

Development considerations and methods for content reference sets

C. Richardson[#], R. Richards, M. Cordell

National Clinical Terminology and Information Service (NCTIS) - NEHTA

[#] Corresponding author: Cathy Richardson Email: cathy.richardson@nehta.gov.au

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

What existing data sources are currently available?

- Nationally or locally
- Suitable for purpose
- Support development of a new reference set
- A terminology currently used for a similar purpose.

SNOMED CT-AU

- Concept model constraints
- Relevant hierarchy/ies
- Current content available
- Level of granularity required
- New content requirements
- Dependencies on other development work e.g. another reference set, a content development project at the national or international level

Development and Maintenance

- Development method/s - more than one may be required for development.
- Rules and/or development instructions
- The module the reference set will reside in.
- Which SNOMED CT^{®1}-AU release will be used for development.
- Metadata requirements
- Will the reference set contain concepts or descriptions?
- Reference set type required e.g. Attribute value, Mapping
- Mandated constraints: Compliance with RF2, naming conventions
- Amount of noise that will be accepted
- Quality assurance processes
- Maintenance requirements

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.

Interaction between the information and terminology models

Information Model	Selected Value	Comment on alignment
Diagnosis	[35566002 <i>hematoma (morphologic abnormality)</i>]	It's important to ensure reference set concepts are drawn from the correct hierarchy. In this instance the concept should actually be drawn from the <i>Clinical finding</i> hierarchy.
Family History	[34000006 <i>Crohn's disease (disorder)</i>]	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	Example 1	Example 2
Anatomical Location	[22335008 <i>left foot</i>]	[56459004 <i>foot</i>]
Laterality	[7771000 <i>left</i>] (or leave blank)	[7771000 <i>left</i>]

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.

Development Methods

Attribute Method

Allowable 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 [442083009 *Anatomical or acquired body structure*] and its descendants.

Using this attribute *Clinical finding* concepts with a FINDING SITE of [80891009 *Heart structure*] and its descendants could be selected as content of a reference set.

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 valueID 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.

Source Data Exclusion Method

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

This is an automated method.

Example: The NCTIS has a Non-human reference set that contain concepts which are not relevant for humans e.g. [15998008 *Animal hair*]. These have been excluded from the SNOMED CT-AU reference sets.

Source Data Inclusion Method

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

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 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 *Dose unit reference set* was primarily developed using this method.

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.

