SNOMED CT Starter Set Translation package Alpha Release Notes - March 2018

Date	**20180331**
Document Version	**0.1**
Release Status	**ALPHA**

SNOMED International

Leading healthcare terminology, worldwide

1 Introduction

IHTSDO passed a translation support policy for Members to assist in increasing translation and adoption of SNOMED CT.

Policy Assumptions:

- IHTSDO will not assume managerial responsibility for any translations other than the existing Spanish translation
- The IP of any derivative work belongs to IHTSDO
- The IP of any translated work belongs to IHTSDO and is to be shared with all Members
- IHTSDO encourages collaboration and will work to make national translation available to all Members
- Minimum core set should be translated for adoption purposes in non-English speaking Member countries

This document deals with the final assumption from the list above, that a minimum core set of terms will be made available, and that this set would be translated into a number of agreed languages to assist new implementers using SNOMED CT.

The ALPHA release package is distributed for evaluation purposes only. It must not be used in production clinical systems or in clinical settings.

2 Background

2.1 Scope

The purpose of the release package is to provide a minimum core set of terms with a number of translations to assist in the implementation of SNOMED CT at a national level. The set is not intended to be a definitive list of required terms, but rather to provide a basic set of terms, which can then be extended in line with Member requirements. The release files contain only conceptId's and descriptions, access to SNOMED CT relationships are only available when the sets are implemented along with a full SNOMED CT release

2.2 Use cases:

The starter set and translations aim to assist end users to access SNOMED CT in languages other than English. This will enable a larger user base to begin to explore SNOMED CT by providing descriptions in four languages provided

3 Motivation

Translation is an essential part of ensuring the continued growth of SNOMED CT worldwide. The starter set and translations aim to support the increasing numbers of end-users who do not have English as their first language. The starter set does not attempt to provide a complete translation of SNOMED CT, but rather provide translations for commonly used terms.

4 Content

4.1 March 2018 Alpha release content

The March 2018 Alpha release contains the following language sets:

- French

- German
- Chinese

For the March 2018 alpha release, the Japanese translation has been removed as it is currently undergoing extensive external quality assurance.

If you have any queries related to the Japanese translation please email info@snomed.org, marking your query "Japanese Translation".

4.2 Minimum Core set (Starter set) creation

The minimum core set (starter set) describes a set of terms identified as commonly used and high usage terms by existing SNOMED CT users. The set is not intended to be a comprehensive list, but rather to provide a defined list of common terms that will support new users of SNOMED CT. The approximate size of the minimum core set is 6,300 concepts.

4.3 Sources of content

To provide the content of the minimum core set, a number of sources have been identified. These are as follows:

NLM top 2500 – 2,500 concepts

Kaiser top 2500 – 2,500 concepts

UK top 1000 terms used in Primary Care – 1,000 concepts

IFP/GP subset – 4,346 concepts

NLM top 2500

Top 2500 problem list – created and maintained by the National Library of medicine (NLM). The list represents the top 2,500 diagnoses codes used in the United States

Kaiser top 2500

Top 2500 concepts used at Kaiser Permanente in the United States – created and maintained by the Kaiser Permanente (KP). The list includes SNOMED CT content created and maintained in the in the Kaiser Permanente SNOMED CT extension.

UK top 1000 terms used in Primary Care

Represent the top 1000 terms used in primary care in the UK. The usage figures represent the use of Read/SNOMED CT codes across multiple UK primary care system suppliers. The content includes mainly content created and maintained in the SNOMED CT UK extension.

IFP/GP (International Family Practice and General Practice) subset

Represents content used in primary care. The content originates from ICPC-2 (International Classification of primary care -2), maintained by the University of Sydney. The number of concepts is just under 4,500, and is reviewed on a release basis by the IFP /GP SIG.

4.4 Process for the initial creation of the starter set

The minimum core set is created through the identification of high use content from the specified data sources. The set uses the specified data sources, and any extension specific content is reconciled as specified in "Section 4.4". The list is then reconciled to remove duplicates, and the final reconciled list represents the minimum data set for publication.

