A State-of-the Art Review of SNOMED CT Terminology Binding and Recommendations for Practice and Research

Anna Rossander¹  Lars Lindsköld¹  Agneta Ranerup¹  Daniel Karlsson²

¹Department of Applied Information Technology, University of Gothenburg, Göteborg, Sweden
²eHealth and Structured Information Unit, National Board of Health and Welfare, Stockholm, Sweden

Address for correspondence  Anna Rossander, Department of Applied Information Technology, University of Gothenburg, 412 96 Göteborg, Sweden (e-mail: anna.rossander@gu.se).

Methods Inf Med 2021;60:e76-e88.

https://doi.org/10.1055/s-0041-1735167
Agenda:

- poll on audience
- walk through how study was performed and overview of results
- poll on what order to discuss the recommendations
- Please ask questions as we go, either by raising hand or typing in chat.
Aim:

We aim to collate existing knowledge on difficulties and possible solutions and to point out relevant future research. The primary audience is practitioners and researchers working with terminology binding.
What is terminology binding?

Establishing links between elements of a terminology and an information model is called terminology binding and is also often referred to as mapping or subset development.

Benson T, Grieve G. Principles of Health Interoperability: SNOMED CT, HL7 and FHIR. Springer International Publishing; 2016
allergies

473011001 | Allergic condition (finding) |

meaning binding

mapping

list of terms in existing system
- nuts
- apples
- egg
- guinea pig

230034002 | Nuts and seeds (substance) |
735215001 | Apple (substance) |
102263004 | Eggs (edible) (substance) |
703927007 | Guinea pig dander (substance) |
PubMed, Scopus, Web of Science:
- snomed
- (map* OR "terminology binding" OR "information model" OR implement* OR subset*)
- NOT ICD*
55 papers published 2004 – 2020

53 papers from Information System outlets

2 papers from medical journals

53% of 1st authpurs women

Projects varied in size: 10 – 37,000 terms
All four authors read 20 randomly selected papers

-> developed topics in picture

Webster J, Watson RT. Analyzing the past to prepare for the future: writing a literature review. Manage Inf Syst Q 2002;26 (02):xiii–xxiii
Stated competence of project participants

- Domain expert: 7
- Terminologist: 8
- Both: 18
- Not stated: 22
terminology browsers:
- CliniClue Browser
- Snomed Internationals browser
- Gephi
- HealthTerm (browser only)
- Nictiz Terminology Browser
- UMLS Metathesaurus
- TermViz (prototype)

terminology binding tools:

<table>
<thead>
<tr>
<th>Name of tool</th>
<th>References</th>
<th>Publication year</th>
<th>Stated use in project</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioPortal Annotator</td>
<td>54</td>
<td>2013</td>
<td>a tool that processes text submitted by a user, recognizes relevant biomedical terms in the text, and returns the annotations to the user.</td>
</tr>
<tr>
<td>eleMAP</td>
<td>55</td>
<td>2011</td>
<td>semi-automatic mapping of research DES [Data Elements] to standardized biomedical vocabularies and metadata registries.</td>
</tr>
<tr>
<td>LexValueSets</td>
<td>56</td>
<td>2008</td>
<td>context-driven value sets extraction.</td>
</tr>
<tr>
<td>Mayo Clinic Tools SAVS/MCSV</td>
<td>57, 58</td>
<td>2004, 2006</td>
<td>The MCVS is a set of tools ... which facilitates health vocabulary indexing.</td>
</tr>
<tr>
<td>Medical Text Extraction, Reasoning, and Mapping System (MTERMS)</td>
<td>59</td>
<td>2016</td>
<td>generic natural language processing (NLP) application ... to process and map local allergy entries to the standard terminologies.</td>
</tr>
<tr>
<td>MoST</td>
<td>31</td>
<td>2008</td>
<td>archetype authoring, semi-automatic SNOMED CT terminology binding assistance and terminology visualization</td>
</tr>
<tr>
<td>SAMT/SSMT</td>
<td>60</td>
<td>2014</td>
<td>... discover mappings between clinical terms and SNOMED-CT concepts.</td>
</tr>
<tr>
<td>Snapper</td>
<td>33, 34</td>
<td>2010, 2011</td>
<td>... creating mappings from an existing terminology to SNOMED CT.</td>
</tr>
<tr>
<td>UMLS MetaMap</td>
<td>28</td>
<td>2010</td>
<td>... map biomedical text to the [UMLS] Metathesaurus.</td>
</tr>
<tr>
<td>Unnamed tool</td>
<td>67</td>
<td>2012</td>
<td>... propose a hybrid approach relying on linguistics as well as structural information for mapping.</td>
</tr>
</tbody>
</table>
• input material
  • local terms, context dependence, ‘other/unspecified’

