

**Reykjavik, Iceland
24-30 October, 2004**

Title: Maintenance and Publication Tool for WHO-FIC Classifications

Authors: Michael Schopen
Stefanie Weber
Susanne Bröenhorst
Bedirhan T Üstün
Can Celik

Purpose: for discussion

Recommendations:

Abstract:

This paper describes the framework and basic design of a maintenance and publication tool for the WHO Family of International Classifications. This project was initiated at DIMDI aiming at a re-design of the existing SGML-based system. The new system will be based on XML and make use of a database for data storage and access.

Work has started with the design of an XML Schema that describes mono-hierarchical classification systems and uses the Classification Markup Language ClaML as a backbone.

The purpose of the paper is to present this project and to get the experts from the WHO Collaborating Centres involved in the further conceptual work.

Maintenance and Publishing Tool for WHO-FIC Classifications

1. Current situation at DIMDI

Since 1993 a system for electronic publishing of the ICD-10 has been implemented at the German Institute of Medical Documentation and Information. It is based on a semantic and structural representation of the ICD using the Standard Generalized Markup Language (1, 2).

Meanwhile, the Extended Markup Language XML is becoming more and more important as an international standard and the support of SGML by commercial software houses is rapidly decreasing. As a consequence the DIMDI system has to be re-implemented. As the markup language approach has proven itself to be extremely flexible and powerful, it has been decided that the new system will be based on XML. Furthermore, it was decided that for data storage the computer file system is not powerful enough and a database will be used to store the XML documents.

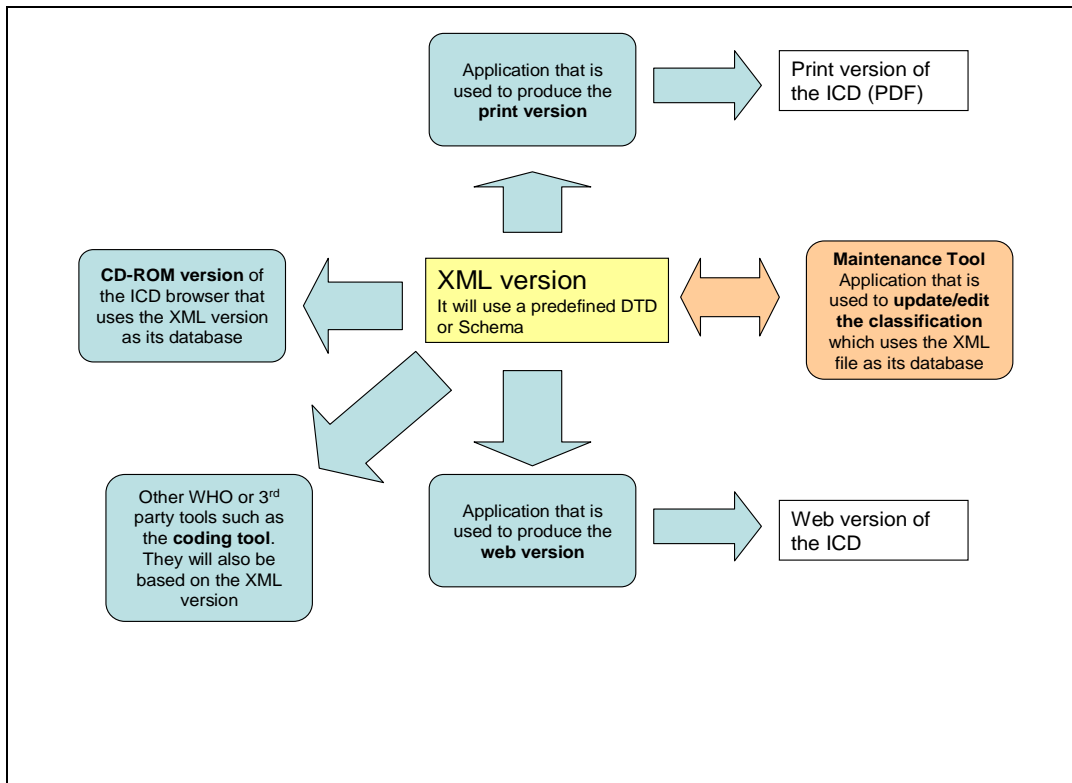
As a first step a generic document structure for mono-hierarchical classification systems has to be designed based on several international experiences (3, 4, 5, 6). As a second step a software system for maintenance and publication has to be developed and implemented.