4.5 Reconciliation of extension specific content

The use of the UK and Kaiser as sources of usage data creates an issue in the specification of an international set of terms, as both data sources include data from local SNOMED CT extensions. Whilst terms specified in extensions cannot be used internationally as they do not exist in the International Release, it is important that conceptually the clinical utterance is provided. For example in the UK extension, there is a concept "Activated partial thromboplastin time (observable entity)". In the international release there is no equivalent term. But there are concepts representing findings related to "Activated partial thromboplastin time" which should be included in the minimum core set

The process to identify a clinically equivalent term in the International Release will be undertaken by a clinician, who is an experienced individual with extensive knowledge of SNOMED CT.

The approach to publication of the minimum core set is to publish the set with an agreed set of translations. The approach is aligned with the approach of both the United Nations (UN) and World Health Organization (WHO). The languages included represent a subset of the common global languages, and also target countries that are not yet members of the IHTSDO or users of SNOMED CT.

4.6 Maintenance of the minimum core set

For the minimum core set to be viable and useful to users, it is essential that only current content is included. As the content of the International Release is updated currently on a six monthly basis, there may be content that becomes inactivated in the release, which can no longer be used in live systems. In terms of the minimum core set, any content that has been inactivated should be removed, to avoid confusion for potential users.

4.6.1 Release schedule

As the minimum core set, also includes translations and is based on usage data, it is not practical to update the set in line with the publication of the International Release. Updates should be undertaken on a two yearly basis, whereby new content can be identified from usage data, the content can be added to the minimum core set, and translations can be provided.

4.6.2 Regular maintenance

The regular maintenance of the minimum core set should be undertaken on a six monthly basis, in line with publication of the International Release. At this time, a difference report should be provided between the International Release and the minimum core set, where inactivated content is identified and removed from the set. In addition to the changes to the minimum core set, a change report will be produced documenting all removals and additions to the set since the previous formal released version.

4.6.3 Agreement of new terms for inclusion

On a two yearly basis, the minimum core set will be formally updated, and a new version published. At this time new content may be added to the set. New content will be identified from usage data, and reviewed through the internal governance process documented in "Section 5".

4.6.4 Agreement of language set

The criteria for agreeing the languages to translated are as follows:

- Language use globally
- Interest in SNOMED CT
- IT infrastructure maturity in countries language is used
- Fiscal stability of the country – ability to pay
- A business case for SNOMED CT expansion

The proposed languages to be used initially are:

- German – Strong interest - used in multiple countries
- French – Strong interest – used in multiple countries
- Japanese – interest – volume
- Chinese – interest – volume

Future development of additional languages to be included should be based on the above criteria, and agreed through the governance process detailed further on in this document.

4.7 Process for translation

The IHTSDO does not have access to translation expertise to allow high quality and timely delivery of translations into the required languages. The delivery of the required translations is therefore best delivered through a third party. There are a large number of providers of translations in the market; it is therefore suggested that we adopt an approach of publishing a request for proposals (RFP) to allow market forces to drive costs down. The process for publication and agreement of contract should take no longer than 3 weeks.

The contract for translation should include translation into the four identified languages, translation of FSN (fully specified name) and PT (preferred term) within the minimum core set and translation of the SNOMED CT Starter Guide. It should be noted that translations of synonyms are not recommended, as these tend to be driven by local user requirements and will not necessarily be of use.

We acknowledge that there are language nuances relating to clinical terms that are sometimes difficult to identify. It is therefore recommended that where possible we attempt to identify a clinician fluent in the required language to review the translated product as a final quality assurance step.

4.8 Maintenance of the minimum core set

For the minimum core set to be viable and useful to users, it is essential that only current content is included. As the content of the International Release is updated currently on a six monthly basis, there may be content that becomes inactivated in the release, which can no longer be used in live systems. In terms of the minimum core set, any content that has been inactivated should be removed, to avoid confusion for potential users.

4.8.1 Release schedule

As the minimum core set, also includes translations and is based on usage data, it is not practical to update the set in line with the publication of the International Release. Updates should be undertaken on a two yearly basis, whereby new content can be identified from usage data, the content can be added to the minimum core set, and translations can be provided.

4.8.2 Regular maintenance

The regular maintenance of the minimum core set should be undertaken on a six monthly basis, in line with publication of the International Release. At this time, a difference report should be provided between the International Release and the minimum core set, where inactivated content is identified and removed from the set. In addition to the changes to the minimum core set, a change report will be produced documenting all removals and additions to the set since the previous formal released version.

