Development considerations and methods for content reference sets

C. Richardson, R. Richards, M. Cordell
National Clinical Terminology and Information Service (NCTIS) - NEHTA

General Considerations

Stakeholders
- Clients, customers and others impacted by the product or who need input into the product.
- What is the use case?
- Who will use the reference set?
  - Setting e.g. Primary Care, Specialist, Acute Care, Research
  - Users e.g. Clinicians, Allied Health, Researchers, Health Departments, Vendors
  - Is it locally or nationally applicable?
- What is required?
  - Scope
  - Size
  - Type of content e.g. clinical findings, procedures
  - Information model specifications to be taken into account?
  - Legislation or national standards to be considered?
- Why is the reference set required?
  - Direct patient care.
  - Secondary use e.g. Research, reporting, decision support

How will the reference set be used?
- User interface or the back end
- Messaging

Information and Terminology Model Considerations

Reference sets can be developed against information model specifications and bound to a data element within the information model. Where a reference set is developed for a specific data element consideration must be given to the following areas.

Information Model

Selected Value

Comment on alignment

Diagnosis
- (35566002) hernoma (morphologic abnormality)

- It is important to ensure reference set concepts are drawn from the correct hierarchy. In this instance the concept should actually be drawn from the Clinical finding hierarchy.

Family History
- (34000006) Crohn’s disease (disorder)

- Concepts selected should relate appropriately to the subject of record. In this instance a family member has Crohn’s disease not the subject of record.

Information model components and their attributes

Defining information can be inherent in some concepts, e.g. laterality of a body part. The information model may be developed to allow capture of this information as one value or separately. Reference set development must take this into account. Consideration must also include the attributes of a data element e.g. whether a data element is optional.

Information Model

Selected Example 1

Example 2

Anatomical Location
- 20335008 left foot
- 56459004 left

Laterality
- 7771000 (left or leave blank)
- 777100 (left)

Example 1: If the reference set for the Anatomical location data element allowed values which included laterality then this information could potentially be captured twice (semantic overlap). Users would need to check both fields for the information and there is potential for different information in each field.

Example 2: Concepts defined with laterality are excluded from the reference set for Anatomical location. If the Laterality data element was optional and was not included in an information system, information about laterality would be unable to be recorded.

Source Data Exclusion Method

Existing reference sets are used as mechanisms for excluding content in another reference set.

This is an automated method.

Example: Editorial rules were created to develop the SNOMED CT-AU relationship to subject of care reference set.

Source Data Mapping Method

Reference set content is determined by mapping source data such as a lists of terms or codesets to suitable SNOMED CT-AU concepts.

This can be a manual or semi automated method.

Example: The SNOMED CT-AU Musculoskeletal finding reference set was derived from the Clinical finding foundation reference set.

Source Data Inclusion Method

Existing reference sets are used as mechanisms for including content in another reference set.

This can be an automated method.

Example: The SNOMED CT-AU Musculoskeletal finding reference set was derived from the Clinical finding foundation reference set.

Attribute Method

Allowed attributes are used to select the hierarchy and concepts that should be included in the reference set.

This can be an automated and/or manual method.

Example: The Clinical finding hierarchy concepts can be defined using the FINDING Site attribute with the allowable values being drawn from (44038009 anatomical or acquired body structure) and its descendants.

Concept Enumeration Method

Inclusion and/or exclusion rules are developed for concept enumeration values appropriate to a certain field or combination of fields in the SNOMED CT-AU core file (tables) and/or structural reference sets.

This is an automated method.

Example: The active field in the concept file and the value field in an attribute value reference set can be selected. Automated rules are then applied for concept enumeration values that represent inactive content. Applying the rule would remove inactive content from a reference set.

Development Methods

Simple Inclusion Method

Hierarchies, sub hierarchies and individual concepts are selected for inclusion or exclusion following editorial rules and guidelines that have previously been developed.

This is a manual method.

Example: Editorial rules were created to develop the SNOMED CT-AU relationship to subject of care reference set.

The design rules were:
- Concepts to be drawn from the Person sub hierarchy of the Social context top hierarchy.
- Within the Person sub hierarchy select:
  - General categories of people sub hierarchy.
  - All descendants EXCEPT for |2600000.089778| (Strictly defined) and |2600000.090890| (Strictly defined).
  - Include |2600000.091919| (Guardians) and descendants.
  - Include |2600000.092929| (Relationship hierarchy).
  - Person in the family sub hierarchy.
  - All descendants EXCEPT for the concept (159341001 Adopted person).
  - Children of this concept are still to be included.
  - Within the Person in the healthcare environment sub hierarchy (159345006 for medical or surgical procedure) to be included.
  - Within the Person categorized by age sub hierarchy (133933007 Newborn) to be included.

- Excluded items:
  - The concept and descendants of |2600000.093933| (member of public) to be excluded, except for |2600000.094944| (member of public involved incidentally).
  - Person in family sub hierarchy due to duplication of concepts from the Person in the family sub hierarchy and also due to duplication of the phrasing ‘relationship to subject of care’ that exists with the data element.

- Added conditions during development:
  - Exclude descendants of |105429005| (elderly persons) and |105429006| (elderly persons) because these are plural and data element indicates another individual.
  - Other considerations:
    - One potential area not considered is that of ‘carers’. There is a Caregiver sub hierarchy that can be considered for this.
    - There are no Person sub hierarchy concepts that address requirements for in situ. Further review required here.

"SNOMED CT® and SNMTS® are registered trademarks of the International Health Terminology Standards Development Organization."