

SNOMED CT - Einstieg in eine neue medizinische Terminologie

30 Jun 2014, 14-17h, Inselspital Bern, Pathologisches Institut, Hörsaal 7, Auditorium Langhans



**Stefan
Schulz**
Medical
University
of Graz
(Austria)

purl.org/steschu



SNOMED CT's relevance for safe and interoperable data exchange in health care

Outline

- Is SNOMED CT relevant in the future?
- Which challenges will characterise the introduction of SNOMED CT?
- How big data and personalised medicine matter for SNOMED CT ?

Outline

- What SNOMED CT is and isn't?
- Is SNOMED CT relevant in the future?
- Which challenges will characterise the introduction of SNOMED CT?
- How big data and personalised medicine matter for SNOMED CT ?

What SNOMED CT is and isn't?

What SNOMED CT is and isn't?

SNOMED CT is NO LONGER a nomenclature

- Originally:
 - SNOMED = “Systematized **Nomenclature** of Medicine”
- Today:
 - The IHTSDO has removed the long form (and the word “nomenclature”) from its documents and educational materials
 - A nomenclature would be prescriptive with regard to naming
 - SNOMED CT is normative with regard to its representational units (“concepts”), and the structure of its concept labels (FSNs), but descriptive with regard to the terms (“descriptions”) in general

What SNOMED CT is and isn't?

SNOMED CT is a large clinical terminology

What SNOMED CT is and isn't?

SNOMED CT is a large clinical terminology

- Standardised „Fully Specified Names“

- Mostly self-explaining
- Unambiguous
- Wordy
- Often not used in clinical discourse

Current Concept:

Fully Specified Name: Viral hepatitis, type A (disorder)

ConceptId: 40468003

- Synonyms

- Context-dependent
- Possibly ambiguous
- Concise
- Typically used in clinical discourse

Descriptions (Synonyms):

Fully Specified Name: Viral hepatitis, type A (disorder)

Synonym: IH - Infectious hepatitis [492634014]

Synonym: Viral hepatitis, type A [67480019]

Synonym: Infectious hepatitis [67481015]

Synonym: Hepatitis A [67483017]

Synonym: Viral hepatitis A without mention of hepatic coma |

What SNOMED CT is and isn't?

SNOMED CT is a large clinical ontology

What SNOMED CT is and isn't?

SNOMED CT is a large clinical ontology

- Collection of clinical “concepts”
 - Language-independent representational units
 - Represent classes of individual things like procedures, substances, organisms, findings etc.
 - Uniquely identified by code (SNOMED ID) and linked to terms (at least to one fully specified name, which is also unique)
- Collection of formal axioms (EL++)
 - Logically declare what is universally true for all instances of a SNOMED CT concept
 - Constituting taxonomies
 - Introduce necessary conditions
 - Fully define concepts (Aristotelian definitions)

What SNOMED CT is and isn't?

SNOMED CT is a large clinical ontology

Multi-
Taxonomy
via subclass
axioms

(Defining)
relationships
“universally true”

[Disease due to Picornaviridae \(disorder\)](#)
[Viral hepatitis \(disorder\)](#)

Current Concept:
Viral hepatitis, type A (disorder)

Child(ren):
(N=7) (Select a child to make it the "Current Concept".)
[Acute type A viral hepatitis \(disorder\)](#)

Current Concept:

Fully Specified Name: Viral hepatitis, type A (disorder)
ConceptId: 40468003
Source: Core

Defining Relationships:

Is a	Disease due to Picornaviridae (disorder)
Is a	Viral hepatitis (disorder)
Causative agent (attribute)	Hepatitis A virus (organism)
Pathological process (attribute)	Infectious process (qualifier value)
Group 1	
Associated morphology (attribute)	Inflammation (morphologic abnormality)
Finding site (attribute)	Liver structure (body structure)

This concept is sufficiently defined.

What SNOMED CT is and isn't?

SNOMED CT is NOT a typical knowledge base

- Restriction to “sustainable” ontology content proper excludes contingent or probabilistic associations such as
 - Diagnostic criteria for diseases
 - Indications or side effects of drugs
- Formal language used for SNOMED CT only expresses what is “universally true”
- Advance of science might lead to the addition axioms but rarely retract axioms

“There are very few interesting items of knowledge that are truly ontological”
(A. Rector)

What SNOMED CT is and isn't?