4.8.3 Agreement of new terms for inclusion

On a two yearly basis, the minimum core set will be formally updated, and a new version published. At this time new content may be added to the set. New content will be identified from usage data, and reviewed through the internal governance process documented in "Section 5".

4.8.4 Updating of translations

When the updated version of the minimum core set has been agreed, the new terms for addition to the set will be identified from the change report. The list of new additions will then be identified within a statement of work, and the translation work outsourced to an external contractor, either through the publication of an RFP or through a formal contract with an existing supplier.

5 Design

5.1 Version

The dependent version of SNOMED CT is the January 2018 International Edition release.

The starter set and translations are provided as an alpha release. The translations have undergone some quality assurance, but we would welcome feedback from end-users on the translations specifically. The starter set has been created as specified in "4 Content", and will be updated on a two yearly basis, based on the availability of usage figures from the sources specified.

5.2 Ownership

The starter set and translations are owned and updated by SNOMED International

5.3 Frequency

The Production version of the SNOMED CT® Starter Set Translation package is currently planned for publication after the July 2018 International Edition release of SNOMED CT, based on the Alpha and Beta feedback content changes. Subsequently it will be maintained in line with each release of SNOMED CT International Edition.

6 Implementation

The starter set and translations can be used to provide a view of the concepts and descriptions. But we would recommend that the starter set and translations be implemented with the complete SNOMED International release, thereby providing access to the relationships.

6.1 Implementation guidance

The release package includes the following:

- Starter set (Refset files)
- French language (Description and language files)
- German language (Description and language files)
- Chinese language (Description and language files)
- Japanese language (Description and language files)

We suggest that the implementation of the files should be in the following order:

1. Implementation of the dependent International Release
2. Implementation of the starter set reset files
3. Implementation of the required Description and Language files

By following the order above, there should be no untoward occurrences linked to the implementation of the files

Please note that the number of records in the language and description files differ between languages. This is due to the addition of synonyms specific to the individual languages, and therefore the number added will differ between languages sets.

Please note that in the starter set there are no language-en and description-en files, as all the required information will be found in the dependent International Release

7 Obtaining the Release

The SNOMED CT® Starter Set Translation package is available from the IHTSDO and is maintained and updated in line with the SNOMED CT International release which is currently 6 monthly.

Access within IHTSDO member countries is provided by the Member National Release Centre in each country, via the relevant Member page. Affiliates of IHTSDO in non-member countries can access the refset through their Member Licensing and Distribution SERVICE (MLDS) account. Please contact info@ihtsdo.org for more information if required.

8 Feedback

As stated in "5 Design", the starter set and translations are provided as an alpha release, and SNOMED International is actively seeking feedback on the translations specifically. Feedback can be provided through the SNOMED International Reset/Translation tool https://refset. ihtsdotools.org/#/refset, or through SNOMED International Helpdesk info@snomed.org, by marking the correspondence "Starter set translations".

9 Technical Notes

RF2 package format

The RF2 package convention dictates that it contains all relevant files, regardless of whether or not there is content to be included in each particular release. Therefore, the package contains a mixture of files which contain both header rows and content data, and also files that are intentionally left blank (including only a header record). The reason that these files are not removed from the package is to draw a clear distinction between:

1. ...files that have been deprecated (and therefore removed from the package completely), due to the content no longer being relevant to RF2 in this or future releases, and
2. ...files that just happen to contain no data in this particular release (and are therefore included in the package but left blank, with only a header record), but are still relevant to RF2, and could therefore potentially contain data in future releases.

This allows users to easily distinguish between files that have purposefully been removed or not, as otherwise if files in option 2 above were left out of the package it could be interpreted as an error, rather than an intentional lack of content in that release.

Approvals

Final Version	Date	Approver	Comments
0.1		Ian Green	
0.2	26/03/18	Ian Green	
1.0			

Draft Amendment History

Version	Date	Editor	Comments
0.1	06/03/2018	Andrew Atkinson	Initial draft
0.2	26/03/2018	Ian Green	Revised draft
0.3			
0.4			

Download .pdf here: