Strategies for Updating Terminology Mappings and Subsets using SNOMED CT®

John Mapoles, Ph.D., Corey Smith, M.S., Jane Cook, R.N., Brian Levy, M.D. Health Language, Inc.
SNOMED CT® provides broad coverage of the medical domain
  • 280,000 active concepts.
Because of this size applications will use collections of SNOMED concepts rather than exposing all of SNOMED to users
Collections will be used for
  • drop-down lists
  • reference sets, e.g. non-human subset
  • mapping sets to reference terminologies
Collections of SNOMED concepts

- Lists or subsets
  - flat list of terms or concepts
  - more complex subsets based on
    - taxonomy
    - descendants of a concept or concepts

- Mappings to and from other standards
  - For example ICD9-CM or ICD10
  - Distributed by IHTSDO, HLI, and others
SNOMED updates

- Applications which support SNOMED implementation must also support SNOMED updates twice a year
- When SNOMED updates collections need to be changed to remain in step with SNOMED
- The update process for production material is time dependent
SNOMED updates

- SNOMED has evolved at a fairly constant rate
- Changes to concepts
  - Add and retire
- New terms are added with new concepts but also to existing concepts
- Remodeling: additions and deletions of relations
- Status changes
How has SNOMED changed, 2006 - 2008

- Magnitude of change over the last 5 versions
  - Constant influx of new concepts and retirement
    - Active concepts have remained at about 280,000
  - Large number of new terms
  - Also changes in hierarchy not shown here

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2006 1 31 - 2006 7 31</td>
<td>2112</td>
<td>1035</td>
<td>16963</td>
<td>11448</td>
</tr>
<tr>
<td>2006 7 31 - 2007 1 31</td>
<td>3029</td>
<td>1363</td>
<td>11725</td>
<td>3444</td>
</tr>
<tr>
<td>2007 1 31 - 2007 7 31</td>
<td>2315</td>
<td>907</td>
<td>17553</td>
<td>12124</td>
</tr>
<tr>
<td>2007 7 31 - 2008 1 31</td>
<td>2065</td>
<td>1071</td>
<td>7854</td>
<td>3077</td>
</tr>
</tbody>
</table>
How has SNOMED changed, 2006 - 2008

SNOMED Changes 2006 - 2008

- New Concepts
- Retired Concepts
- New Terms
- New Terms Existing Concepts

Date Ranges

Number of Changes

- 2006 1 31 - 2006 7 31
- 2006 7 31 - 2007 1 31
- 2007 1 31 - 2007 7 31
- 2007 7 31 - 2008 1 31
Updating SNOMED collections

- Flat lists are updated most easily
  - Retired and new concepts must be accounted for
  - Mappings can be updated similar to lists but both the source and target may change
- Subsets and other structures that leverage the SNOMED hierarchy must account for changes made in the hierarchy
  - Careful attention must be given to where concepts have moved and what has moved into the hierarchy
Updating lists, retired concepts

- All retired concepts must be removed (or flagged) and considered for replacement
  - The SNOMED historical relations are very valuable
Updating lists, new concepts

- New concepts must be considered for inclusion
- Concepts should be sorted by a ranking algorithm
  - Algorithm must reflect the requirement of the list so that the most likely candidates are considered first
  - SNOMED hierarchy can be valuable in ranking
The MSS consist of 45 subsets containing 10,000 SNOMED terms.

MSS subsets were built in pairs - one subset for the diagnoses related to a medical specialty and the other for the procedures.

Exception psychiatry

The top approximately 150 diagnoses or procedures applicable to a specialty were incorporated into a subset.

The subsets contain “clinically friendly” SNOMED terms most often encountered in clinical practice.

Specific to the level of granularity commonly needed by clinicians in that specialty practice.

Built in the HLI database LE® and managed in LExScape®
Asthma, unspecified

Bronchial structure

Disease of bronch... Respiratory obstr...

Coalworker's pneu...

Courses

Episodicities

Acute asthma

Byssinosis

Exercise-induced...

Mixed asthma

Asthma attack

Cardiac asthma

Extrinsic asthma

Pneumopathy due...

Asthmatic bronchitis

Chemical-induced...

Hay fever with asth...

Weavers’ cough

Asthma unspecified

Childhood asthma

Intrinsic asthma

Asthma without st...

Drug-induced asth...

Late onset asthma

Brittle asthma

Exacerbation of as...

