Connecting the System to Enhance the Practitioner and Consumer Experience in Healthcare
E. Cummings et al. (Eds.)

doi:10.3233/978-1-61499-890-7-126

Representation of ICD-10-AM/ACHI Using Classification Markup Language (ClaML)

Robert David SMITH[a,1], Jeewani Anupama GINIGE[a] and Christos BOULAMATSIS[a]
[a]*School of Computing, Engineering and Mathematics, Western Sydney University, Australia*

Abstract. This paper outlines the Australian experience in adopting the Classification Markup Language (ClaML) to represent two clinical classification systems: ICD-10-AM (International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification) and ACHI (Australian Classification of Health Interventions). The primary goal of this process is to share classification data efficiently with relevant parties in a consistent format. This paper outlines extensions implemented in ClaML to facilitate the representation of the above classifications, with validation of the resulting ClaML files verified through XML stylesheet transformation (XSLT) to render the data on a browser, in the same format of printed ICD-10-AM and ACHI books.

Keywords. ClaML, XML, Clinical Classification, ICD, ICD-10-AM, ACHI

Introduction

ICD-10-AM (International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification) and ACHI (Australian Classification of Health Interventions) are two of the clinical classification systems used in acute care setting, in public and private hospitals in Australia. These classification systems are used to code acute care episodes covering diagnosis and healthcare interventions. ICD-10-AM is the country specific version of the international standard ICD-10[1] and is used for coding diagnosed diseases. ACHI is a unique Australian classification for coding healthcare interventions, for which currently there is no approved equivalent international classification. The ICD-10-AM and ACHI coded healthcare data, is used for a variety of purposes such as: statistical tracking across all regions, territories and states, funding, insurance claim processing and research.

In Australia, biennial releases of ICD-10-AM and ACHI are disseminated to various classification users as printed hardcopy books by IHPA (Independent Hospital Pricing Authority)[2] or through the software systems used within hospitals for clinical coding such as TurboCoder[3], Codefinder[4] and Codexpert[5]. Currently, the National Centre for Classification in Health (NCCH) at the University of Sydney leads the Australian Consortium for Classification Development (ACCD), in collaboration with Western Sydney University (WSU) in the development of the ICD-10-AM and ACHI, on behalf of IHPA. The current method of sharing the latest classification systems with the vendor organisations, is through CSV files, database scripts or PDF files. These formats are not easy to be parsed by systems, in an automated manner, which uses this clinical

[1] Corresponding Author.

classification data. Due to ongoing evolving nature of the classification systems, classification data is heavily subject to change. Further, the sharing of ICD-10-AM and ACHI classifications as database scripts leads to misinterpretation of tables and data fields that leads to errors in the classification system software. Therefore, based on feedback from the vendors, the clinical classification developers have identified a requirement for providing a standard that would facilitate the proper structuring of classification data, which is machine consumable and human readable.

ClaML (Classification Markup Language)[6], is an XML based mark-up language recommended by WHO-FIC (World Health Organization Family of International Classifications Network) for the maintenance, publishing and sharing of clinical classifications. The ClaML schema was developed by DIMDI (Deutsches Institut für Medizinische Dokumentation und Information)[6] and approved as an ISO standard[7]. By definition, ClaML is suitable for the presentation of a range of WHO classifications such as ICD (International Statistical Classification of Diseases and Related Health Problems), ICF (International Classification of Functioning, Disability and Health), and ICHI (International Classification of Health Interventions) the ACHI equivalent international standard, currently under development. The developers of the ClaML specification have provided an example of the representation of the base ICD-10 classification with ClaML[8].

This paper presents details on how ICD-10-AM and ACHI, the two most commonly used clinical classification systems in Australia, are represented using ClaML. The paper highlights the various extensions suggested to be used in ClaML when representing the ICD-10-AM and ACHI classification systems. By using ClaML for representing ICD-10-AM and ACHI, better opportunities and benefits are expected to be seen from the self-describing nature of XML, and its suitability for complex hierarchically structured data.

1. Methodology

This three-phased project identifies the usage of ClaML to capture the complex data associated with ICD-10-AM and ACHI.