This paper gives an overview of the system and reflects the current state of the project.

2. Overview of the System

The software system aims to simplify the maintenance and publication of the *WHO Family of International Classifications (WHO-FIC)*. In order to achieve this goal, the classifications have to be transformed into a uniform format so that the same software system can be used for all classifications and their versions. After the triumphant advance of the Extended Markup Language (XML) it was felt that this format is very much suitable for the representation of the WHO-FIC classifications. This approach will also allow providing updated versions of the classifications to 3rd party software producers in a uniform format and in a timely fashion.

The core of the system is the XML document. This document will contain the classification in a predefined format which will be provided as a Document Type Definition (DTD) or as an XML Schema. The document will be stored in a database for easy access. The software system will consist of a tool that can browse and edit the classification and some other tools that can read the XML document and produce the web, print or CD-ROM versions. The structure of the XML document will be generic enough to hold any of the classifications in the WHO-FIC. The format and the software tools are described in more detail in the next sections.



3. Structure for the XML file

There are several options for the generic structure to use for the representation of the WHO-FIC.

OPTION A

We decide which structure to use.

OPTION A-1: We simply use another existing structure that may work for our case, e.g. the Classification Markup Language ClaML or the structure used for the ICD-10-CA (Canadian Adaptation)

OPTION A-2: We decide to design a new structure from scratch

OPTION A-3: We design an expanded version of the existing structures until it meets our needs.

OPTION B

We may involve a company in the process of selecting the structure to use or even ask them to develop a suitable one for the needs of our system.

4. Data Conversion

The existing classifications need to be converted into the format mentioned in section 3. WHO will provide electronic versions of their documents. The classifications that need to be converted are as follows:

- ICD-10 (English and French with highest priority, other WHO official languages with secondary priority, further languages with third priority)
- ICF (same priorities)

It is desirable that the conversion process is supported by those centres being in charge of the language versions at present. The final goal, however, is that the system can be used directly in all Collaborating Centres.

5. Maintenance Tool

The maintenance tool is the software that will be used to update/modify the XML document of the classification. The following are the specifications for this tool:

1. It should be able to read the classification in XML format from the database as defined in section 3.
2. After modifications are performed it should be able to save the document in the same format in the database.
3. Browsing the classification
 - 3.1. The tool should be able to browse the classification hierarchy in a tree format.
 - 3.2. It should also have simple and advanced search capabilities.
4. The user should be able to enter new items, modify existing items, add or delete index items.
5. The tool should be multi-lingual.
 - 5.1. The user interface needs to be in English only, but the tool should be able to handle all of the major languages in the world, although capabilities for other languages might be added little by little.
 - 5.2. All of the functionality of the tool described in this section should be available in multiple languages.
 - 5.3. The tool should be able to display two languages at the same time.
 - 5.4. The tool should have a feature to compare two language versions. For example it should be able to list the items that are not included in one language version but are included in the other language version.
 - 5.5. The tool should be able to handle two versions of one classification so that clinical modifications can be maintained in one system.
6. The tool should have a versioning system.
 - 6.1. The tool should store a timestamp of the last edit that has been made to the classification.
 - 6.2. It should allow rolling back to previous versions.
 - 6.3. It should also keep the information on the updates of each and every item in the classification.

6. Web version builder

This tool should be able to read the XML file and produce a web version of the classification that can be browsed or searched using a web browser.

