Development of an Information Portal for the Lake Winnipeg Basin Initiative

Isaac Wong[1], William G. Booty[1], Phil Fong[1], and Sarah Hall[2]

[1] Environment Canada, 867 Lakeshore Road, Burlington, Ontario, L7R 4A6, Canada
[2] Environment Canada, 200 Blvd Sacré-Coeur, Gatineau, Quebec, K1A 0H3, Canada
`{Isaac.Wong,Bill.Booty,Phil.Fong,Sarah.Hall}@ec.gc.ca`

Abstract. The Lake Winnipeg Basin Initiative was created as part of Canada's Action Plan on clean water. Its focus is to deal with excessive lake inputs of nutrients from surface runoff and municipal wastewater. Understanding the dynamics of nutrient loading, the associated algal blooms and resulting changes in fish populations, beach closures, and ecosystem imbalance requires access to various sources of data, information, knowledge, expertise and tools. Such critical components are delivered through the Lake Winnipeg Basin Initiative Information Portal, which integrates multiple geospatial and non-geospatial datasets of information pertaining to the basin and serves as a data, information and modelling portal. With data coming from many disparate sources, the Canadian Geospatial Data Infrastructure standards are applied to ensure interoperability. The "Community of Models" allows the modellers to post their model and results, and also allows the portal users to comment on the results to ensure a healthy dialogue.

Keywords: Lake Winnipeg; environmental information portal; knowledge discovery; data mining; metadata.

1 Introduction

Lake Winnipeg is the sixth largest freshwater lake in Canada. For several years water quality in Lake Winnipeg has been deteriorating. Nutrient loading from agricultural run-off and municipal wastewater effluent had led to advanced eutrophication of the lake; this is the primary issue. These excessive nutrients are contributing to the growth of toxic blue-green algae which, in turn, deplete lake oxygen levels and affect the local commercial fishing industry and recreational activities. Recognizing the severity of Lake Winnipeg's environmental situation, the Canadian government decided to invest $18 million towards cleaning up the lake. Part of the funding for this Lake Winnipeg Basin Initiative (LWBI) is devoted to both monitoring and research, and is aimed at understanding and solving some of the water quality challenges facing Lake Winnipeg. Solutions will be found by identifying and assessing key water quality issues within the lake and its contributing watersheds.

To help ensure decision-makers and water managers have access to relevant scientific information and tools to guide and evaluate water management, a single window information portal to promote data sharing with key federal partners is being

developed. This portal will ensure consistent and reliable delivery of relevant information and will also facilitate data sharing among partners and contributors. In addition, the Portal also provides a mechanism to compile and review information on point and non-point sources of pollution. The result will also lead to better informed decision support and priority setting for compliance promotion and enforcement activities. Additionally, the Portal will enable the identification of gaps and priorities for further collection of information through monitoring and research. Finally, the LWBI Information Portal will serve as a hub to integrate various land and water models being used in the LWBI research and provide their results to the project partners and the public.

2   LWBI Information Portal Considerations

As stated in the introduction, the key objective of the LWBI Information Portal is information sharing and discussion among research personnel in the web environment [1]. Since the Portal must be designed to receive data from various partners whose data formats and methods of data delivery are not the same, multiple considerations must be made at the design level. Information sharing is an important component of the Portal. For proper decision-making, it is essential to have the data and information as accurate as possible. The desire of individuals, groups and organizations working within the watershed to share their information has a direct impact on the success of the Portal. The most common factors of information sharing rely on trusting relationships and good personal contacts [2]. The Portal must present solutions which encourage collaboration and data sharing with an enabling infrastructure that promotes open exchange of data, information and knowledge. This approach will provide a good comfort level to the users and will encourage multi-way partnerships.

The Portal must also be designed to handle information overload. Users will want to pinpoint the appropriate data at the temporal and spatial scale needed to answer their specific question or query. Also, there is the added pressure that users sometimes feel that ownership of their information is lost when it is shared. Users express concerns that information might be misused or misrepresented and in some cases, choose to withhold information, especially raw data, if they haven't yet completed a full interpretation of the information.

To assist decision makers and to support complex environmental issues such as Lake Winnipeg eutrophication, environmental modellers are often asked to investigate and understand the inter-relationship of the ecosystem across multi-media, such as the watershed and lake. No single model can handle such complex conditions alone and it is required to have a Community of Models framework (CoM) to provide a better understanding of the impact to Lake Winnipeg.

Since each model is operated by domain experts in their own field, it is a challenge to integrate various models from various sources and personnel. The models may have to integrate together in a certain configuration and model to model data exchange is an ongoing challenge. The CoM needs to ensure that model results are synchronized seamlessly among all models. This includes guaranteeing consistency among the variables and respective units of measure, temporal and spatial scales. It is important for the CoM that all key personnel, including modellers and data holders, "buy-in" to the approach to avoid any pitfalls and failure.

