ODLS 2016

Workshop on Ontologies and Data in Life Sciences,
Sep 29-30, Halle (Saale), Germany

Qualitative assessment
of SNOMED CT annotations

Jose Antonio Miñarro-Giménez,
Catalina Martínez-Costa, Stefan Schulz

Institut für Medizinische Informatik, Statistik und Dokumentation,
Medzinische Universität Graz, Austria

Context of study: ASSESS-CT

- European support action on the fitness of purpose of SNOMED CT as a core reference terminology for the EU: www.assess-ct.eu

- Feb 2015 – Jul 2016

- Scrutinising clinical, technical, financial, and organisational aspects of reference terminology introduction

- Main recommendations:
  - "SNOMED CT is the best candidate for a core reference terminology for cross-border, national and regional eHealth deployments in Europe."
  - Must be part of an ecosystem of semantic assets

Annotation of clinical narratives

- Resources
    - Parallel corpus: 60 clinical text samples from 6 languages,
    - For each language:
      2 annotators * 40 samples →
      20 samples annotated twice
    - Annotators
        - trained by webinars
        - follow annotation guideline
- Comparing
    - SNOMED CT vs.
    - UMLS derived terminology

Nitroglycerin pump spray as required	387404004;385074009;225761000
Amantadine bds	372763006;229799001
Allopurinol 300 ½ tablet every other day (last dose on 20091130)	387135004;385055001;225760004
Mefenamic acid 500 mg up to 3x daily for pain in conjunction with	387185008;258684004;229798009;22253000
simultaneous administration of a drug to protect the stomach e. g.	79970003;416118004;373517009;69695003
Pantoprazole 40mg.	395821003;258684004
Torasemide bds	318034005;229799001
Melperone 50 mg p. m.	442519006;258684004;422133006
§ 7 Intact teeth are in the mouth.	11163003;245543004;123851003
Fractures are visible on the medians of Mandible and Maxilla	263172003;263156006;260528009
the fragments are dislocated.	123735002
Normal mucous membranes in mouth pharynx and on the larynx.	17621005;33044003;71248005
Hyoid and thyroid cartilage are intact.	21387005;52940008;11163003
Fragmental fractures of the two upper vertebrae of the cervical spine.	13321001;207984009;207983003
Otherwise the cervical spine is intact.	122494005;11163003
Oesophagus as well as trachea are torn at the lower end of the neck.	262793000;282459005;261122009;123958008

Principal quantitative results

Concept coverage [95% CI]	SNOMED CT	Alternative
Text annotations – English	.86 [.82-.88]	.88 [.86-.91]

Term coverage [95% CI]	SNOMED CT	Alternative
Text annotations – English	.68 [.64; .70]	.73 [.69; .76]

Inter annotator agreement Krippendorff's Alpha [95% CI]	SNOMED CT	Alternative
Text annotations	.37 [.33-.41]	.36 [.32-.40]

Krippendorff, Klaus (2013). Content analysis: An introduction to its methodology, 3rd edition. Thousand Oaks, CA: Sage.

