SNOMED CT Reference Set Tooling

RFP202011.2 - Functional Specification

v1.0

2020-11-02

SNOMED International
Registered in England and Wales ⋄ Company Registration Number 9915820
Reg. address: One Kingdom Street ⋄ Paddington Central ⋄ London W2 6BD ⋄ United Kingdom
Tel: +44 (0) 203 755 0974 ⋄ info@snomed.org ⋄ www.snomed.org

SNOMED International is the trading name of the
International Health Terminology Standards Development Organisation,
a private company limited by guarantee

Table of Contents

1. Introduction

This document provides the functional specification for a major new version of the solution provided by SNOMED International for the creation, management, and distribution of reference sets.

Functionality and requirements are based on experience and learning from the current production tooling already in use, as well as input from stakeholders. This document is intended to start the discussion and provide a starting point for these requirements, before opening up a Request For Proposal process.

Reader note: *Whilst the term reference set or refset is used throughout the document for sake of brevity, this refers to SNOMED CT reference sets, HL7 value sets, free sets and local or simple subsets, unless explicitly stated.*

1.1. SNOMED CT Translation Tooling

The current live SNOMED International Reference Set tool also includes functionality to aid translation of SNOMED CT content. For this new version of reference set tooling, translation is not considered to be part of this solution and therefore is not covered in this document.

2. Functional Requirements

Functional requirements are broken into the following categories:

- Reference set specific functionality
- Content collaboration functionality
- Out of scope functionality

2.1. Reference Set Functionality

The proposed solution/service should meet all of the following requirements

2.1.1. Reference set directory

A. The solution must provide a reference set directory that is made available to all users, including those without accounts.
B. The solution must provide features to search and view the entire catalog of available sets, based on well known search techniques and user interfaces, including across different languages where applicable.
C. The solution must also provide a way for sets that are not hosted on this service to be included in the directory of sets.
D. All reference set metadata is searchable.
E. Searching by reference set members must be available based upon term searches across all active terms of the concepts in sets (including different languages where

applicable). Results for a term search should list all the reference sets where the relevant concept or term appears.

F. Only reference sets that the user has access to should be displayed. The results should make clear if a reference set is published or not, and the responsible organization, as well as the date of the previous release and the dependent edition/extension.
G. The tags from the reference set metadata should be used to display categories of reference sets

2.1.2. Displaying reference sets

A. When viewing a reference set, all the members of the reference set should be visible. Historical versions of the reference set should also be visible.
B. For large sets, paging functionality should be available as well as the ability to see the members of a reference set as a hierarchy.
C. The user should be able to choose which terms are displayed in the list of set members. These could be any of the following, with the option to display two or more:
    a. FSN
    b. Preferred term in any/each of the available language reference sets of the edition/extension on which the reference set is dependent (e.g. US, SE for Swedish… US, FR, NL, DE for Belgium, etc). The user should be made aware of the available languages & dialects that could be displayed
D. When viewing a reference set within a user's own project, the audit trail of the changes made to a set must be easily accessible
E. Users should be able to search the members of a set using term searches as well as ECL within the context of the relevant set. These term searches should match any active term on concepts in the set (not just those displayed as above),.

2.1.3. Set types

The platform should support the creation and maintenance of the following types of sets of SNOMED CT content:

A. SNOMED CT extensional (simple type reference set) and intensional (simple query specification) reference sets
B. SNOMED CT sets of data represented as free sets, including the Global Patient Set (GPS)
C. SNOMED CT sets of data represented as HL7 FHIR value sets
D. Local sets, which are reference sets that may not have a SNOMED CT identifier and will not be published as part of an official SNOMED CT extension or edition release. These may be used by users who wish to create sets of SNOMED CT content without the overhead of creating an extension and managing releases.

2.1.4. Extensional reference sets

A. Creation of a new set must start through any of the following ways:

    a. Selecting an existing SNOMED CT concept which represents the reference set or creating a SNOMED CT concept for the reference set, asking the user for the parent concept, which is stored on the terminology server. The latter option may only be available if the user has permissions to create new SNOMED CT concepts.
    b. When not creating a formal SNOMED CT reference set, the user should be able to provide a local reference identifier to create a set which can be used as a subset or value set.

B. Users must be able to import an existing set of members from an RF2 (simple reference set) file or from a given list of SNOMED CT concept identifiers.
C. Users must be able to create an initial set of members from an ECL query, with inclusions or exclusions, using search features currently provided by the Snowstorm SNOMED CT terminology server. The solution will re-use the SNOMED International ECL query builder to aid end-users without detailed ECL knowledge.
D. Users must be able to manually add or remove refset members by entering or selecting known codes, text-based searches, or hierarchical searches.
E. Users must be able to recompute a previously entered ECL definition against changes to the underlying SNOMED CT edition/extension. During this process, referred to as migration, the solution should identify manually added/edited members that are no longer active.
F. The solution will provide a feature for migrating a reference set from its existing dependent edition version to a newer version. The user will choose the new version of the relevant SNOMED CT edition/extension to migrate to. During this migration process, a report will be produced for the user of the changes. The user will be able see potential new members, potential deprecated members, and be able to accept or reject these changes.
G. The migration of reference sets can be done in bulk in situations where a user/project has multiple reference sets to update at the same time.