High quality data minimize errors and reduce uncertainties about the model results. In general, successful calibration and validation of the models requires large amounts of high quality data and a strong background knowledge on the data inputs.

Timeliness of the data is another issue. In many cases modellers frequently update their models with better calibration and new model results will follow. New input data may also become available occasionally. With each new update of new model results and data, it requires all stakeholders within the CoM to know and revise their own results. This trickles down within the CoM until all the results are updated.

3   LWBI Information Portal Architecture

In designing the Lake Winnipeg Basin Initiative Information Portal, significant effort has been invested in the capture and delivery of standards-based metadata, portal functionality, web personalization and a seamlessly integrated modelling framework. Fig. 1 illustrates the components of the LWBI Information Portal. It highlights the use of the CGDI (Canadian Geospatial Data Infrastructure) metadata standard concept, the rich functionality allowing users moving from one component to another while retaining personal preferences by means of web personalization, and finally, a CoM allowing modellers to share information among models.
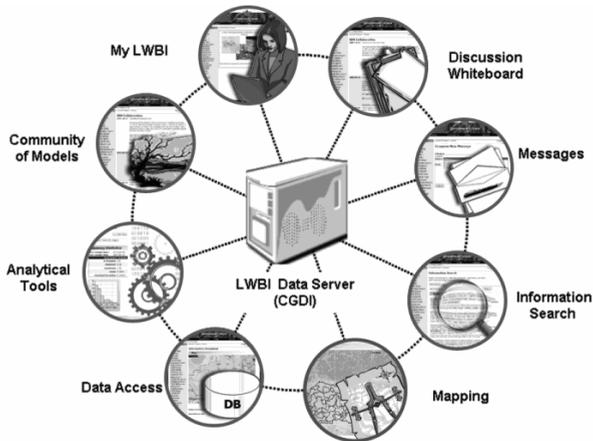

Fig. 1. The LWBI Information Portal components

3.1   Using Standards as the Backbone of the LWBI Information Portal

The Portal endorses data ownership, encourages collaboration among users and provides decision support as well as handles integrated environmental modelling. Compliance with metadata and geospatial standards is important [3]. The backbone of the Portal is based on the CGDI standards. The CGDI standardizes the way information in many of the databases is stored, accessed, and presented online. These standards are endorsed by the Government of Canada and are the same standards often applied to geospatial data infrastructures operated by other countries. Thus, the interoperability of

data from standards-based Canadian and foreign sources is intact. This is an important aspect because it enables the users to understand exactly what data that they are retrieving. The Information Portal is a distributed system that attempts to leave data at source. Sharing data directly from the source, rather than adopting a data warehouse approach, ensures that users receive the most accurate and current data from a single, authoritative source. This removes problems with data duplication and redundancy. Users know that the data complies with the standards and that there will be no surprises, i.e. non-standard data delivery that often requires a huge effort to make it usable.

Depending on the nature of the dataset, there are up to two levels of metadata information. The first level is the overarching collection level metadata. At this level, information such as the description of the dataset, data owner, access privileges, bounding extents and information contacts are among the key mandatory metadata elements. The second level allows definition of station level metadata. This permits the capture of more detailed metadata information from an environmental dataset containing station information.

The LWBI Information Portal is built on the foundation of the CGDI standards. Ideally, data should be stored at source. In practice, however, this can be quite challenging because of implementation issues that may have to be addressed. There may be differences in the levels of information management and technology (IM/IT) expertise from the various partners. While some of the partners are capable of following the CGDI standards and providing data through web services such as Web Map Service and Web Feature Service, this may not be the case for others who have low IM/IT capacity. Although these partners have valuable data to share, they may be constrained by a lack of technical expertise and often require assistance in building up their capacity. In extreme cases, there are partners who are simply unwilling to invest their time and resources in building up that capacity because it is not seen as a critical component of their core business. To secure access to this critical information, the Portal must provide acceptable hosting options for partner's data.

3.2 Portal Functionality

Understanding complex environmental problems and making informed resource management decisions requires the integration of scientific data and knowledge across multiple disciplines and diverse landscapes. Ever increasing demands for timely, accurate and spatially explicit information requires that environmental modellers deploy the latest information technology. Such modern technologies provide decision support capability to the Portal and enable informed decisions surrounding point source pollution, non-point source pollution, and lake eutrophication. The Portal uses the Decision Support System (DSS) approach [4] to assist users in the various decision making processes. For example, the Portal utilizes data and models to tackle the nutrient problem in the LWBI.

There are many design issues for the LWBI Information Portal. These include the problem of information overload, the reluctance of data and information sharing, the wide spectrum of technical capacity of the partners and the consistency of heterogeneous datasets. We will describe some of the functionality currently built into the Portal which provides tools for overcoming these barriers.

3.3 Web Personalization

As more and more data is added to the Information Portal over time, a situation of information overload may result. The expanding knowledge base may lead to usability challenges, perhaps limiting users' ability to efficiently obtain the information that they need. It will be critical to ensure provision of quick access to data for these various users, in order to maintain their long-term interest in the Portal.

To improve the usability and secure user retention of the Portal, web personalization is implemented [5]. The principle components of web personalization provide an environment for categorizing and pre-processing of data and information, linking heterogeneous datasets together logically and providing services that are tailored for the web personalization.

"My LWBI" is an avenue in the Portal to allow the users to save their favourite settings. Users can save the weather information of a city within the Lake Winnipeg Basin, Know Your Watershed information, geospatial boundary of a region of interest, datasets, stations and water quality variables as a "My LWBI" project. Users can save as many projects as they want and can designate any one of their favourite projects as their default project.

The use of metadata searching tools to pinpoint appropriate data and information is an important feature of the Portal, especially when large amounts of information are involved. Since the metadata XML records are based on CGDI standards, it is possible to search information using all the elements of these metadata records. The Portal makes use of a Catalog Service for the Web to hold and query all of the metadata. Users can easily search for relevant information against these collection level metadata records.

Station level metadata of the collection dataset can subsequently be used to select water quality stations. Once the water quality stations are selected, the Portal will dynamically return the list of the available water quality variables for these stations.

The metadata approach does not stop at the collection and station levels. The Portal also provides metadata information about the water quality variables that are chosen. It is capable of identifying the number of sampling points among each station and the duration of sampling. This feature allows users to further refine their information requests. Since the Portal uses the CGDI standards, it is, by default, designed to increase interoperability. Heterogeneous datasets can be combined to provide additional information. In this manner, the Portal not only deals with information overload, but is also capable of information aggregation. User retention is the key to the success of any information portal. In order to promote regular user revisits, the LWBI Information Portal employs a number of strategies. First, the Portal offers a single window to the Lake Winnipeg Basin information. It is critical that the users understand that the Portal is providing up-to-date data and information about the basin. By combining various datasets through the CGDI standards, the Portal offers interoperability among datasets. This allows users to access and retrieve various datasets even when the data are stored at source. Secondly, the components of the Portal are tightly integrated. Users can use "My LWBI" to define the areas of interest and to select appropriate datasets and water quality stations. All of these selections are "memorized" and migrate automatically to other Portal components including analytical tools and models. Users can expect the same responses from all the components of the Portal without

any surprises. Thirdly, the Portal makes use of data selection methods and provides water-related analytical tools to allow the users to understand the data better. These tools include environmental statistics, statistical graphs, tables and thematic mapping.

3.4 Knowledge Discovery Platform

Since knowledge is power, a discussion whiteboard is made available for the users to share information and local knowledge, to ask questions about specific LWBI issues and to expand networking with other like-minded users. Users can discuss specific topics regarding the LWBI. They can also upload relevant information such as spatial layers, photos and spreadsheets to share among their peers. This greatly increases the usability of the Portal since new information will be created dynamically.

The additional supporting tools include a message system allowing users to get help from the Portal system administrators, a "Who's Who" page allowing users to discover the expertise of individuals and organizations that are participating in the Portal, the "What's New" page conveying new information from the Portal administrators to the users and the "Knowledge Exchange" module which allows the users to learn from each others' experiences.
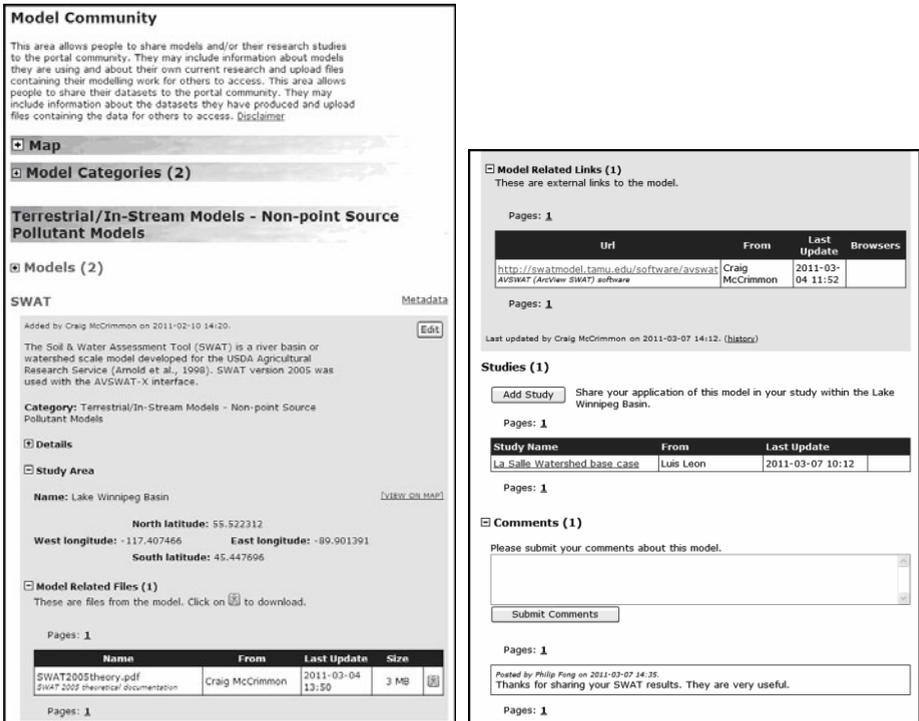

Fig. 2. Community of Models web page

3.5 Community of Models

The use of computer modelling to simulate environmental processes and the interactions between ecological and anthropogenic systems is very common. In the Lake Winnipeg Basin, there are many known models being applied to evaluate, understand and predict the environmental fate of pollutants, their impacts on human health and the environment, and the costs and benefits of alternative policies. However, it is a challenge to set up a modelling framework within the Information Portal that caters to all levels of users [6].

In some modelling, the model may be reasonably accurate to a certain spatial or temporal threshold or both, beyond which point the output becomes thoroughly undependable. These thresholds, where they may exist, tend to be poorly defined. Often, the effectiveness of the communication between the portal user and the modeller is lacking, resulting in the modeller attempting to answer the wrong question or the user asking a question in a way which makes it extremely difficult for the answer to be produced by the model. To alleviate this problem, a common modelling framework of communication for the Lake Winnipeg Basin called "Community of Models" page as shown in Fig. 2 is created. Modellers can deposit their models and model results into the portal in the way that they would like to communicate with the portal users. The assumptions, limitations, inputs and outputs of the models are described and discussed. Users can comment on the model use and directly engage with the modellers.

4   Conclusions

The LWBI Information Portal requires careful consideration to overcome information overload, user buy-in and user retention. In designing the LWBI Information Portal, we employ a number of strategies. To overcome information overload, we use web personalization so that users can customize their preferences with the "My LWBI" option. Session memory is also retained from component to component to allow for better navigation. Search of relevant information and useful environmental data analysis are also important functionalities to combat information overload.

The Information Portal is a distributed system that leaves the data at source but offers a single window access point for the users. The deployment of the CGDI metadata standard concept allows the Portal to address the issue of data interoperability. It is now possible for heterogeneous data to be aggregated. The synergy encourages more buy-in.

The whiteboard forum to exchange information, knowledge and ideas is one of the key retention strategies. It allows like-minded users to have a platform for discussion. In addition, a message system is set up to assist users with technical problems. The response time is kept to the minimum so as not to irritate users. Finally, data and information download is the main goal of the Information Portal. It allows users to search and retrieve data in a friendly data format.

The "Community of Models" framework allows the modellers to communicate their models and results with the Information Portal users. This exchange allows the knowledge about the pollutant transport, eutrophication and algae growth issues in the lake to be shared. The end result of the LWBI Information Portal is to develop

wisdom among users with data mining, information sharing and modelling exchange on how to protect the basin.

References

1. Xiang, X., Madey, G., Huang, Y., Cabaniss, S.: Web Portal and Markup Languages for Collaborative Environmental Research. In: Scharl, A. (ed.) Environmental Online Communication, pp. 113–126. Springer, London (2004)
2. Forslund, J., Larsson, J.: A Portal Implementation for Information Sharing. Master Thesis in Informatics, IT University of Göteborg, Sweden (2005)
3. Kruse, F., Uhrich, S., Klenke, M., Lehmann, H., Giffei, C., Topker, S.: PortalU, a Tool to Support the Implementation of the Shared Environmental Information System (SEIS). In: Proceedings of the European conference TOWARDS eENVIRONMENT, Prague, Czech Republic, March 25-27 (2009)
4. Alter, S.L.: Decision Support Systems: Current Practice and Continuing Challenges. Addison-Wesley, Reading (1980)
5. Eirinaki, M., Vazirgiannis, M.: Web Mining for Web Personalization. ACM Transactions on Internet Technology 3(1), 1–27 (2003)
6. Gordov, E.P., Lykosov, V.N., Fazliev, A.Z.: Web Portal on Environmental Sciences "ATMOS". Adv. Geosci. 8, 33–38 (2006)