Translating SNOMED CT

Tutorial Phoenix Arizona, May 31th 2008
Asta Høy and Ulrich Andersen
Presentation

• Asta Høy
  – Terminologist, PhD
  – Consultant with the Danish National Board of Health, Socialstyrelsen, Sweden and INFOWAY, Canada.

• Ulrich Andersen
  – MD, MPA
  – Initiator of Danish Translation Project. Inventor of the translation workflow. Established the IHTSDO as a Danish association in Copenhagen.
  – Consultant in SNOMED CT implementation and Translation
Agenda

- Importance of Concept based translation
- Web based application
- Machine translation?
- Term requirements
- Hands on: What does it mean?/What’s the difference?
- Break
- Solutions to exercise
- Process - Workflow - Management Guidelines
- Linguistic Guidelines
- Hands on: Add synonyms
- Closing discussion

uan@SORANO.net
Why choose to translate SNOMED CT®?

- All health care professionals must be able to use and understand the terminology
- Language must fit the culture of each individual country
- English is not sufficiently used in either writing or speech within the health professional domains
- Patients and relatives must be able to understand the terminology used
What will be translated?

- Start with the national need for concepts
- Decide on one or more IT architectures
- Choose related concepts in subsets for translation
- Translate one term for each concept
- Be aware of the fact that historical data contained in SNOMED CT® often corresponds to concepts not in current use
- Be aware that limiting concept selection is resource-intensive and carries risk

uan@SORANO.net
How do you put together the entire team?

- Obtain competencies within: health care, medical informatics, project management and administration, language, terminology and SNOMED CT*
- Partner with an IT supplier with focus on process support and workflow management
- Establish a management group with political support
- Establish an editorial group with health care professional and terminological competencies
- Develop a close partnership and coordination effort between the organisation responsible for the project and the supplier
Which tools will the team use?

- Develop a web-based IT system to support translation workflow processes
- Require that the IT system can: distribute the process, maintain the history, support translation workflow processes, be accessed quickly and easily
- Require that the IT system is closely connected to terminology maintenance, extension, distribution, subset management and mapping to other classifications
- Require that the IT system offers functionality that supports fast and accurate translation, including concordance searches
- Require that the IT system is transparent and can display progress and updated language guidelines for the translation
- Require that the IT system can distribute all processes involved in development and maintenance of the health care terminology
How to ensure high quality translation?

- Emphasise the importance of a concept-oriented approach
- Support interdisciplinary teamwork
- Ensure both clinical and terminological review
- Maintain a timely translation and review workflow to ensure quality continuity
- Involve clinical test departments for the final acceptance
- Facilitate the exchange of feedback between the translation agency and the organisation responsible for the project
Translation

- Language (text) level
- Referential level
  - Ogden's triangle
- Cohesive level
- Level of naturalness

{ Psychological Acceptability

uan@SORANO.net
Knowledge

• Tacit knowledge
• Explicit knowledge
  – Organization build on trust
  – Common language
    = Reference terminology
Concept based translation

• Language (text) level
• Referential level
• Cohesive level
• Level of naturalness

uan@SORANO.net
Caminos desiderata

I. Content

II. Concept Orientation

III. Concept Permanence

IV. Nonsemantic Concept Identifiers

V. Polyhierarchy

VI. Formal Definitions

uan@SORANO.net
II Concept orientation

• Concepts before terms
• Ogden-Richards
  TRIANGLE
• Concept (thought)
• Referent (in the world)
• Symbol (term)

a rose

uan@SORANO.net
SNOMED CTs Triangle

SCT_CONCEPTS
- CONCEPTID
- FULLY_SPECIFIED_NAME
- CTVID
- SNOMED_ID
- IS_PRIMITIVE
- CONCEPT_STATUS

SCT_DESCRIPTIONS
- DESCRIPTION_ID
- DESCRIPTION_STATUS
- CONCEPT_ID
- TERM
- INITIAL_CAPITAL_STATUS
- DESCRIPTION_TYPE
- LANGUAGE_CODE
Concept based translation

• Meaning of Concepts rather than “word to word translation”

• Eg, IRIS
  – as in the concept of IRIS the anatomical structure
  – And
  – IRIS as the flower.
Content

• Snomed CT has a content that has been build up by clinicians over more than three decades.

• The content is balancing between clinical demands and systemacy.