2.1.5. Intensional reference sets

A. Users must be able to choose to create an intensional reference set, instead of an extensional reference set.
B. All functionality listed for extensional reference sets in 2.1.4 is relevant aside from the following differences listed here.
C. The solution must support manual override with “include” and “exclude” lists for outlier codes alongside the ECL query.
D. The solution must use the SNOMED International ECL query builder to aid end-users without detailed ECL knowledge.
E. Users must be able to convert an intensional refset to an extensional refset.

2.1.6. Publishing reference sets

A. If a user is part of an organization involved in the SNOMED International Managed Service, the user should be able to flag an official reference set as complete and ready to be released in the relevant extension or edition.
B. Users should be able to self-publish sets that are not necessarily included in a managed service extension to MLDS for those National Release Centers using MLDS. The API for this is documented here - https://confluence.ihtsdotools.org/display/MLDSDOC/MLDS+REST+API+Documentation

2.1.7. Reference set export/download

A. Users must be able to download any of the reference sets either from their own project(s) workspaces or in the public directory from other projects.
B. The format of download must include the following:
    a. SNOMED CT RF2 format
    b. SNOMED CT free set format
C. When downloading an intensional reference set, users should be given the option to choose from downloading the definition as ECL (including all inclusions and exclusions, collated into the ECL definition) and/or download the list of members of the reference set in any of the formats above.
D. The user should be able to select if terms are included in the download and whether these are the FSNs or PTs or both in a particular language reference set.
E. Visitors to the site who are not logged in can download sets which will only contain SNOMED CT component identifiers. The exception to this is if a set is marked as a freeset (such as GPS), then visitors can download the freeset format, once they have provided a minimum set of information (see current GPS request form).

2.1.8. Reference set metadata

A. Ability to store and display reference set metadata, including
    a. version, which would be indicated by a published or effective date (e.g. 2020-07-31)
    b. type (refset, freeset, valueset, local set)
    c. identifier (SNOMED CT id, value set id, or local set id)
    d. name (preferred term in chosen language)
    e. the edition or extension of SNOMED CT that was used to create it, otherwise known as the “dependent edition or extension”,
    f. an indicator of whether it has been published including the date of publication,
    g. an indicator of its workflow status (in development, beta, published),
    h. owning and responsible organization

i. any relevant narrative about the reason for the reference set including links to any external reference sites
j. tags or labels that can be used in the categorisation of reference sets
k. direct URL (preferably as an HL7 FHIR resource) and URI, that directs a user into the solution where the reference set itself can be visualized, browsed, and searched.
l. audit trail, only visible to project members, to see the history of work on all aspects of the reference set

B. Support for attaching related artifacts to a “published” reference set to allow the solution to be used as a distribution portal (if desired). This would include:
    a. documentation
    b. any other ancillary or auxiliary attachments.

2.1.9. Reference set comparison

A. The solution must allow users to compare two given reference sets.
B. Users should be given a comparison screen or report in which members in common, members added, and members removed were all clearly shown.
C. Users should be able to add or remove members between each refset on either side of the comparison.
D. This feature is intended to support these use cases (among others):
    a. Comparison of a published reference set of a particular edition of SNOMED CT against a different edition (e.g. the international version of a reference set vs. the US version).
    b. Comparison of two reference sets developed by different authors (or organisations) that are intended for the same or similar purpose.

2.1.10. Common functionality

A. Users must be able to use any SNOMED CT Edition or Extension that is available on the underlying terminology server, including using SNOMED CT content that is still being developed and has not yet been published.
B. Unless the set is new and unpublished, all SNOMED CT RF2 rules must be observed, i.e. reference set members removed from a published set should be inactivated in the refset, etc.
C. Reference sets which have not been published, I.e. those without an effective date, can be deleted, with suitable warnings to the user that once deleted, they are no longer retrievable.
D. Reference sets which have been published and have an effective date can be inactivated by users and they can choose the relevant historical association. All relevant refset members should also be made inactive within the terminology server directly if not done already.
E. All reference sets must be saved, stored and maintained on the SNOMED International terminology server (Snowstorm).

