

# Applying Topic Maps to the Classification of Health Interventions

Derek Millar <info@newbook.com>

## Abstract

This paper describes how a health information application in Canada, the classification of health-related interventions, can benefit from the application of Topic Maps [ISO 13250], an international standard for encoding subjects and the relationships between them.

## 1. Introduction

The Canadian Classification of Health Interventions (CCI) is a multi-axial classification of health-related interventions developed and maintained by the Canadian Institute for Health Information (CIHI). CIHI has a mandate to foster better data and technical standards for gathering information and on obtaining consensus on the indicators and determinants of good health in Canada. In response to that mandate, CIHI publishes CCI in conjunction with the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Canada (ICD-10-CA). These publications [CCICD10CA], are used in hospitals and health care centres across Canada as a reference to facilitate the implementation of new diagnosis and intervention standards in Canada.

CCI is a tool that is used to classify the types of interventions applied in the course of treating patients. This information is collected and then analyzed to provide insight into the population's level of health, as well as the state of the country's health system. The tabular listing of interventions in CCI contains over 16,000 codes that provides comprehensive coverage of diagnostic, therapeutic and other associated healthcare interventions.

In order for CCI to be a useful and successful tool it must,

- provide a simple and fast method for locating the code for a particular

intervention,

- help the user to confidently select the correct code for a particular intervention,
- be compatible with other classification schemes for health interventions from other countries and organizations.

This paper describes how Topic Maps, an international standard for codifying subjects and the relationships between them, can help the Canadian Classification of Health Interventions meet these requirements by

- explicitly capturing the semantics of the classification,
- supporting additional methods for navigating the classification,
- providing a mechanism for collaboration and sharing of information.

## **2. Background - Health Information in Canada**

### **2.1. The Canadian Institute for Health Information (CIHI)**

CIHI is an independent, national, not-for-profit organization in Canada that plays a critical role in the development of Canada's health information system. CIHI is responsible for providing accurate and timely information that is needed to establish sound health policies, manage the Canadian health system effectively and create public awareness of factors affecting good health. It does this by collecting, processing and maintaining data for a comprehensive and growing number of health databases and by setting national standards in Canada for financial, statistical and clinical data.

### **2.2. Health Information**

There are two broad types of health information, mortality and morbidity. Mortality relates to causes of death. Morbidity relates to the diagnosis of diseases. Once a disease is diagnosed, it must be treated appropriately.

Information about mortality and morbidity, including diagnosis and intervention information, is collected by Canada for use nationally and provincially to monitor Canada's health and health system. Since Canada is a member state of the World Health Organization ([WHO](#)), the information that is collected and the statistics derived from that information is also shared with the international health community.

Health information is collected from a wide variety of sources, including large, urban hospitals, small clinics, teaching environments and research laboratories. In order to aid in the collection of data from such a wide variety of sources, and to ensure that meaningful statistics can be derived from the collected data, the information must be standardized in some way. The [WHO](#) has developed such a standard for cause of death and the diagnosis of diseases, the International Classification of Diseases ([ICD](#)).

In addition to cause of death and disease diagnosis, information on the procedures used to treat diseases are also collected. These medical procedures are called "interventions". There is currently no widely used international classification for interventions. As a result, Canada has undertaken to develop its own procedure classification that is sensitive to Canada's needs and can be maintained and updated according to medical and technological advances. The Canadian Institute for Health Information has developed this classification of health interventions as part of its mandate to develop and maintain national health information standards.

### **2.3. Definition of "Health Intervention"**

A "health intervention" is defined in [CCI](#) as

“

"A service performed for or on behalf of a client whose purpose is to improve health, to alter or diagnose the course of a disease (health condition), or to promote wellness."

”

### **2.4. Using [CCI](#)**

In a health institution, such as a clinic or hospital, as part of the duties of the health

care providers, information about the diseases, diagnoses and treatments must be collected and recorded. For treatments, or interventions, CCI is the standard reference. When an intervention is performed as part of a patient's treatment, this fact is recorded in the patient's records and is also entered into what is called "abstracting" software that is used to populate a database. The database is then used to compile statistics about occurrences of diseases, their diagnoses and treatment.

