

Tsoi-Ching Yip*, Ashij Kumar, and Michael Minuk
Meteorological Service of Canada, Downsview, Ontario, Canada

1. INTRODUCTION

The Meteorological Service of Canada (MSC) is in the process of modernizing the National Climate and Water Archive (NAS). The goal is to create a paperless archive so that all the operational atmospheric and water quantity monitoring data will be captured and quality controlled in near real-time and will be made available to users in a timely fashion. The current MSC NAS contains long-term operational climate, weather, water, marine, lightning and some air quality monitoring data in both electronic digital format and paper format.

2. METADATA

Metadata will play an important role in the modernization of NAS. Metadata is data about data. It describes the content, quality, condition, and other characteristics of data. With proper metadata, users can locate, understand and use the data properly.

3. APPROACH

The NAS adopts a hierarchical approach to the creation and management of metadata employing the Dublin Core standards and the Content Standards for Digital Geospatial Metadata (CSDGM), which will provide three levels for users to search by. These three levels include Data Discovery, Data Access and Data Exchange.

Discovery Level Metadata - a high level description of Environment Canada's (EC) data and information holdings. This basic level will provide a general description of a data collection using 15 Dublin Core elements. It can also describe single components like reports, web sites, new releases, etc. This is one record which describes the entire dataset, repository or application. These records will be used for high level searching by departmental search engines such as, Environment Canada's Green Lane (<http://www.ec.gc.ca/>).

Access Level Metadata - a mid-level description of departmental data or information holdings that provides detailed access to data holdings (i.e. where to get the data). This description will be at the record level within a data collection system. We adopted about 25 mandatory elements from the CSDGM standard.

Exchange Level Metadata - this level provides detailed information about the data (e.g., type of instrument used in the monitoring). In addition, it uses geospatial and/or biological profiles for full electronic data exchange. We are planning to use as many CSDGM elements as possible for this level.

Using these three levels will allow a user to search for more detailed information as required and not be inundated by numerous returns of information if not needed, which will allow for more efficient and quicker searches. In addition, to ensure consistency in both the cataloguing or indexing and searching of Environment Canada metadata repositories, controlled vocabularies (such as keywords) from standardized thesauri will be used during subject searches.

In addition, NAS contains a detailed operational database on monitoring station information (station location, opening and closing dates, instrumentations and instrumentation changing dates, maintenance schedule, operational schedule, observing program, station owner, etc.). A lot of information contained in the station information database is required by the metadata system. We are in the process of interfacing the operational station information system with the standardized metadata to improve efficiency of development the metadata database.

An off-the-shelf software package is used to interface with the station information database to generate XML (Extensible Markup Language) files. Since XML is a universal, digital format for data exchange, it will allow for ease of searching of these standardized metadata records and information exchange over the Internet.