SNOMED CT is NOT a typical knowledge base

What SNOMED CT is and isn't?

SNOMED ALSO is a collection of clinical statements

What SNOMED CT is and isn't?

SNOMED ALSO is a collection of clinical statements

- metastasis to peritoneum of unknown primary tumor
- Suspected autism
- Suspicion of gastritis
- No antenatal care: not known pregnant
- No drug side effect reported
- T1a (IA): Invasive carcinoma of uterine cervix diagnosed by microscopy only
- Diabetes mellitus excluded
- Surgical pathology consultation and report on referred slides prepared elsewhere
- Previous known suicide attempt
- Medication not administered
- Helicobacter blood test negative
- Poor condition at birth without known asphyxia
- Natural death with probable cause suspected
- Unlikely diagnosis
- Operating room unavailable

What SNOMED CT is and isn't?

SNOMED ALSO is a collection of clinical statements

- No “terms” in a proper sense
- Statements that contain terms and place them into a specific context (plan, suspicions, exclusions, past history, family history etc.)
- Mostly inherited via Read Codes / CTV3
- Typically found in the “situation” subhierarchy
- Driven by need for one-click encoding of complex but recurrent pieces of information in the EHR
- Difficult to formally represent
- Obstacle to interoperability: Context *should* ideally be represented by the information model to which the terminology is bound

Outline

- What SNOMED CT is and isn't?
- Is SNOMED CT relevant in the future?
- Which challenges will characterise the introduction of SNOMED CT?
- How big data and personalised medicine matter for SNOMED CT ?

Is SNOMED CT relevant in the future?

Is SNOMED CT relevant in the future?

- Yes, provided
 - It is safe
 - It supports relevant interoperability use cases
 - It is harmonized with other terminologies

Is SNOMED CT relevant in the future?

SNOMED CT's safety

- Ontological content
 - Reliable maintenance and quality control by IHTSDO
 - Concept permanence, non-disruptive update
 - Reliable inferences (still issues to be fixed)
 - Clear concept definitions (formal and text)
- Terminological content
 - Self-explaining fully specified names
 - Quality of translated content (translation guidelines)
 - Awareness of ambiguity (especially short hand expressions)
 - RTA: renal-tubular acidosis vs. road traffic accident
 - HWI: Hinterwandinfarkt vs. Harnwegsinfekt

Is SNOMED CT relevant in the future?

Interoperability support

- Terminological interoperability
 - Multilingual support
 - Different languages
 - Different language registers
 - Coverage of synonyms, term variants
- Ontological content
 - Ability to detect equivalence between syntactically different expression
 - Clear rules of interplay with different information models (e.g. HL7, EN13606, openEHR), ideally supported by formal semantics

Is SNOMED CT relevant in the future?

Interoperability support

- Computational detection of equivalence between pre-coordinated and post-coordinated expressions

**Organ Failure
Diagnosis**

Organ Heart ▼

Status Suspected ▼

**Caused by
ischaemic
heart
disease** Yes X
No
Unknown

Diagnosis

Suspected heart failure cause
by ischaemic heart disease ▼

Diagnosis

Heart Failure ▼

Status Suspected ▼

Cause Ischaemic hear
disease ▼

Example from SemanticHealthNet Network of Excellence (www.semantichhealthnet.eu)