Features of the web version builder:

1. Production of an updated web version whenever an updated version of the XML document is provided

Features of the web version itself:

1. Basic and advanced search functions
2. Support of multiple languages or at least of the official WHO languages.
3. Hypertext links for the cross-references within the classification.

7. Print version builder

A tool which converts the XML data into a printable format (e.g. RTF or PDF) is necessary. Its product will be a paper version as close as possible to the existing versions.

8. CD-ROM version

A CD-ROM version of the classification that is based on the XML file is needed. Its features should allow:

1. Browsing the classification hierarchy.
2. Support of multiple languages.
3. Basic and advanced search functions.
4. Hypertext links for the cross-references within the classification.
5. The CD-ROM should be able to update itself from the Internet.
6. Production of an updated version of the CD-ROM whenever an updated version of the classification is developed.

9. Present state of the project at DIMDI

The overall design of this system was drafted on June 22, 2004 during a visit of representatives of the World Health Organization to DIMDI. It is highly desirable to use one system in all Collaborating Centres that works on the same representation for all WHO-FIC classifications in all languages so that parallel work can be avoided and resources can be shared.

DIMDI is in urgent need of redesigning its existing system and thus has decided to use XML as described above and to start the project independent of a joint WHO-FIC decision. However, DIMDI offers to collaborate with other Collaborating Centres to avoid parallel work.

DIMDI is in favour of using international standards whenever possible and has thus decided to use the ISO standard ClaML (Classification Markup Language, 6) as a backbone for the new document structure. As ClaML has certain weaknesses in

representation below the code level, DIMDI has started to draft a more elaborate structure. This draft should be discussed internationally as soon as possible to make sure that the experiences made all over the world feed into the current developments. The developers of ClaML have already agreed to take part in this process. WHO-FIC Collaborating Centres with experience in the electronic maintenance of the classifications or experience in XML are very much invited to participate in the process. This collaboration will ensure that the final product is useful for all Collaborating Centres.

Dr. med. Michael Schopen
Dr. med. Stefanie Weber
Susanne Bröenhorst
WHO Collaborating Centre for the Family of International Classifications
c/o DIMDI
Waisenhausgasse 36-38A
50676 Köln
Germany
Telephone: +49 221 4724 297
Telefax: +49 221 4724 444
Email: schopen@dimdi.de

Dr. T. Bedirhan Üstün
Can Celik
World Health Organization
Classification, Assessment and Terminology
Avenue Appia 20
1211 Geneva 27
Switzerland
Telephone: + 41 22 791 36 09
Telefax: + 41 22 791 31 11

References

1. Schopen M: Electronic Publishing of the German Language Edition of ICD-10. Annual Meeting of the Heads of the WHO Collaborating Centres for the Classification of Diseases 1998 in Paris. WHO/GPE/ICD/C/98.24
2. Schopen M: Electronic Publishing of the Alphabetical Index to ICD-10. Annual Meeting of the Heads of the WHO Collaborating Centres for the Classification of Diseases 1999 in Cardiff. WHO/GPE/ICD/C/99.33
3. Stahl C, Walker SM, Garrett C, Truran D, Roberts R, Schopen M: Electronic Maintenance of Clinical Classifications. Annual Meeting of the Heads of the WHO Collaborating Centres for the Classification of Diseases 2000 in Rio de Janeiro. WHO/GPE/ICD/C/00.33
4. Caulfeild D, Renahan M: Development of ICD-10-CA and CCI in a Database Format. Annual Meeting of the Heads of the WHO Collaborating Centres for the Classification of Diseases 2002 in Brisbane. WHO/GPE/CAS/C/02.36
5. Health Informatics – A syntax to represent the content of medical classification systems (ClaML). http://www.centc251.org/WGII/N-01/WGII-N01-03rev%20_2_.pdf (Last visited September 23, 2004).

6. Ten Napel H: Applicability of Dutch ICD-10 Electronic tool for publication of ICD-10 updates and control of derived and related classifications. Annual Meeting of the WHO Family of International Classifications Network 2003 in Cologne.
WHO/HFS/CAS/C/03.36