• SNOMED CT
  • next slide

• information models
  • need for stated information model to share burden of terminology binding and to share information

• tools
  • lack of tools, different tools give different results
• SNOMED CT
  • failure to find concept mentioned 60% of papers
  • lack of terms
  • incorrect concepts/terms
  • concept model issues in 50% of papers
    • could not model concept needed
    • choice of hierarchy difficult
    • incorrect subsumption
    • navigational concepts both good and bad
    • boundary problem (information model or terminology?)
  • solutions stated that show lack of knowledge about SNOMED CT
Validation

60% of papers described validation
- manual: independent reviewers
- supported by tools: automatic controls, “results presented to clinical modeler”
- evaluation of applicability 29%
- descriptions of implementations 29%
- recommendations 22%
- development of tools 20%
1. Ensure Domain Knowledge and Informatics Competence
2. Follow a process including validation
3. Plan for maintenance
General Process

RECOMMENDATION FOR PRACTICE: Sort input data, select relevant concepts, validate.

RECOMMENDATION FOR RESEARCH: Test and refine the process described above. Evaluate balance between keeping legacy structures and resolving technical debt.
Understanding the Meaning of Terms

RECOMMENDATION FOR PRACTICE: Involve domain experts and invest time in educating them in informatics and SNOMED CT.

RECOMMENDATION FOR RESEARCH: Design supportive tooling geared to domain experts without in-depth knowledge about informatics or SNOMED CT.
Unspecified Terms
The problems reported regarding explicitly unspecified terms in the input material cover two principally different types of entities, not otherwise specified (NOS) and not elsewhere classified (NEC).  

RECOMMENDATIONS FOR PRACTICE: Use general concepts for nonspecified terms. Enable free text entries for situations where the subset might be incomplete, and data must be entered.

RECOMMENDATIONS FOR RESEARCH: Find effective processes for analysing data entered as free text to either bind it to an existing concept in the subset or add the requested concept to the subset.
5888003 | Keratitis (disorder) |
MINUS all of the other options in list
Incorrect Terms

RECOMMENDATION FOR PRACTICE: Report incorrect terms to National Release Centre (NRC) or SNOMED International. Develop local extensions when needed.

RECOMMENDATION FOR RESEARCH: Design processes for managing reported incorrect terms considering both promptness and quality.
Lack of Terms or Concepts

RECOMMENDATION FOR PRACTICE: Engage in relevant CRGs. Request new content as needed.

RECOMMENDATION FOR RESEARCH: Design supportive software for authoring of content close to the implementation setting. Continue developing maps and demarcations between terminologies.
Using Precoordinated Concepts or Post-coordinating or Using the Information Model

<table>
<thead>
<tr>
<th>body site</th>
<th>body site</th>
<th>body site</th>
</tr>
</thead>
<tbody>
<tr>
<td>left hand</td>
<td>hand: laterality = left</td>
<td>left</td>
</tr>
</tbody>
</table>

**pre-coordinated**  **post-coordinated with SNOMED CT**  **two separate parts of an information model**

Fig. 5 Using precoordinated concepts or post-coordinating or using the information model.

**RECOMMENDATION FOR PRACTICE:** Be consistent regarding what to store with terminology and what to store with an information model. Use the same demarcation as those with whom you will share information if possible.

**RECOMMENDATION FOR RESEARCH:** Develop tooling to facilitate post-coordination and comparison of pre- and post-coordinated SNOMED CT content. Compare different demarcation lines between terminology and information model in the search for an optimal compromise.
Choice of Information Model

RECOMMENDATION FOR PRACTICE: It is beyond the scope of this paper to recommend an information model. Prioritise internal information structure.

RECOMMENDATION FOR RESEARCH: Evaluate feasibility of DCMs and similar solutions.
Intentional Subsets Do Not Meet Expectations

RECOMMENDATION FOR PRACTICE: Manually validate subsets developed with ECL.

RECOMMENDATION FOR RESEARCH: Participate in enhancing the SNOMED CT concept model to improve ECL searches. Develop methods to minimise primitive content.
Proficiency with SNOMED CT

RECOMMENDATION FOR PRACTICE: Participate in education and user for a develop targeted education toward different types of users.

RECOMMENDATION FOR RESEARCH: Design supportive tooling geared to domain experts without in-depth knowledge about informatics or SNOMED CT.
Tools

RECOMMENDATION FOR PRACTICE: Use dedicated tooling where such exists.

RECOMMENDATION FOR RESEARCH: Design methods and tools suited for supporting terminology binding.
Thank you for listening!