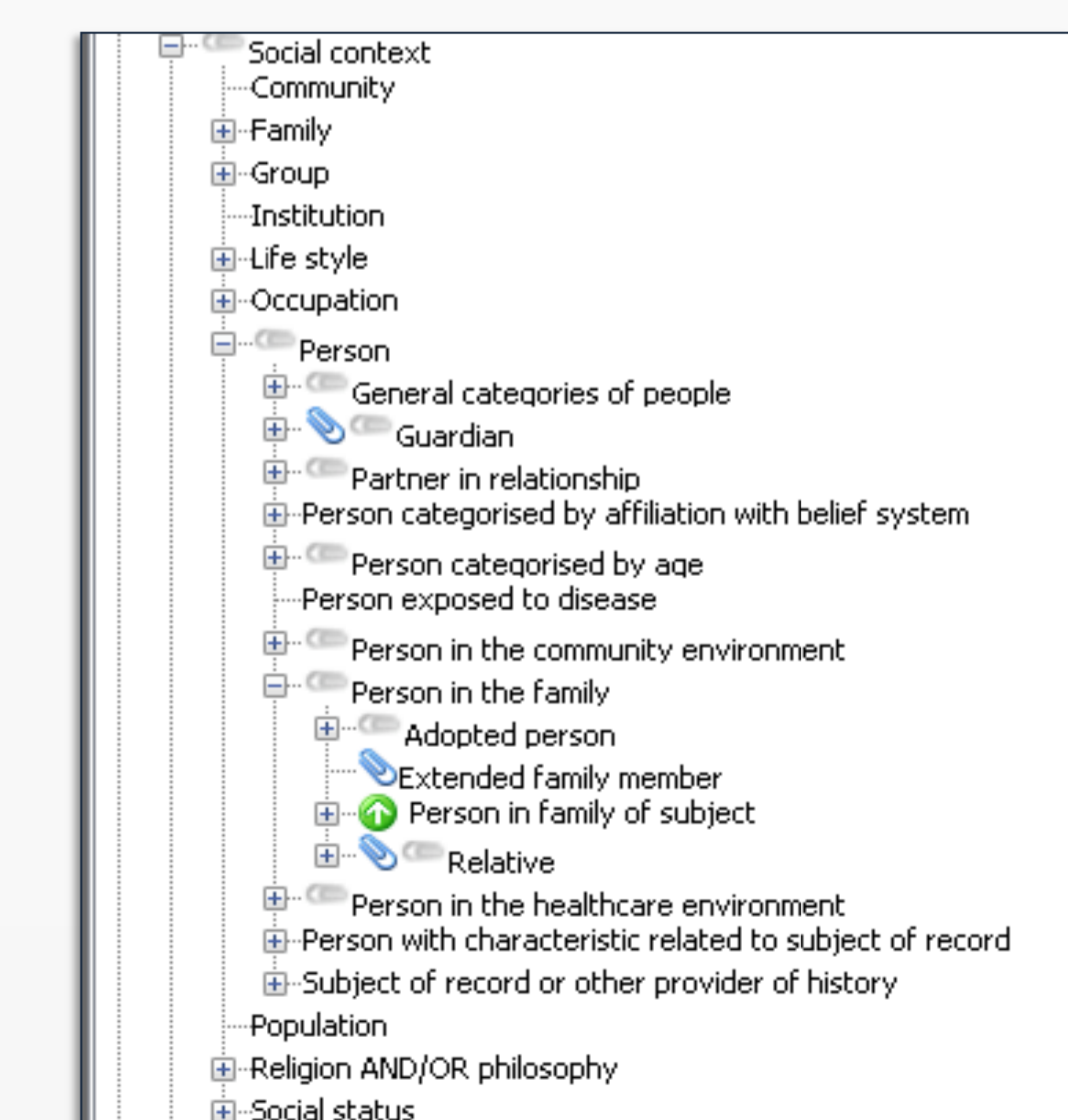
Example: Editorial rules were created to develop the SNOMED CT-AU *Relationship to subject of care reference set*.

The design rules were:

- Concepts to be drawn from the *Person* sub hierarchy of the *Social context* top level hierarchy.
- Within the *Person* sub hierarchy select:
 - General categories of people* sub-hierarchy. All descendants EXCEPT for [224619008 *Migrant*], [310543008 *Graduate*] and [385435006 *Individual*].
 - Include [394619001 *Guardian*] and descendants.
 - Include *Partner in relationship* sub-hierarchy
 - Person in the family* sub-hierarchy: All descendants EXCEPT for the concept [105431001 *Adopted person*]. Children of this concept are still to be included.
 - Within the *Person in the healthcare environment* sub-hierarchy [105455006 *Donor for medical or surgical procedure*] to be included.
 - Within the *Person categorized by age* sub- hierarchy [133933007 *Newborn*] is to be included.
- Excluded items:
 - The concept and descendants of [257513009 *Member of public*] are to be excluded, except for [257514003 *Member of public involved incidentally*].
 - Person in family of subject* 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 name.
- Added conditions during development:
 - Exclude descendants of [105449001 *sick relative*] because specific disease contexts such as [demented relative] are not needed for this data element.
 - Exclude [105429005 *elderly parents*] and [105428002 *homosexual parents*] because these are plural and data element indicates an individual.

Other considerations:

- One potential area not covered 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 foetus. Further review required here.



Blue paper clips represent members and gray paper clips represent parents of members.

Source ID	Source Description	Target Concept ID	Target Description	Match type
C0287	Acute MI	[57054005]	<i>acute myocardial infarction</i>	Equivalent
D0025	Left fractured NOF	[5913000]	<i>fracture of neck of femur</i>	Specialised
D0021	Right fractured NOF	[5913000]	<i>fracture of neck of femur</i>	Specialised

¹ SNOMED CT[®] and IHTSDO[®] are registered trademarks of the International Health Terminology Standards Development Organisation.