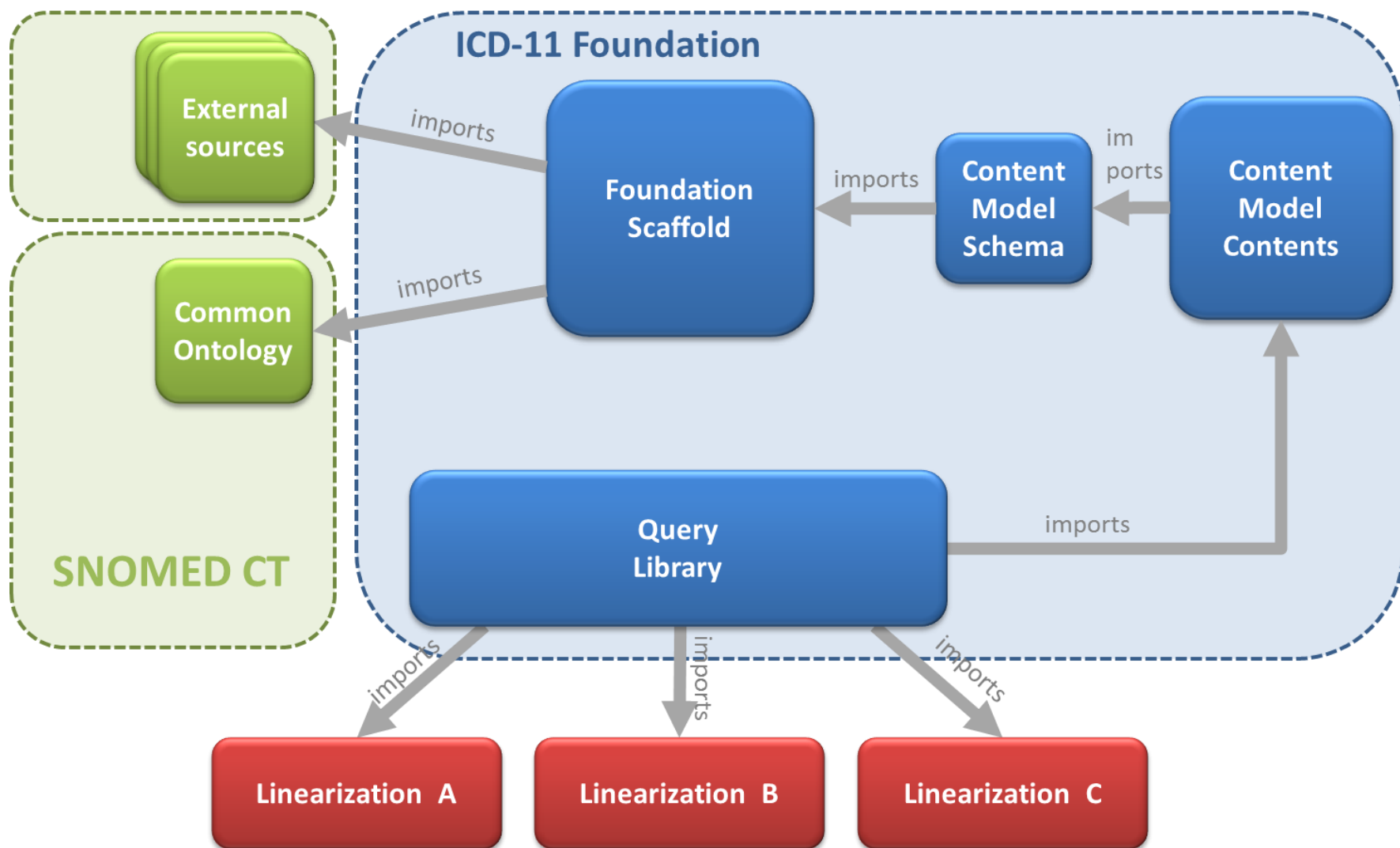
Is SNOMED CT relevant in the future?

Harmonization with other terminologies

- Agreement with LOINC for Lab Medicine – harmonization with SNOMED CT Observables hierarchy
- Agreement with WHO
 - ICD-10
 - ICD-11
 - ICF
 - ICD-O
- Classification-like terminologies: no direct mapping but formulation of queries against joint ontology (subset of SNOMED CT)

Is SNOMED CT relevant in the future?

Example: Harmonization SNOMED CT - ICD-11



Outline

- What SNOMED CT is and isn't?
- Is SNOMED CT relevant in the future?
- Which challenges will characterise the introduction of SNOMED CT?
- How big data and personalised medicine matter for SNOMED CT ?

Challenges of SNOMED CT introduction

- Find appropriate use cases
- Identify the users
- Develop realistic and affordable localisation strategy

Challenges of SNOMED CT introduction

Find appropriate use cases

Challenges of SNOMED CT introduction

Find appropriate use cases

- Two major use cases to “snomedize” the electronic health record
 1. Linkage of elements in structure data entry instruments to SNOMED CT expressions (e.g. patient summaries)
 2. Automated semantic annotation of free text content

Challenges of SNOMED CT introduction

Semantic annotation of structured information

Patient: Franz Müller
DoB: 17/01/1965
Encounter: 23/08/98

Tobacco Use Summary
"Cigarette smoker, 14 /d"



Hospital

-- Form

InformationItem

and **isAboutSituation** only **CigaretteTobaccoSmoking**
and **isOutcomeOf** some **HistoryTaking**