F. Users should be able to clone/copy existing reference sets from their own projects, or from other publicly visible reference sets into their own projects, creating the relevant SNOMED CT concept where necessary and if the user has permissions to do so (unless cloning as a local set)
G. The SNOMED International release validation framework (RVF) will be used at this time to perform validation checks on all members against the updated SNOMED CT version. Things like members now inactive will be flagged and the author given an opportunity to make relevant changes.
H. The release process for reference sets that are part of an extension or edition will be managed as direct exports from the terminology server as for any release. The Reference Set tool is not expected to be part of this process.

2.2. Projects, Users, Collaboration and Workflow Functionality

2.2.1. Projects & Users

A. All work will be associated with a project which can be set up and administered by users with the appropriate access.
B. Projects can be private or public so that any published work is included in the public directory
C. Projects can be associated with an Organization or group. These organizations or groups can have multiple projects.
D. Users can be added to projects either individually or as a given user-group.
E. If a User does not have an account in the system, project administrators can invite new users, using their email address, providing the relevant information to apply for an account. Using that given email address for that user to be automatically added to the project on account creation.
F. Users with an account on the system can request to be added to a project that is visible in the public directory and would be approved by project administrators
G. Project-based mechanism to allow for grouping the maintenance of multiple reference sets under the same “organization”. This streamlines user-role management and binds together various reference sets that may belong to or be maintained by the same organization.

2.2.2. Collaborative and crowdsourcing features

A key feature of the solution/service will be for users to collaborate amongst themselves, within organizations/groups, across different organizations/groups as well as providing the ability for external users to provide feedback. The following features will support this functionality:

A. Providing feedback
    a. Ability for users to provide feedback on any reference sets in the directory. This feedback should be possible at the level of the reference set or for each of the members of the reference set.
    b. Users can decide to make their feedback visible publicly or only for themselves and the responsible organization/group

    c. Other users will be able to see any previous feedback that has been marked as public at the reference set level or against the components in the reference set.
    d. The owners of the reference set in question will receive notifications that feedback has been received.

B. Receiving feedback
    a. Owners of reference sets can seek feedback from other groups/organizations and users without having to formally publish a reference set. This would involve making the reference set available in a ‘beta’ stage.
    b. Feedback can also be sought on published reference sets although owners must be able to set the level of feedback that can be given (reference set members, reference set or none at all).
    c. The beta reference set can be made visible in the reference set directory or kept private, depending on the requirements of the owning organization/group
    d. ‘Share set’ functionality to make it easy to request the feedback of specific users/organizations by sending a direct URL to the reference set for feedback
    e. Any component of a reference set which has received feedback should have a visual indication that feedback has been left.
    f. When reviewing feedback, there should be a way for users to go directly to the next feedback without needing to leave a review screen/dialog box.
    g. The owners of the reference set can respond to feedback within the solution (much like comments in a document) and/or resolve feedback which removes the feedback from the public view. This feedback should be retrievable after being resolved/removed on request.

2.2.3. Workflow

A. Users should be able to define a simple workflow based on given workflow templates. These templates will reflect the number of reviews that are expected on a set before being published and will range from workflow with no extra review to those with multiple review steps involving different users.
B. Both external and SNOMED International users can have different or multiple roles across workflows (author, reviewer, etc) so that at different stages in the workflow, users only have access to sets depending on their role.
C. When a user has multiple roles, it should be very clear in the application which of the user’s roles is applicable at a given point in the workflow.
D. There should be a full audit trail of all actions within the workflow, including authoring, submitting for review, and reviewing. This audit trail should be available either as a download or a separate screen for use by administrators.

2.3. Out of scope functionality

The following are out of scope of this tool:

A. Map reference sets
B. Translation of SNOMED CT terms

3. Non-functional Requirements

A. The solution must store and retrieve SNOMED CT data and reference sets from the provided SNOMED International Terminology Server. This server will be an instance of Snowstorm, https://github.com/IHTSDO/snowstorm.

B. Performance is expected to not impact the user experience and testing against performance for higher volumes of concurrent users is required.

C. Where necessary, the functionality for reference sets on Snowstorm may need to be enhanced. These enhancements will be in the scope of this project.

D. Authentication is through the current SNOMED International identity management solution, https://github.com/IHTSDO/identity-management-service, but other more standard methods should be considered, including OAuth2, to allow login using other web services (including, but not limited to, Google, GitHub, etc. Facebook is not expected to be included in this list)

E. User interface look and feel should be similar to existing SNOMED International authoring solutions

F. The solution must provide functionality to migrate all user data and sets from the current SNOMED International reference set & translation platform.

G. Whilst the organization has an existing technology stack preference based on Java, Spring Boot, Angular and Linux, responses to the RFP can propose other technologies with rationale as to the choice.

H. Any solution will be expected to follow SNOMED International's infrastructure devops procedures.

I. All software developed under this work ideally will be made available under an open source license, specifically Apache v2.