In order to make this process efficient, the classification must be:

1. easy to navigate, so that a given intervention can be located quickly,
2. easy to understand, so that the correct code is used for a given intervention.

For example, in the current CCI electronic product, there are a number of tables ("rubric tables") that effectively act as cross-reference tables and indexes. These mechanisms aid in navigating the classification and locating specific interventions. Conceivably, there are many different types of indexes that can be constructed around the CCI.

## **2.5. The CCI Classification Scheme**

Figure 1 illustrates how health interventions are classified in CCI.

<b>Section 3 — Diagnostic Imaging Interventions</b>	
...	
<b>3.AN.^^.^^</b>	<b>Diagnostic Imaging Interventions on the Brain</b>
...	
<b>3.AN.40.^^</b>	<b>Magnetic Resonance Imaging, Brain</b>
Includes:	That for meninges, ventricles, cerebellum, brain stem, cisterna [of brain], posterior fossa MRI, brain
3.AN.40.VA	without contrast
3.AN.40.VC	following intravenous injection of contrast
3.AN.40.VZ	following percutaneous injection of contrast

**Figure 1. Example of CCI Classification Scheme**

In **CCI**, interventions are classified by a code that is composed of as many as six discrete code "components". In the example above, only four components are illustrated.

The first component, "3", indicates one of seven broad "realms" of intervention. In this example, the realm is "Diagnostic Imaging".

The second component, "AN", indicates a group or region of focus. This component allows generic intervention types to be grouped into some reasonable arrangement. In this example, the region is "the brain". In some of the seven realms of interventions, these groups correspond to anatomical sites, or areas of the body. In other sections, the groups reflect some reasonable breakdown of the realm.

The third component, "40", indicates the generic type of intervention, in this case, "magnetic resonance imaging". The fact that this code refers to MRI of the brain is explicit in the title, as well as implicit in the first two components.

The fourth component indicates the specific approach and technique used in the intervention. In some of the sections, as in the example, this fourth code is sufficient to uniquely identify the intervention. In other sections, up to two additional components, to indicate the tools used and the tissue used in the intervention, are required to uniquely identify an intervention. For the purposes of this paper, only the first four codes are addressed.

To summarize and simplify somewhat: [CCI](#) is a comprehensive list of health interventions, each given a unique, identifying code. The code has been constructed so that the components that comprise a given code imply certain aspects and information about each intervention with the intention that this makes the codes easier to locate and infer for a given intervention, and vice versa.

It is possible to perform analysis on health interventions looking at just one or any combination of the "components". This makes the classification into a very flexible and powerful tool for searching and organizing information about healthcare interventions. The multi-dimensional aspect of the classification of interventions and the many possible methods of locating an intervention make this a compelling application for topic maps.

### **3. Overview of Topic Maps**

Topic Maps [[ISO 13250](#)] is an international standard that defines a way of encoding information subjects and the relationships that exist between and among them. It provides a mechanism for organizing unstructured information. A topic map is most easily conceptualized as the digital equivalent of an index that might be found at the back of a printed book. However, a topic map supports many more organizing strategies than does a simple index.

A topic map provides a mechanism for establishing a network of links layered over a collection of information. These links can be used to navigate the information in many different ways. The network can be extended to grow with the collection, or can be merged with other topic maps to provide additional paths through the information. A single collection of information may have many topic maps, as a topic map just provides a particular view or path through the information. Finally, a topic map explicitly captures additional, tacit information. It is this point that has captured the

interest of those working on knowledge management issues, because a topic map can be seen as a mechanism that helps capture what could be considered "knowledge" from a set of information.

