WHO ICD Revision Information Note

No: 17
DATE: 30 January 2013
VERSION: FINAL
TO: RSG-SEG; RSG; TAGs; WGs;
SUBJECT: Multilingual Generation

Keywords: Multilingual, Translation, Translation Tool

ISSUE: Multilingual Generation

A goal of ICD revision is to be able to represent ICD content in multiple languages while maintaining conceptual integrity. At the same time, the revision process requires the ability to respond efficiently and dynamically to changes in the classification across all of the language versions as the classification evolves. This ability is gained by means of suitable multilingual generation tools.

Definition:
This note describes the way ICD-11 will be represented in multiple languages and the tooling environment that will be used to support this.

Decisions - Options:

The multilingual representations of ICD-11 will be based on the Foundation Component. The semantic relationships that exist in the Foundation Component will be preserved during the multilingual generation process while allowing for flexible representation of the concepts in each individual language.

The multilingual generation of ICD-11 will be facilitated by the creation of a web-based software platform.

Features of this platform include that it will:

- Make use of existing translations
  - Using a "translation memory" populated with ICD-10 translations, the tool will include a start-up list with many proposed multilingual representations already in place. The system is capable of using these historical translations even though concepts have been moved around within the classification during the revision efforts already completed.
- Support multilingual representations of titles, definitions, synonyms and narrower terms.
  o Inclusions and index terms are also included in the platform, as subsets of the above.
  o Other content model parameters may be added in the future.

- Provide a flexible multilingual generation workflow
  o Multiple user roles will allow a flexible workflow for the multilingual generation. The ability to indicate the precise status of each item as work progresses will support this, as well. User rights enforced by the platform include:
    ▪ There are three types of users (public, general and reference multilingual content generators). Not all of the types of users may be used in each situation. For example, Arabic multilingual generation may include only general and reference multilingual content generators, whereas French multilingual generation may include only Reference multilingual content generators.
    ▪ Public multilingual content generators will not have the authority to change the content made by general or reference multilingual content generators or from the ICD-10 start-up list.
    ▪ General multilingual content generators will not have the authority to change translations made by reference multilingual content generators or the ICD-10 start-up list.
    ▪ A multilingual representation must be endorsed by at least one reference multilingual content generator (depending on the policy to be developed) before receiving final approval.
  o This platform and workflow will allow for the engagement of a wider range of contributors without the risk of undermining the quality of the multilingual representation made by experts.

- Allow multiple multilingual content generators to work simultaneously via a web-based platform.

- Be suitable for an evolving classification.
  o The platform will keep track of what content has been entered, as well as changes in the source content.
  o Multilingual content generators will be notified about relevant updated content in the Foundation.

- Base the multilingual content on the Foundation component.
  o No need to develop multilingual linearizations separately.
  o No need to develop a separate multilingual Index.

- Support commenting on the multilingual representations.

- Handle special terminology requirements of multiple multilingual representations
- A multilingual representation may use a different title (preferred term) than the exact translation of the English title.
- A multilingual representation may use a different number, larger or smaller, of synonyms than the English source. The important component is the linkage to the ICD entity.

- Multilingual representations generated by this tool can be automatically integrated into the existing software
  - ICD-11 Browser
  - ICD-11 Print Generation (volume I and index)
  - REST APIs based on URIs

- Will support the user by showing similar multilingual representations in order to guide the user and to achieve a higher consistency in the multilingual representations. An example of this might be allowing the user to view the Danish language representation when working on the Norwegian representation, if so desired.