Agreement map: text annotations

SNOMED CT

UMLS SUBSET
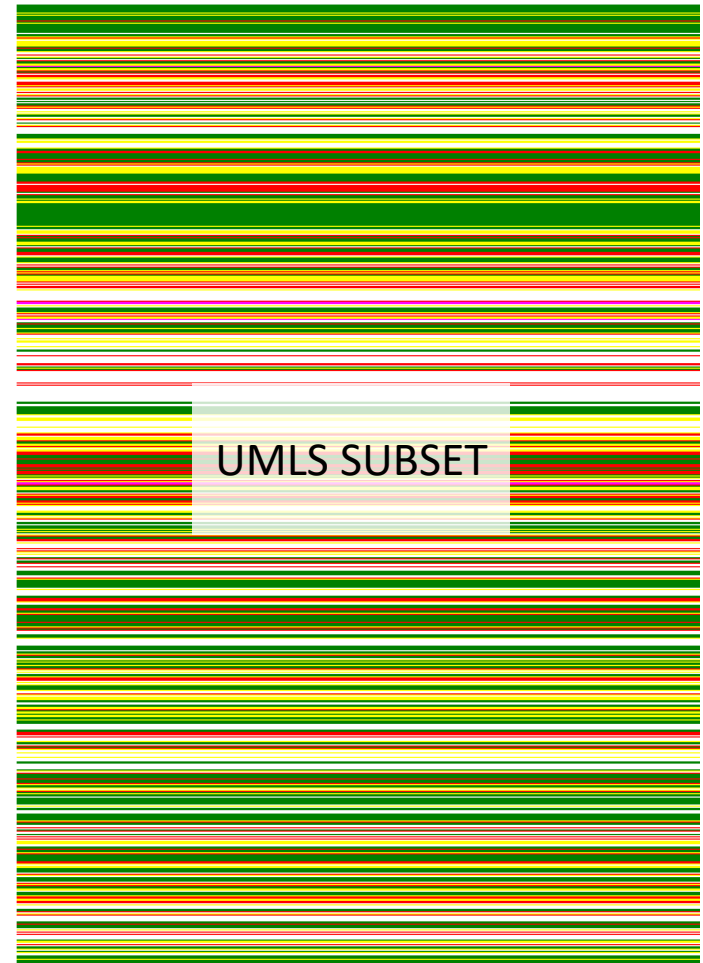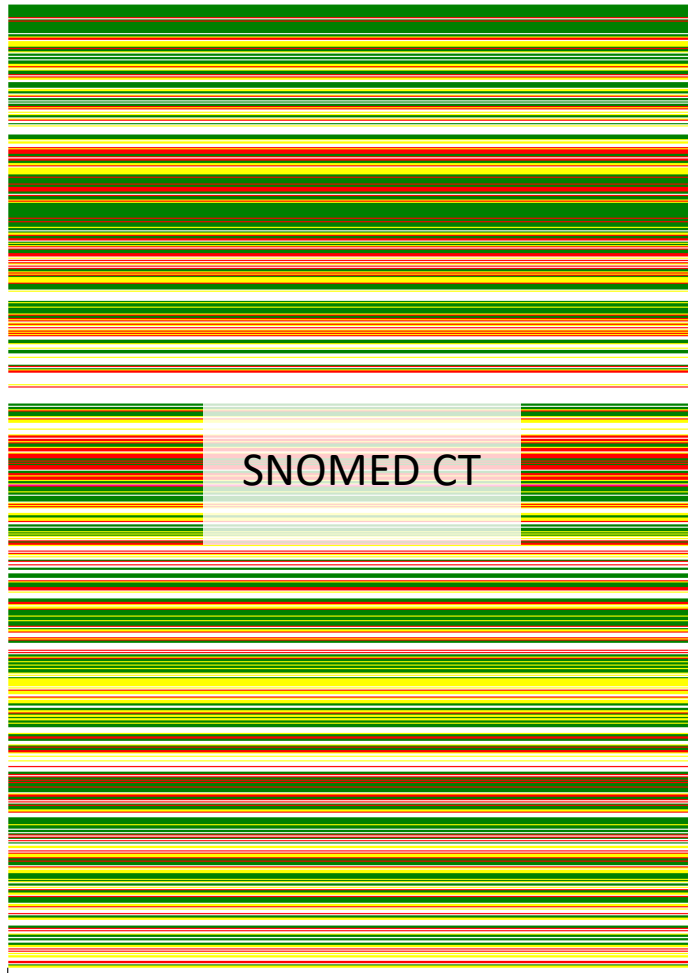
green: agreement   –   yellow: only coded by one coder   –   red: disagreement

Systematic error analysis

- Creation of gold standard for SNOMED CT
  - 20 text samples annotated twice →208 NPs
  - Analysis of English SNOMED CT annotations by two additional terminology experts
  - Consensus finding where disagreements, following pre-established annotation guidelines
    (e.g. maximally pre-coordinated concept)
- Inspection, analysis and classification of text annotation disagreements
- Ongoing study, first results presented here
- Presentation of disagreement cases for SNOMED CT

Reasons for disagreement

Human issues

- ## Lack of domain knowledge / carelessness

Tokens	Annotator #1	Annotator #2	Gold standard
IV	53240008 |Structure of abductor hallucis muscle (body structure)	80622005 |Abducens nerve structure (body structure)	80622005 |Abducens nerve structure (body structure)

- ## Retrieval error (no interface term found)

Tokens	Annotator #1	Annotator #2	Gold standard
Glibenclamide	384978002 | Glyburide (substance)	–	384978002 | Glyburide (substance)

Annotation guideline issues

- Underspecification
  - e.g. put anatomy concept always in procedure or disorder context

Tokens	Annotator #1	Annotator #2	Gold standard
IV	39322007 |Trochlear nerve structure	171674005 |Exploration of trochlear nerve (IV) (procedure)|	171674005 |Exploration of trochlear nerve (IV) (procedure)|

  - more general: avoid isolated primitive concepts

Ontology issues (I)

- ## Polysemy ("dot categories")*

Tokens	Annotator #1	Annotator #2	Gold standard
Lymphoma	118600007 |Malignant lymphoma (disorder)|	115244002 |Malignant lymphoma - category (morphologic abnormality)|	118600007 |Malignant lymphoma (disorder)|

