German-Language Content in Biomedical Vocabularies

Stefan Schulza,e, Josef Ingenerfb, Sylvia Thunc, Philipp Daumked, Martin Boeker e

aInstitute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria
bInstitute for Medical Informatics, University of Lübeck, Germany
cUniversity of Niederrhein, Krefeld, Germany
dAverbis GmbH, Freiburg, Germany
eInstitute of Medical Biometry and Medical Informatics, University of Freiburg, Germany
Biomedical Vocabulary

- **Umbrella term for**
  - Classification, Nomenclature, Terminology, Ontology, Catalogue

- **Building blocks**
  - **Representational units**: the primary identifiers of meaning (RUs, concepts, classes, synsets)
  - **Links**: the connections between RUs, such as broader/narrower, is-a, part-of
  - **Codes**: alphanumeric identifiers for a concept
  - **Terms**: words or group of words that describe the meaning of a RU in human language.
  - **Hierarchies**: network of links that constitute a partial order (single or multiple hierarchies)
  - **Axes**: systems of independent, non-overlapping hierarchies
  - **Glossary entries** (scope notes): natural language elucidations of concept meaning and use
  - **Axioms**: sentences expressed in logic providing formal, computable definitions for concepts
  - **Rules**: directives that specify exclusion and inclusion criteria in classifications
Genres of Biomedical Vocabularies

- Thesauri: describe the meanings of pre-existing terms and relate them using semantic relations
  - synonymy / broader...narrower

- Classifications:
  - Subclass hierarchies
  - tendency to artificial names
  - arbitrarily restricted meanings due to class disjointness

- Ontologies
  - logic-based foundation
  - concepts represent types of entities rather than meaning of words
<table>
<thead>
<tr>
<th>Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>■ ICD-10</td>
</tr>
<tr>
<td>■ OPS</td>
</tr>
<tr>
<td>■ ICD-O</td>
</tr>
<tr>
<td>■ MeSH</td>
</tr>
<tr>
<td>■ MEDDRA</td>
</tr>
<tr>
<td>■ LOINC</td>
</tr>
<tr>
<td>■ ATC</td>
</tr>
<tr>
<td>■ ICF</td>
</tr>
<tr>
<td>■ SNOMED /Wingert Nomenclature</td>
</tr>
<tr>
<td>■ RADLEX</td>
</tr>
<tr>
<td>■ EDQM</td>
</tr>
<tr>
<td>■ UCUM</td>
</tr>
<tr>
<td>■ ICPC</td>
</tr>
<tr>
<td>■ ICNP</td>
</tr>
<tr>
<td>■ Fee catalogues for procedure-based reimbursement</td>
</tr>
<tr>
<td>■ Sources for Latin terms</td>
</tr>
<tr>
<td>■ Domain-independent terminology sources</td>
</tr>
</tbody>
</table>
ICD 10

- Two German versions:
  - Mortality (WHO): 14,400 RUs
  - Morbidity (DIMDI): ICD-10 GM (12,500 concepts) extensions, pruning of subhierarchies

- Alphabetical Index: 77,000 terms linked to ICD-10 GM identified by Alpha ID

- Synonyms
  - more specific entry terms, e.g.
  "Amoxicillin-Allergie" [I112547] → T88.7 (Unspecified adverse effect of drug or medicament).
**OPS**

- Procedure coding system, only used in Germany and Austria
- Adaptation of the WHO's International Classification of Procedures in Medicine (ICPM)
- 28,300 concepts
- Alphabetical index
- Watch out: Abbreviated class labels

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-497</td>
<td>Transvenöse oder transarterielle Biopsie</td>
</tr>
<tr>
<td>1-497.0</td>
<td>Endokard</td>
</tr>
<tr>
<td>1-497.1</td>
<td>Endomyokard</td>
</tr>
<tr>
<td>1-497.2</td>
<td>Myokard</td>
</tr>
<tr>
<td>1-497.3</td>
<td>Leber</td>
</tr>
<tr>
<td>1-497.x</td>
<td>Sonstige</td>
</tr>
<tr>
<td>1-497.y</td>
<td>N.n.bez.</td>
</tr>
</tbody>
</table>
Indexing vocabulary for biomedical literature
(not recommended for EHR)

- 28,300 concepts (main headings)
- German version: and 60,067 German entry terms (compared to 193,000 English entry terms)
- No translation of scope notes
- Annual update (DIMDI)
SNOMED CT

- emerging international standard, maintained by the IHTSDO (International Health Terminology Standards Development Organisation), approx. 311,000 concepts, ontology based, biannual update
- German speaking countries so far no IHTSDO members
- Non-released German translation from 2004 (362,000 concepts, ~1M terms)
## Wingert nomenclature