• The content has the potential for transfer of knowledge between different health-professions.
Concepts based translations ABC

• Understanding concept
• **Sequence of Basics: A; B; C; A**

• Constructing term
• **Sequence of Basics: B; C; A**

Concept based translation Basics
A: Term reading/writing
B: Is-a positioning
C: Associative relationships assessment

uan@SORANO.net
Use case for Product (terms)

- Fully Specified Name (FSN)
  - Theoretical approach

- Preferred terms (PT)
  - Near to Clinical users

```
<table>
<thead>
<tr>
<th>SCT_CONCEPTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CONCEPTID</td>
<td>79</td>
</tr>
<tr>
<td>FULLYSPECIFIEDNAME</td>
<td>A</td>
</tr>
<tr>
<td>CITATIONAL</td>
<td>A</td>
</tr>
<tr>
<td>NORMED</td>
<td>A</td>
</tr>
<tr>
<td>ISPRIMITIVE</td>
<td>79</td>
</tr>
<tr>
<td>CONCEPTSTATUS</td>
<td>79</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SCT_DESCRIPTIONS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DESCRIPTIONID</td>
<td>79</td>
</tr>
<tr>
<td>DESCRIPTIONSTATUS</td>
<td>79</td>
</tr>
<tr>
<td>CONCEPTID</td>
<td>79</td>
</tr>
<tr>
<td>TERM</td>
<td>A</td>
</tr>
<tr>
<td>INITIALCAPITALSTATUS</td>
<td>79</td>
</tr>
<tr>
<td>DESCRIPTIONTYPE</td>
<td>79</td>
</tr>
<tr>
<td>LANGUAGECODE</td>
<td>A</td>
</tr>
</tbody>
</table>
```

uan@SORANO.net
Why not machine translation?

- The ontology model
- Terms generation
- Spanish Translation
- Translation memory
- [http://demo.healthterm.com/](http://demo.healthterm.com/)
- Advice

uan@SORANO.net
The model

<table>
<thead>
<tr>
<th>Concept</th>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>angina</td>
<td>Is a</td>
<td>Heart Disease</td>
</tr>
<tr>
<td>angina</td>
<td>Is a</td>
<td>Disorder characterized by pain</td>
</tr>
<tr>
<td>angina</td>
<td>Finding site</td>
<td>Heart Structure</td>
</tr>
<tr>
<td>angina</td>
<td>Definitional manifestation</td>
<td>Pain</td>
</tr>
<tr>
<td>angina</td>
<td>Severity</td>
<td>Severities</td>
</tr>
<tr>
<td>angina</td>
<td>Episodicity</td>
<td>Episodicities</td>
</tr>
<tr>
<td>angina</td>
<td>Course</td>
<td>Courses</td>
</tr>
</tbody>
</table>
Term construction

• Each concept IS A something with distinct characteristics
• The terms of The precoordinated concepts could be a construct of the terms of its elements
• Similar to the problem of screen representation of postcoordinated concepts
Angina: Heart disease characterized by pain

<table>
<thead>
<tr>
<th>Concept</th>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>angina</td>
<td>Is a</td>
<td>Heart Disease</td>
</tr>
<tr>
<td>angina</td>
<td>Is a</td>
<td>Disorder characterized by pain</td>
</tr>
<tr>
<td>angina</td>
<td>Finding site</td>
<td>Heart Structure</td>
</tr>
<tr>
<td>angina</td>
<td>Definitional</td>
<td>Heart Structure</td>
</tr>
<tr>
<td></td>
<td>manifestation</td>
<td>Pain</td>
</tr>
<tr>
<td>angina</td>
<td>Severity</td>
<td>Severities</td>
</tr>
<tr>
<td>angina</td>
<td>Episodicity</td>
<td>Episodicities</td>
</tr>
<tr>
<td>angina</td>
<td>Course</td>
<td>Courses</td>
</tr>
</tbody>
</table>

Angina is a disorder [of the] heart structure [with] pain as definitional manifestation

uan@SORANO.net
Spanish Translation

- Pretranslation: Words –phrases-ad hoc memory -translation
- 20% of concepts could be constructed automatically
- 25% needed manual correction
- 55% manual translation process

Reynoso, G A et al: Development of the Spanish Version of Systematized nomenclature of Medicine: Methodology and Issues

uan@SORANO.net
Permutations

• A Caveat

fracture coccyx with 208082008
Advice!

• Start out with translation of the attribute values

• Use a translation memory (look up concordance)

• Create linguistic guidelines
★ Process Workflow Management

- Huge task
- Demands
- Quality

uan@SORANO.net
Process

Decision for translation
- 18 months
- 2 MDs

Design of Workflow
- Selection of subsets
  - Translation
  - Review
  - Accept for validation
  - Gant diagram

Design of web application
- Tenders were invited

Selection of translators
- Resources
  - 3 managers
  - Consultants
  - 3 months

Education
- Resources
  - SNOMED INTERNATIONAL
    - Danish National Board of Health
    - IT developer

Clinical acceptance
- Validations
Recommendations

- A platform for knowledge share
- Consensus
- Management
A platform for knowledge share

• Distributed process
• Multidisciplinary process
• Transparency

uan@SORANO.net
Consensus

• Multiple solutions are possible
  – Systematic approach
  – balance between systematic approach and psychological acceptability
Management

- Monitoring batch 26.2
- Feedback
- Stats

uan@SORANO.net