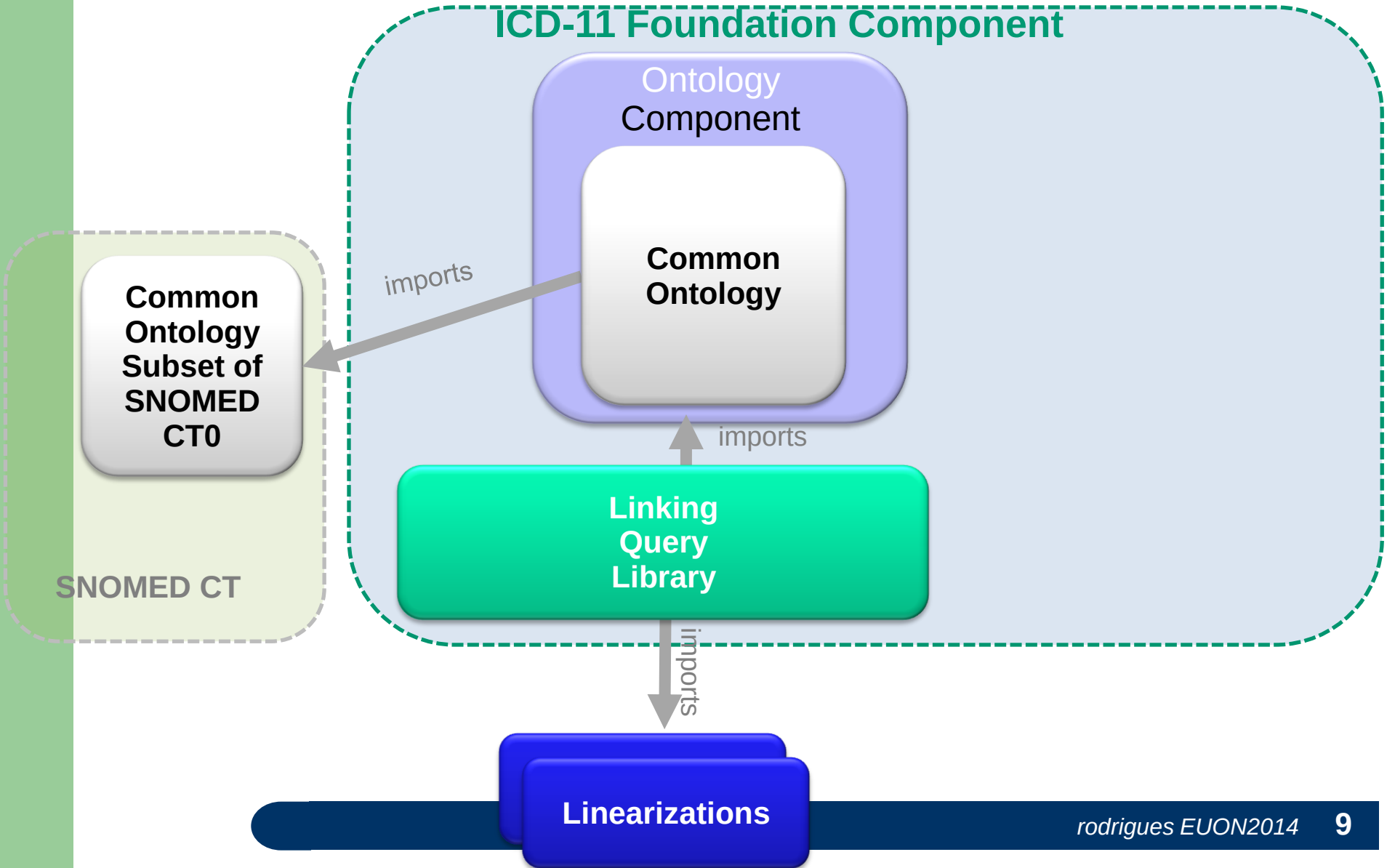


- FALSE, if X means “pathological entity”
- TRUE, if X means “situation with X” or “patient having X”

Semantic approach definition

- The Common Semantic is a set of ICD Foundation Component terms harmonised with a subset of SNOMED CT and formulated in SNOMED CT's description logic.
- OWL/DL SubClass/SuperClass
- Non Is_A Definition relations (axioms)
  - Fully defined terms (necessary and sufficient)
  - Partially defined terms (necessary plus manual assertion)
  - Poly-hierarchy

Simplified Architectural Diagram



Linearizations differences with the common ontology

- Mono-hierarchy(different from Co Is_a)
- Specific preferred term
- Inclusion/exclusion
- Residuals
- Related to Common Semantic by
  - a disjunction of terms in the *Common Semantic*
  - *Queries*

the basic structure of most queries

- the basic structure of most queries are of the form “A but not B or C or …”, *i.e.*
- LinearizationClass → BaseOntologyClassExpression
  MINUS OntExpression1
  MINUS OntExporession2
- where “MINUS” can be read as “not classified under”.