- **Origin:** SNOMED v2
- **multiaxial nomenclature**
- **German translation (1984)** available as "Wingert nomenclature", maintained by private SW company

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>D5-46210</td>
<td>Acute appendicitis, NOS</td>
</tr>
<tr>
<td>D5-46100</td>
<td>Appendicitis, NOS</td>
</tr>
<tr>
<td>G-A231</td>
<td>Acute</td>
</tr>
<tr>
<td>M-41000</td>
<td>Acute inflammation, NOS</td>
</tr>
<tr>
<td>G-C006</td>
<td>In</td>
</tr>
<tr>
<td>T-59200</td>
<td>Appendix, NOS</td>
</tr>
<tr>
<td>G-A231</td>
<td>Acute</td>
</tr>
<tr>
<td>M-40000</td>
<td>Inflammation</td>
</tr>
<tr>
<td>G-C006</td>
<td>In</td>
</tr>
<tr>
<td>T-59200</td>
<td>Appendix, NOS</td>
</tr>
</tbody>
</table>
RadLex

- Vocabulary of standardized radiological terms for structured reporting in medical imaging
- 30,000 concepts
- 6,240 German translations
MedDRA

- Terminology for the drug regulatory process
- signs, symptoms, diseases, procedures, investigation results and therapeutic indications
- German translation: 92,000 German terms
- different kinds of synonymy relations
**ICNP - International Classification of Nursing Practice**

- Vocabulary for nursing diagnoses, interventions, and results
- 2800 concepts
- Multiaxial terminology system, increasingly ontology and logic-based.
- German version is maintained by the German ICNP User Group
Special cases

- Catalogues for reimbursement, drugs, devices
- Sources for Latin terms
- Domain-independent terminology sources
Catalogues

- Fee schemes for the ambulatory sector
- Terms + defined descriptions, exclusions inclusions
- Similar for devices (eClass) and drugs (ABDA)

34291 Herzkatheteruntersuchung mit Koronarangiographie
Obligator Leistungsinhalt
- Herzkatheteruntersuchung mit Koronarangiographie,
- Begleitleistungen, die im unmittelbaren Zusammenhang mit der Leistungserbringung stehen,
- Kontrastmitteleinbringung(en),

Fakultativer Leistungsinhalt
- Selektive Darstellung auch bei Patienten mit einem oder mehreren Bypässen und/oder bei Patienten mit Herzvitium,
- Angiokardiographie (Nr. 34290),
- Gerinnungsuntersuchung(en) (z. B. aktivierte Gerinnungszeit),
einmal im Behandlungsfall 8865 Punkte
Sources for Latin terms

- Use of Latin terms (with Latin inflections) in German text
  - Organisms ("Homo sapiens", "Drosophila melanogaster")
  - Anatomy ("Vena cava superior", "Papilla Vateri")

- International sources:
  - NCBI taxonomy for organisms
  - Terminologia Anatomica (Latin and English)
Lexical German resources

- GermaNet
  - Lexico-semantic network, similar to WordNet under construction
  - 85,000 concepts (synsets), 111,000 terms, 100,000 linkages

- Coverage not domain specific, but potential source for layperson terms

- Rubrics in medical lexicons
  - Pschyrembel
  - Roche Lexicon

- Wikipedia
General observations

- Localised versions modify the international version (ICD)
- No international version available (OPS)
- Localized version incomplete (RADLEX)
- Localized version outdated (SNOMED CT)
- Localized version unavailable (Wingert, SNOMED CT)
- Localized versions maintained by public body (DIMDI): ICD, MeSH
- Localized versions maintained by user groups (ICPM)

Additional source of terminology

- National catalogues (procedures, drugs, devices):
  - Latin Sources
  - Medical and general lexicons
Machine vs. Human Translation of SNOMED CT Terms

Stefan Schulz\textsuperscript{a,c}, Johannes Bernhardt-Melischnig\textsuperscript{a}, Markus Kreuzthaler\textsuperscript{a}, Philipp Daumke\textsuperscript{b}, Martin Boeker\textsuperscript{c}

\textsuperscript{a} Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria
\textsuperscript{b} Averbis GmbH, Freiburg, Germany
\textsuperscript{c} Institute of Medical Biometry and Medical Informatics, University of Freiburg, Germany

Abstract

Objective: In the context of past and current SNOMED CT translation projects we compare three kinds of SNOMED CT translations from English to German by: (i) professional

Survey of current translation projects

Besides the US and UK versions, a complete (Latin American) Spanish translation of SNOMED CT is being maintained.