The first phase initiated a focus group involving the clinical classification experts and IT specialists to understand the structures of ICD-10-AM, ACHI and ClaML. Two subgroups within this focus group took the initiative to understand how ICD-10-AM and ACHI classification system structures (e.g. chapters, blocks, code, etc.) can be mapped to the tags of ClaML (e.g. classes, modifiers, rubrics, etc.). This exercise was carried out through workshop mode over three months. As part of the workshop activities, relatively simple chapters were manually entered, into a prototype XML that follows the ClaML standard. The purpose was to simulate the migration of ICD-10-AM and ACHI classification data and evaluate any potential issues that would occur during the transition to ClaML. In this exercise, ClaML 2.0 was used as there is more material available to utilise, in comparison to the newer ClaML 3.0 standard.

Once the nuances of ICD-10-AM and ACHI to ClaML were determined, scripts were written to automatically generate the ClaML version of these two classifications from the databases. Since these classification systems are large, it required additional verification to ensure that the representations were accurate. To do that, XSLT (eXtensible Stylesheet Language) files were created to consume the ClaML data, and format the classification data, similar to the format it is represented in printed books (see figure 1). This format is more familiar for human-users which allows for visual

verification and quality assurance of the generated data. The XSLT contains the instruction on how the ClaML should be parsed, in generating HTML content, representing the full ICD-10-AM and ACHI classifications on a web page. The XSLTs use multiple templates to format each part, based on pre-existing formatting rules.

In the third phase, the clinical classification specialists checked the accuracy of representation of this formatted data, through a web-interface. This process revealed certain missing and inaccurate representation of data elements, which led to revising of the ClaML representations as well as the XSLTs. At the time of this writing, the classification specialists are checking the generated content through the said web-based system. Once the classification specialists are satisfied with the ClaML representation, ClaML files for ICD-10-AM and ACHI will be distributed to vendor organisations to test those against their products. The feedback received from this process also, will be used for the further refinement of ClaML representation of ICD-10-AM and ACHI.

A36	**Diphtheria**
A36.0	Pharyngeal diphtheria Diphtheritic membranous angina Tonsillar diphtheria
A36.1	Nasopharyngeal diphtheria
A36.2	Laryngeal diphtheria Diphtheritic laryngotracheitis
A36.3	Cutaneous diphtheria ***Excludes*** erythrasma (L08.1)
A36.8	Other diphtheria Diphtheritic: • cardiomyopathy† (I43.0*) • conjunctivitis† (H13.1*) • myocarditis† (I41.0*) • polyneuritis† (G63.0*)
A36.9	Diphtheria, unspecified

Figure 1. ClaML data rendered on a web page after applying XSLT.

1.1. Business Drivers for the Use of ClaML

The reasons for the use of ClaML and XML are that there was no current standard in place to share the classification to those that needed access to the classification data, such as the vendors of ACCD, one of the only forms of communicating the complete classification data to the vendors was via a database export that had to be done by IT staff which is not good as the export takes time to prepare and requires on-going IT support for the exports if they need to be re-exported or updated.

1.2. ClaML Export versus Existing Methods

With the ClaML export, after any changes are made to the classification an export will be able to be done by any staff member of ACCD. With the database export, IT support would have to be contacted to re-export the classification by running 23 SQL scripts on the production SQL server to generate the Vendor database this would then be copied to removable media and given to the requesting staff. Also to note, ClaML is not used for modification of the classification which is done within an intranet application on the ACCD web site and publication for human use is provided by books printed by IHPA and software applications created by the vendors of ACCD. As a result of our efforts we find that there is no unified methodology for modification, distribution and publication for human use of the classification as each needs different levels or types of access to the classification.

2. Structure and Complexity of Clinical Classification Systems

2.1. Structure of ICD-10-AM and ACHI

In ICD-10-AM, there are 22 chapters and each chapter represents diseases or injuries associated with a major body system such as: respiratory, gastrointestinal or peripheral nervous system. These chapters in total have 16,875 exported entities. The chapters (e.g. Chapter 2 - Neoplasms) are organised in a hierarchical manner, where the top most level is represented by the chapter name: Chapter 2 - Neoplasms, followed by sub-chapters. The sub-chapters (e.g. Malignant Neoplasms) are further specialised areas of the chapter. Under the sub-chapters there are sections (e.g. malignant neoplasm of lip, oral cavity and pharynx) which group similar codes under a specified afflicted body part. Categories (e.g. C00 – malignant neoplasm of lip), defined under sections, are 3 character codes which serves as a grouping of codes under a very specific area. Assignable codes exist under categories, and may be in 2 levels: four digit codes (e.g. C00.0) or five digit codes (e.g. C88.00).