-- Amount

ObservationResult

and **isAboutQuality** some (*MassIntake*
and **inheresIn** some **CigaretteTobaccoSmoking**
and **projectsOnTo** some (*ValueRegion*
and **isRepresentedBy** only (**hasValue** value 14
and **hasInformationAttribute** some **PerDay**)))

```

EVALUATION[at0000] matches { -- Tobacco Use Summary
  data matches {
    ITEM_TREE[at0001] matches { -- Tree
      items cardinality matches {0..*; unordered} matches {
        ELEMENT[at0002] occurrences matches {0..1} matches { -- Smoking
          Status
          value matches {
            DV_CODED_TEXT matches {
              defining_code matches {
                [local::
                  at0003, -- Current Smoker
                  at0027, -- Never Smoked
                  at0005] -- Ex-smoker
                ]}}}
            CLUSTER[at0029] occurrences matches {0..*} matches { -- Smoking
              Details
              items cardinality matches {1..*; unordered} matches {
                ELEMENT[at0028] occurrences matches {0..1} matches { -- Form
                  value matches {
                    DV_TEXT matches {*}}}}
                ELEMENT[at0030] occurrences matches {0..1} matches { -- Typical
                  smoked amount
                  value matches {
                    C_DV_QUANTITY <
                      property = <[openehr::382]>
                      list = <
                        ["1"] = <
                          units = <"/d">
                          magnitude = <|>=0.0|>
                          precision = <|0|>
                        >>}}}}...}}}}

```

openEHR archetype

Challenges of SNOMED CT introduction

Semantic annotation of texts

Challenges of SNOMED CT introduction

Semantic annotation of texts

Physical examination on admission revealed purpura of the upper and lower extremities, swelling of the gums and tonsils, but no symptoms showing the complication of myasthenia gravis. Hematological tests revealed leucocytosis: WBC count 68 700/ μ l (blasts 11.5%, myelocytes 0.5%, bands 2.0%, segments 16.0%, monocytes 65.5%, lymphocytes 4.0%, atypical lymphocytes 0.5%), Hb 7.1 g/dl (reticulocytes 12%) and a platelet count of 9.1×10^4 / μ l. A bone marrow aspiration revealed hypercellular bone marrow with a decreased number of erythroblasts and megakaryocytes and an increased number of monoblasts

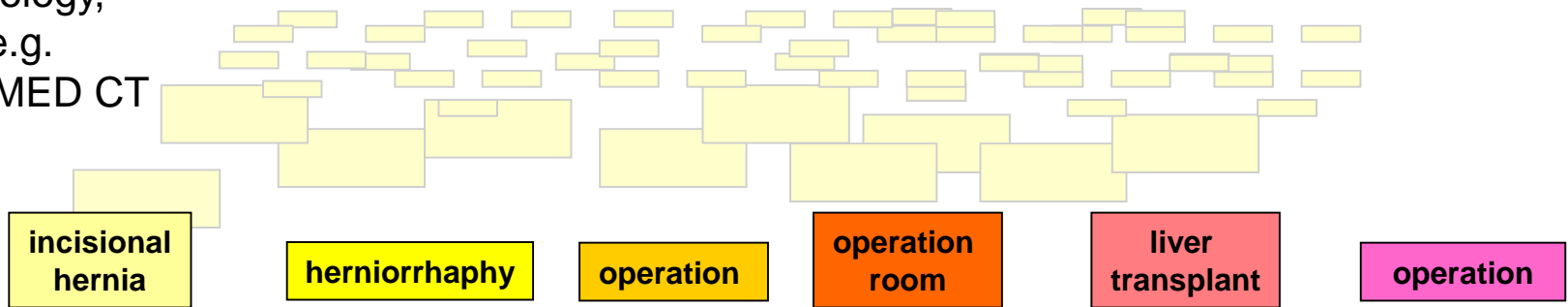
```
ADM{419620001  
110714004  
65124004 113279002  
NEG{116223007 91637004 252275004  
111583006} 767002[68700] 271040006[11.5]  
313696224[0.5] 313696667[2.0]  
313696009[16.0] 271037006[65.5]  
271036002[4.0] 271036013[0.5]  
365809007[7.1] 45995003[12]  
365632008[91000] }  
49401003 76197007 14016003  
420510009 103213002  
53945006 35105006
```

- Numerous challenges: context, negation, ambiguities..

Challenges of SNOMED CT introduction

Ontology Mapping is not enough → CONTEXT!