Queries examples

- ICD:'Essential Hypertension' →
  FROM sct: Findings SELECT
  sct: 'Hypertensive Disease'
- ___MINUS sct Secondary hypertension

- ICD;'Hypertensive disease'
- FROM sct: Findings SELECT
  - Sct Hypertensive disease
- MINUS sct neonatal hypertension
- MINUS sct complication of pregnancy, childbirth or puerperium

Circulatory system pilot workflow

- identify candidate SCT concepts for FC classes.
- Construct valid SCT formal expressions for ICD-11 textual definitions,
- Specify the queries for ICD-11 FC classes.

Common Semantic data base

- SQL database
  - Lexical matches
  - Automatic input of SCT terms and Short Norms Form
  - Import export process via WHO URL
- References
  - e.g. Medscape/OMIM/Orphanet
- Internal review approach
- External review
  - American College of Cardiology for circulatory

Identifying SCT candidates to Common Semantic

Map Type	Description	Common Semantic(CO)
SCTID	ICD11 maps to existing SNOMED CT, synonymous	Short Normal Form
0/A	pre-coordinated concept without change	Definition using SNOMED CT concepts and compositional grammar
0/E	post-coordinated expression without change in the model of meaning	Definition using SCT attributes only e.g. site, morphology, expression provided in SNOMED CT compositional grammar
0/R	residual phrase	Not in CO but queries
0/X	pre-coordinated concept with changes	Change to SCT model of meaning
0/EX	post-coordinated expression with changes	Change to content – object/value

– Exact match examples

Map Type	ICD11 Rubric	Common Semantic (Short Normal Form)
SCTID	Cerebral venous thrombosis	95464572001|Disease (disorder)|:{116676008|Associated morphology (attribute)|=396339007|Thrombus (morphologic abnormality)|363698007|Finding site (attribute)|=68351006|Structure of cerebral vein (body structure)|}
SCTID	Coronary vaso spastic with angina	194828000|Angina (disorder)|+23687008|Coronary artery spasm (disorder)|:{363698007|Finding site (attribute)|=74281007|Myocardium structure (body structure)|}

– Pre coordination examples

Map Type	ICD11 Rubric	Common Semantic
0/A	Aldosterone-producing carcinoma	116680003 | Is a |88213004 | Hyperaldosteronism, |42752001 | Due to|255035007 | Adrenal carcinoma
0/A	Acute myocardial infarction, STEMI, anterior wall	401303003 | Acute ST segment elevation myocardial infarction | + 54329005 | Acute anterior myocardial infarction

Post coordination – examples

Map Type	ICD11 Rubric	Common Semantic
0/E	Asymptomatic stenosis of extracranial carotid artery	116680003 | Is a |230738008 | Asymptomatic cerebrovascular disease, |363698007 | Finding site|17999001 | Structure of cervical portion of internal carotid artery, |116676008 | Associated morphology|415582006 | Stenosis
0/E	internal auditory artery occlusion	116680003 | Is a |2929001 | Occlusion of artery, |363698007 | Finding site|89471000 | Structure of labyrinthine artery

Results

Map Type	Number	Common Semantic
SCTID	792 (50.1%)	Short Normal Form
0/A	534 (33.8%)	To be developped with SCT grammar
0/E	86 (5.4%)	To be developped with SCT grammar
0/EX	16 (1.0%)	To be developped with changes SCT values
0/X	12 (0.8%)	To be developped with changes in SCT model of meaning
0/R	116 (7.3%)	Not in CO but queries
Unclear definition	26 (1.6%)	Definition by domain experts

Conclusion

- Around 84 % of circulatory chapter ICD11 codes can be represented by a subset of existing SCT concepts and pre-coordination
- 5,4 % can be represented with SCT attributes (or defining relations or model of meaning)
- 1,8 % needs modification of the model of meaning (attribute or values sets)
- 7,3 % and 1,6 % cannot be represented in CO

Questions?

Tak

Vielen Dank für Ihre Aufmerksamkeit

Terima Kashi

Thank you

Merci

rodrigues@univ-st-etienne.fr

Tak
Grazie
Gracias
Obrigado
Efcharisto
ARIGATO GOZAIMASU
Multumesc.