In ACHI, there are 20 chapters and each chapter deals with a different body system similar to ICD-10-AM. In ACHI, there are 8,644 exported entities. Each chapter contains: anatomical site axis, procedural type axis, blocks and codes. The anatomical site axis is the area of the body (e.g. Peripheral nervous system) and contains the procedural type axis, which describes the types of procedure that could be performed on that body system. Under the procedural type axis, there exists blocks (e.g. 370 - Examination procedures on nose) that further divides the procedural type axis into the various sub-types of the procedure. Furthermore, blocks contain codes (e.g. 41764-00 - Nasendoscopy), at the lowest level that represents a procedure performed on a patient.

Both ICD-10-AM and ACHI have two sets of printed books, one representing the *alphabetic index* and the second that is called the *tabular list* that shows the full details of the code listing. When sharing classification data with vendors, it is important to provide data associated with the alphabetical indices and tabular lists, for the accurate re-production of electronic versions of these books.

2.2. Structure of ClaML and Mappings into ICD-10-AM and ACHI

Table 1 illustrates how each ClaML element is used in representation of entities associated with ICD-10-AM and ACHI. Additional information exists in the ICD-10-AM classification that is not in ICD-10 (base version upon which ICD-10-AM is created).

Table 1. Illustrating how ICD-10-AM, ACHI and index elements relate to ClaML elements.

ClaML	ICD-10-AM Tabular	ACHI Tabular	ICD-10-AM/ACHI Index
Class	Chapter , sub-section, code	Chapter , sub-section, code	N/A
Rubric	Attributes	Attributes	Attributes
Label	Attribute text content	Attribute text content	Attribute text content
Meta	Link to other data	Link to other data	N/A
SubClass	Child code links	Child code links	N/A
SuperClass	Parent code link	Parent code link	N/A
Modifier	Attribute, to be moved	N/A	N/A
ModifierClass	Attribute, to be moved	N/A	N/A
Index	N/A	N/A	Index with type set
Term	N/A	N/A	Term text
Reference	N/A	N/A	Term to code (Class) link
TableItem	N/A	N/A	Multiple term to code link.
IndexKinds	N/A	N/A	List of index types
Tables	N/A	N/A	List of table types

2.3. Compatibility of ClaML with Alphabetic Index

The ClaML structure allows for the representation of ICD-10-AM tabular content, with minor changes to the core structure. However the indices of both ICD-10-AM and ACHI could not fit into ClaML as is, due to the use of parent and child links in ClaML as the terms are not unique across the classifications. To provide a structure for the index, ClaML had to be extended (see figure 2).

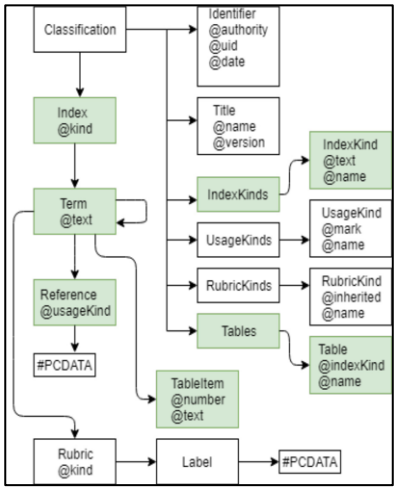

Figure 2. ClaML structure representing index elements. The green elements as extension to ClaML for index representation.

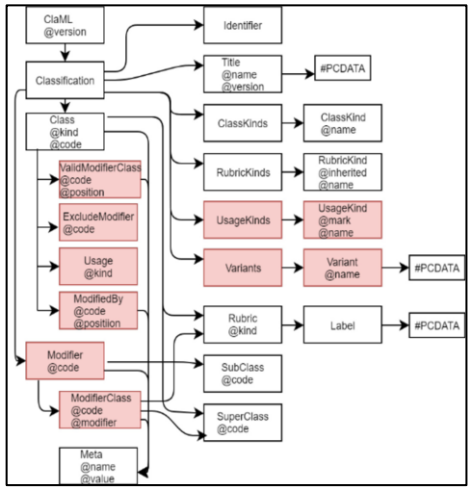

Figure 3. Tabular XML Structure same as ClaML Standard, red not used in ACHI.