Ontology,
e.g.
SNOMED CT



- exists
- doesn't exist
- reference to a suspended plan
- exists
- reference to a plan
- reference to a new plan
- NOT referring to the same patient
- not yet executed

Patient with **incisional hernia** admitted for **herniorrhaphy**, but **operation** was suspended because **operation room** was urgently needed for **liver transplant**. Discharged with orientation and rescheduled **operation**.

Challenges of SNOMED CT introduction

Understanding medical language is hard

3. St.p. TE eines exulc.
sek.knot.SSM (C43.5) li Lab.
majus. Level IV, 2,42 mm
Tumordurchm.

Challenges of SNOMED CT introduction

Find appropriate use cases

- Two major use cases to “snomedize” the electronic health record
 1. Linkage of elements in structure data entry instruments to SNOMED CT expressions (e.g. patient summaries)
 2. Automated semantic annotation of free text content
- **Direct coding by physicians is not a primary use case**
 - Coding load already seen as sufficient by physicians
 - Coding is purpose-oriented and context dependent
 - Coding bias (e.g. co-morbidities not coded as seen as irrelevant for reimbursement)

Challenges of SNOMED CT introduction

Identify the users

Challenges of SNOMED CT introduction

Identify the users

- Clinicians: “indirect” users of SNOMED CT within advanced EHR systems. They want to see interface terms, not codes
- Standards developers, policy makers, epidemiologists: interested in international patient summaries, interoperable registries
- Software manufacturers interested in semantically enhanced, interoperable clinical information systems interested in ready-to-use, multilingual value sets
- Clinical controllers: querying EHR content for business analytics
- Researchers: making use of secondary use scenarios, retrospective studies, patient recruitment for clinical trials, using SNOMED CT as semantic reference for data in clinical trials

Challenges of SNOMED CT introduction

Localisation strategy

Challenges of SNOMED CT introduction

Localisation strategy

- Localisation... not just terminology translation
- Learn from others: Denmark: full SNOMED CT FSN translation available but not in use

“The standardized terms available at present are inadequate for clinical use”
(Result from focus group interview in Denmark)

- Criticisms:
 - no close-to user synonyms: complicates search and the creation of user friendly value sets
 - Current use cases well served by terminologies in use

Outline

- What SNOMED CT is and isn't?
- Is SNOMED CT relevant in the future?
- Which challenges will characterise the introduction of SNOMED CT?
- How big data and personalised medicine matter for SNOMED CT ?

Big data and personalised medicine

Big data



NSA



IBM Watson

Big data and personalised medicine

Big data

- Big data: automated learning from huge amounts of patient records from disparate sources
 - Finding “similar patients”
 - Predictive models for planning
 - Decision support
 - Discovering new associations between e.g. drugs and diseases, diseases and genes, metabolites etc.
- Problems
 - Ethical and legal obstacles to secondary use of clinical data
 - Multicentric big data approaches require structured data: again use case for natural language processing of clinical text

Big data and personalised medicine

Big data

“Once every assessment, dictated note or scanned document is processed through a natural language processor and stored with machine readable formats the possibilities are only limited by our imagination”

Ellen Harper, Cerner Corporation

Big data and personalised medicine

Personalized medicine

Big data and personalised medicine

Personalized medicine

- Huge increase of data expected (lab, imaging biomarkers)

“A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention”

Biomarkers Definitions Working Group

- Focus on molecular entities including proteins, genes, gene variants, metabolites – not core domain of SNOMED CT
- Inclusion of biomarker concepts into SCT would require significant content addition – currently no issue for IHTSDO
- Need for semantically precise clinical data in the PM context

Conclusions

Conclusions

- SNOMED CT is and will be relevant
- SNOMED CT and other terminologies are increasingly harmonised
- Clinicians are not the primary customers
- Introduction should be use-case driven
- Importance of text mining use cases
- Thoughtful use of resources for localization
- Future of biomedical research and health care will be data driven: better structuring of data indispensable

SNOMED CT - Einstieg in eine neue medizinische Terminologie

30 Jun 2014, 14-17h, Inselspital Bern, Pathologisches Institut, Hörsaal 7, Auditorium Langhans



**Stefan
Schulz**
Medical
University
of Graz
(Austria)

purl.org/steschu



Merci !

Slides will be made available at

Purl.org/steschu

Comments / Questions →

stefan.schulz@medunigraz.at