### 3.1. Topic Map Building Blocks

The primary unit of a topic map is a "topic". A topic can be literally anything one might think of. Generally, any noun within a collection of information, or any noun that might relate to a collection of information, can be included in a topic map as a topic. In the example above, "*diagnostic imaging*", "*the brain*", "*magnetic resonance imaging*", and "*magnetic resonance imaging of the brain following intravenous injection of contrast*" are all examples of topics.

Topics can have types. A topic type is a category to which sets of topics belong. This is where things can get confusing, because a topic type is also a topic, just a special kind of topic. Continuing with our example, we can identify additional topics such as "*realm*", "*group or region of focus*", "*generic intervention*" and "*intervention*" and earmark these topics as topic types. Then

- "diagnostic imaging" is a topic of type "realm",
- "brain" is a topic of type "group or region of focus",
- "magnetic resonance imaging" is a topic of type "generic intervention",
- "magnetic resonance imaging of the brain following intravenous injection of contrast", is a topic of type "intervention".

In addition to topics and links to occurrences of topics, a topic map contains associations. An association is something that relates topics together. The advantage of associations lies in the fact that an association is independent of any occurrences of the topics that it relates. So associations in a topic map provide a mechanism of capturing additional information about the information, or "metainformation" that the topic map is concerned with. An association is also a topic, according to ISO 13250.

In our example, a realm of interventions "*encompasses*" groups or regions of focus.

This is an example of an association between two topics that are actually topic types. Also, each topic that is involved in an association has an "association role". An association role explicitly defines the role that each topic plays in an association.

Normally, a topic has links that point to occurrences of that particular topic. An occurrence is simply a portion of a collection of information that is about a topic. For purposes of this paper, CCI is being considered as a "pure" index, so there are no links to occurrences of the topics in CCI. For the moment, only the topics and the relationships between them are of interest.

### **3.2. Topic Map Templates**

One source of difficulty in understanding and building a topic map is that almost everything involved is defined as a "topic". This makes for an elegant and self-contained standard. However, it also makes it difficult to distinguish the "declarative" parts of a topic map, i.e., the part that designates what and how the map classifies the information, and the actual topics and other objects in the map. In order to clarify these two aspects of a topic map, a good approach is to construct a topic map "template". The template contains the topics, association types and other objects that will be the basis for the actual topic map.

In order to create a topic map for the CCI, a good approach would be to define a topic map template for the classification scheme. Once this is done, a topic map can be implemented containing the actual topics and associations.

## **4. Applying Topic Maps to the CCI Classification Scheme**

In order to build a topic map for the CCI, the following approach can be taken:

1. Identify the topic types, association types, association roles, required to generate the topic map.
2. Build a topic map template that can be used to implement an "actual" topic map.
3. Create a topic map for CCI that is based upon the template.

The following topic types have been identified in CCI for purposes of this

investigation:

- Realm of Intervention
- Group or Region of Focus
- Generic Intervention
- Intervention
- Part of Anatomy

The following associations have been identified in [CCI](#) for purposes of this investigation:

A realm of intervention

- "*encompasses*" Groups or Regions of Focus

A Group or Region of Focus

- "*is subdivided into*" generic interventions

A generic intervention

- "*is a*" intervention type
- "*includes*" interventions
- "*excludes*" interventions

An intervention

- "*involves*" part of the anatomy
- "*is related to*" another intervention
- "*is included by*" a generic intervention

- *"is excluded by"* a generic intervention
- *"is a"* generic intervention

A part of Anatomy

- *"is involved in"* an intervention
- *"is located in"* a region of focus

Among these associations, some are symmetric, some are the inverse of others, and some imply a parent-child, hierarchical relationship between topics. These are the most obvious topics and associations in CCI that capture the conventional structure of the classification scheme. There are more topics and associations that are implicit that can be identified and added to this collection. Encoding these in a topic map allows the set of information to be navigated based upon these associations.

The creation of a topic map for CCI involves encoding the information according to the Topic Maps DTD. Currently, an XML representation of the CCICD10-CA data set is produced as part of the publishing process that generates the print and electronic products. This XML data can be processed to produce a topic map.