2.4. ICD-10-AM/ACHI Attributes vs. ClaML Rubrics

ClaML defines rubrics as a term to describe the extra information associated with a code, such as '*see also'* (which references to another closely associated code) and 'definitions'. The attributes in ICD-10-AM and ACHI predominantly consist of plain text, and sometimes HTML (HyperText Markup Language) formatted content. The term attribute in ICD-10-AM and ACHI is similar to what ICD-10 (the base version) refers to as a rubric. During the process of translating ICD-10-AM to ClaML, the original ICD-10 ClaML rubric terminologies were used. However, the attributes in ICD-10-AM did not perfectly correspond to the ICD-10 ClaML rubric types. ICD-10-AM has striven to differentiate attributes from ICD-10 rubrics, providing more specificity. Thus, less styling tags (resembling HTML) is required in ICD-10-AM compared to ICD-10's representation of ClaML. For instance, in ICD-10-AM, there exists '*Code Also*', '*Code First*', '*Inclusion*', '*See*' and '*Use Additional*' attributes. For ICD-10, there is only '*coding-hint*', which corresponds to all five of these attributes. This poses an issue as there are different styles used to display each of these attributes. For each attribute a mapping to an existing ClaML rubric was created as identified in table 2. The varied attribute styles will cause issues when attempting to use the XML as there would be no means to instruct how to process each attribute.

Table 2. Illustrating how ICD-10-AM and ACHI attributes map to existing ClaML rubrics.

ClaML	ICD-10-AM Attribute
coding-hint	code also, code first, inclusional, see, use additional
text	text
footnote	-- not used --
definition	definition
introduction	introduction
modifierlink	see
note	note, styled note
exclusion	excludes
inclusion	includes
preferredLong	long code name
preferred	code name
meta tag type range	range
meta tag type ACS	ACS No
modifiers	code subdivision

ClaML	ACHI Attribute
coding-hint	codeAlso, codeFirst, codeAlsoWhenPref, inclusional, see
definition	definition
exclusion	excludes
inclusion	Includes
note	note, styledNote
meta tag type ACS	acsNo
preferredLong	long code name
preferred	code name

3. Conclusion

After several workshops and investigations into ClaML, it was found that it is highly compatible to represent ICD-10-AM and ACHI data. This paper explored the use of the ClaML ISO standard for the representation of ICD-10-AM and ACHI. However, this exploratory work indicated that there is no one-to-one map between ClaML structures and ICD-10-AM/ACHI. Several parts of ICD-10-AM classification: codes, code attributes and term attributes were successfully mapped to the ClaML structure elements such as Classes, Rubrics, and Modifiers. Conversely, the index representations of ICD-10-AM and ACHI were not compatible with ClaML. For these outlined reasons, extensions were required to facilitate the transition of ICD-10-AM and ACHI to ClaML standard. In addition, a set of transformation files (XSLT) based on a set of classification data rendering rules were developed, and when combined with the ClaML files, allowed the generation of content to be published on a web browser in a readable manner.

References

[1] *Icd-10 online versions*. 2018. [accessed 2018 13-03-2018]. http://www.who.int/classifications/icd/icdonlineversions/en/.

[2] *Ar-drg classification system product sales*. 2014. [accessed 2018 02-03-2018]. http://ar-drg.laneprint.com.au/.

[3] *Turbocoder*. [accessed 2018 02-03-2018]. http://www.turbocoder.com.au/.

[4] *Health information systems*. 2018. [accessed 2018 02-03-2018]. https://www.3m.com.au/3M/en_AU/his-au/applications/coding-and-reimbursement/.

[5] *Codexpert™ - electronic coding reference application*. 2017. [accessed 2018 02-03-2018]. https://pavilion-health.com/codexpert/.

[6] *Das dimdi - medizinwissen online*. 2018. [accessed 2018 02-03-2018]. https://www.dimdi.de/static/de/index.html.

[7] *Health informatics -- syntax to represent the content of healthcare classification systems -- classification markup language (claml)*. 2013. [accessed 2018 02-03-2018].

[8] *Downloadcenter klassifikationen, terminologien und standards*. 2015. [accessed 2018 02-03-2018]. https://www.dimdi.de/dynamic/de/klassi/downloadcenter/icd-10-gm/vorgaenger/govdata/.