- ## Pseudo-polysemy

Tokens	Annotator #1	Annotator #2	Gold standard
Former	410513005|In the past (qualifier value)	392521001 |History of (contextual qualifier) (qualifier value)|	8517006|Ex-smoker (finding)
Smoker	77176002|Smoker (finding)	77176002|Smoker (finding)	8517006|Ex-smoker (finding)

* A. Arapinis, L. Vieu: Complex categories in ontologies, FOIS 2014 Workshop on ontology and linguistics

Ontology issues (II)

- ## Incomplete definitions

Tokens	Annotator #1	Annotator #2	Gold standard
Diabetes	73211009 |Diabetes mellitus (disorder)|	170742000 |Diabetic monitoring (regime/therapy)|	170742000 |Diabetic monitoring (regime/therapy)|
monitoring	360152008 |Monitoring - action (qualifier value)|	170742000 |Diabetic monitoring (regime/therapy)|	170742000 |Diabetic monitoring (regime/therapy)|

- ## Navigational concepts (not for coding)

Tokens	Annotator #1	Annotator #2	Gold standard
palpebral fissure	301916005 |Finding of measures of palpebral fissure (finding)|	595000 |Structure of palpebral fissure (body structure)|	363934008 |Measure of palpebral fissure (observable entity)|

Ontological issues (III)

- ## Normal findings, no full definitions

Tokens	Annotator #1	Annotator #2	Gold standard
Motor:	127954009 |Skeletal muscle structure (body structure)|	106030000 |Muscle finding (finding)|	298300008 |Skeletal muscle normal (finding)|
normal bulk and tone	17621005 |Normal (qualifier value)|	17621005 |Normal (qualifier value)|	298300008 |Skeletal muscle normal (finding)|

- ## Fuzziness of qualifiers

Tokens	Annotator #1	Annotator #2	Gold standard
Significant	386134007 |Significant (qualifier value)|	24484000 |Severe (severity modifier) (qualifier value)|	6736007 |Moderate (severity modifier) (qualifier value)|
bleeding	131148009 |Bleeding (finding)|	131148009 |Bleeding (finding)|	131148009 |Bleeding (finding)|

Interface term issues

Tokens	Annotator #1	Annotator #2	Gold standard
Pain	406189006 |Pain observable (observable entity)|	22253000 |Pain (finding)|	22253000 |Pain (finding)|

"pain observations"

Tokens	Annotator #1	Annotator #2	Gold standard
Blood	87612001 |Blood (substance)|	50960005 |Hemorrhage (morphologic abnormality)|	50960005 |Hemorrhage (morphologic abnormality)|
extravasation	76676007 |Extravasation (morphologic abnormality)|		

"extravasation of blood"

Tokens	Annotator #1	Annotator #2	Gold standard
anxious	48694002 |Anxiety (finding)|	79015004 |Worried (finding)|	48694002 |Anxiety (finding)|

"anxious cognitions"

Language issues

- Ellipsis / anaphora
  - "Cold and wind are provoking factors as well." (provoking factors for angina)
  - "These ailments have substantially increased since October 2013" (weakness)
- Context
  - "No surface irregularities" (breast)
  - "Significant bleeding" (gastrointestinal bleeding)
  - "IV" (intravenous? Forth cerebral nerve? Type 4)
- Co-ordination:
  - "normal factors 5, 9, 10, and 11 "
- Negation
  - "no tremor, rigidity or bradykinesia"

Prevention of annotation disagreements

- Users (humans, text processing algorithms)
  - Training
  - Tooling
    - Guideline enforcement by appropriate tools
    - Post-co-ordination
  - Machine-processable annotation rules
  - Context awareness, scoping (e.g. looking back for anaphora resolution, identification of content of text passages)
  - Support by comprehensive, well-curated interface terminologies, tailored to the specific sublanguage to be analysed

Preventive measures (SNOMED CT structure)

- Fill gaps
  - equivalence axioms (reasoning)
  - Self-explaining labels (FSNs)
  - Scope notes where necessary (e.g. what means "entitic")
- Remove unnecessary ambiguity
- Flag concepts that should not be used (navigational concepts, anatomic "entire" concepts)
- Strengthen ontological foundations
  - Upper-level ontology alignment
  - Formalize constraints (SNOMED CT concept model)
  - Ontology / information model boundary
  - Overhaul problematic subhierarchies, especially qualifiers

Preventive measures
(SNOMED CT content maintenance)

- Include large-scale analysis of real data in routine maintenance process
  - Harvest notorious disagreements between notorious text passages and value sets with concepts
  - Compare concept frequency across institutions and users to detect imbalances
- Stimulate community processes for ontology-guided content evolution:
  - SNOMED CT ontological content
  - Interface terminologies for languages, specialties, users
  - Linking interface terminologies / value sets with SNOMED CT codes or expressions

- Contact: stefan.schulz@medunigraz.at