It is important to keep in perspective that only additional indexing and cross-referencing is going on here. In order to produce a usable and valid topic map, subject matter experts and skilled indexers or cataloguing specialists must be involved in identifying these additional access points into the information. The topic map mechanism allows for capturing these additional "views" into the information set and provides greater flexibility in manipulating and using the information.

One possible path for the evolution of the current database publishing system that produces the CCICD10-CA would be to use the topic map as the logical data model for these publications. Information products could then be generated from this data model which would incorporate many different navigation paths based upon the explicit encoding of topics and their associations.

## **5. Benefits of Using Topic Maps**

## **5.1. Capture of Information**

The topic map encodes the classification scheme explicitly, capturing the relationships between the information and the grouping of the interventions. The code becomes an identifier for the intervention. This makes the data more easily processable in different contexts and positions the classification scheme for the use of natural language processing and inference engines to apply the classification.

## **5.2. Location of Information**

The explicit encoding of implicit relationships and associations among topics can provide additional navigation paths in an electronic CCICD10-CA CD-ROM product that is based upon or generated from a topic map.

## **5.3. Exchange of Information**

As previously mentioned, [WHO](#) has not developed a standard classification for health procedures and interventions. Consequently, national organizations have developed classifications of their own according to their regional and national requirements.

Having multiple classifications in different countries makes it difficult to:

- compare statistics between different countries,
- combine data from different countries to produce results based on a larger sample of data,
- combine data from different countries to produce results for a larger region,

Topic maps provide a mechanism for uniquely identifying an intervention in an application independent, portable manner, by assigning each intervention (topic) a unique "identity" attribute. An equivalency decision can then be made automatically by an application. This would allow for the comparing and combining of sets of intervention data.

Taking this concept further, other relationships or rules could be defined via a topic map that would not just identify equivalency relationships, but could indicate to an application how intervention data could be combined in more sophisticated methods. This could be used as a "normalization" process for different data sets.

## 6. Summary

The Canadian Classification for Health Interventions (CCI) is a complex, multi-dimensional classification scheme for health interventions. This paper has illustrated how topic maps can be applied to CCI and has discussed some potential benefits of encoding the CCI as a topic map.

These benefits include:

- more explicit capture of information,
- easier navigation of the classification,
- support for exchange of information with other health organizations,
- support for merging of information with other health organizations.

## Acknowledgements

I would like to thank the Canadian Institute for Health Information for its kind permission to use CCI to explore the emerging technology of Topic Maps and to present this talk.

## Bibliography

[CCICD10CA] *CCICD10 Canada, Version 2001*, The Canadian Institute for Health Information, Toronto, 2001

[ISO 13250] *International Organization for Standardization: ISO/IEC 13250:2000 Document description and processing languages - Topic Maps*, ISO, Geneva, 2000

## Glossary

CCI	Canadian Classification of Health Interventions
CIHI	Canadian Institute for Health Information
ICD	International Classification of Diseases
ICD-10-CA	International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Canada
WHO	World Health Organization

## Biography

### Derek Millar

Director, Professional Services  
NewBook Production Inc.  
Mississauga  
Canada  
Email: [info@newbook.com](mailto:info@newbook.com)

*Derek Millar* - Derek Millar is Director of Professional Services at NewBook Production Inc. He has many years of experience providing consulting services and technical expertise to companies implementing solutions for publishing and document management. He has helped develop solutions for issues relating to data creation and maintenance in the aerospace, legal publishing and reference publishing industries.

NewBook Production Inc is a management and technical consulting company located in Ontario, Canada specializing in the field of electronic publishing systems involving SGML and XML. The company has delivered many solutions where the manipulation and re-purposing of information is central to the business requirement and is now focusing on the new processes and techniques that can be applied to data once it is marked up and stored in flexible, structured formats.

This includes emerging knowledge technologies such as Topic Maps.