Millers’ asthma

Taxonomies: SNOMED CT Clinical Finding; Disease; Finding by site
All 23 MSS areas of specialty, each with a disease and procedure subset, except psychiatry

<table>
<thead>
<tr>
<th>Subset</th>
<th>Subset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergy/Immunology</td>
<td>Infectious Disease</td>
</tr>
<tr>
<td>Anesthesiology</td>
<td>Internal Medicine</td>
</tr>
<tr>
<td>Cardiology</td>
<td>Neurology</td>
</tr>
<tr>
<td>Critical Care</td>
<td>Neurosurgery</td>
</tr>
<tr>
<td>Dermatology</td>
<td>Obstetrics</td>
</tr>
<tr>
<td>Endocrinology</td>
<td>Oncology</td>
</tr>
<tr>
<td>Family Practice</td>
<td>Ophthalmology</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>Orthopedics</td>
</tr>
<tr>
<td>General Surgery</td>
<td>Pulmonology</td>
</tr>
<tr>
<td>Gynecology</td>
<td>Rheumatology</td>
</tr>
<tr>
<td>Hematology</td>
<td>Urology</td>
</tr>
<tr>
<td>Psychiatry</td>
<td></td>
</tr>
</tbody>
</table>
Term requirements

- Disease subset members are taken from SNOMED Clinical Findings taxonomy.
- Procedure subset members are taken from SNOMED Procedure taxonomy.
- Terms must be active on an active concept: $(\text{STATUS} = 0)$.
- A subset can contain only one term from a concept.
- Concepts with a term in a disease subset must have a map to a billable ICD9-CM (IHTSDO cross-maps).
- Concepts with a term in a procedure subset must have a map to CPT (HLI cross-maps).
Updating the MSS, a list update

- **Strategy**
  - Each subset needs to be evaluated so that only the specific groups of changes that immediately affect the content are considered.
    - This analysis is essential for performing the update quickly and efficiently.
  - The most important changes reviewed first
Updating the MSS, a list update

- Change type causing terms to be removed from a subset
  - Type - 1
  - Retired
  - Status change, status is no longer 0
  - Concepts that are no longer mapped to reference terminology or the target is not billable in the case of the disease subsets
  - Concepts is not longer in correct taxonomy

- Change types for term to be considered for inclusion into subset
  - Type - 2: New terms on concepts in the subset
  - Type - 3: New terms on concepts that are descendants of a concepts in the subset
    - These terms are likely to be closely related to the term in the subset
    - May be a good replacement or new term
  - Type - 4: New terms not in type-2 or type-3
    - All new terms must be considered
Updating the MSS

- Update analysis is done via the LE java APIs.
  - Accessing the MSS and latest versions of SNOMED
  - SNOMED concepts and MSS accessed using standard java code
- Java scripts build lists of concepts / terms of type-1, -2, -3, and -4.
  - Stored externally for review
  - or in LE microglossaries
- Analyzed by modelers in LExScape for updating MSS
  - Reviewed by modeler and QA
Updating the MSS

- Changes are reviewed in order
- Type-1 must be reviewed and repaired, they break the requirements of the subsets
- Type-2 and -3 follow
  - Type-2: these terms are on concepts that are in the subsets
  - Type-3: terms are of lesser interest but are descendants of concepts in the subsets
- Type-4 are reviewed last as a batch for all 45 subsets.
  - Reviewed only once.
## Sample data

<table>
<thead>
<tr>
<th>Subset Name</th>
<th># of Members</th>
<th>Type-1 Changes</th>
<th>Type-2 Changes</th>
<th>Type-3 Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical Care - Disease Subset</td>
<td>558</td>
<td>3</td>
<td>3</td>
<td>137</td>
</tr>
<tr>
<td>Neurology - Disease Subset</td>
<td>367</td>
<td>4</td>
<td>2</td>
<td>185</td>
</tr>
<tr>
<td>Gastroenterology - Procedure Subset</td>
<td>185</td>
<td>6</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Neurosurgery - Procedure Subset</td>
<td>327</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subset Name</th>
<th>Type-1</th>
<th>Type-2</th>
<th>Type-3</th>
<th>Type-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical Care - Disease Subset</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Neurology - Disease Subset</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Gastroenterology - Procedure Subset</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Neurosurgery - Procedure Subset</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
Conclusions

- Updates are post-production and under the time constraint that new data must be available as soon as possible
- All changes cannot be reviewed for each entry
  - The changes relevant to the subset must be identified
- Rank changes in order of importance
- Use automated processes where possible
- Leverage SNOMED hierarchies
Complexity of hierarchy

- Lists or subsets that are based on the SNOMED hierarchy are valuable because they leverage the work done by SNOMED
  - The SNOMED hierarchy can be used as the local definition of a disease or a group of concepts
  - Example: All descendants of Diabetes mellitus (SNOMED: 73211009) are considered to be diabetes and trigger a system response to a diabetic patient
- When the descendants of the defining concept are remodeled a decision must be made
  - To include other targets in the list definition
Complexity of hierarchy

Old version

```
List
All descendants of C1
```

C1

New version

```
List
All descendants of C1
```

C1
Further complexities

• Excludes: subsets that support inheritance may consider supporting excludes so that only part of a concept’s descendants are included in a subset
  • This leads to great complexity during updates because remodeling of the hierarchy can move excluded concepts as well as included concepts
• History: maintaining concepts that are out of scope in the list may be valuable for future reference
  • concepts need to be flagged with dates or date ranges