

# **Integrated web portal: A case study for Integration of spatial data and domain knowledge**

Rajan Man Bajracharya  
ICIMOD, G.P.O. Box: 3226, Khumaltar, Kathmandu, Nepal  
Tel +977 1 5003222 Fax +977 1 5003277  
rbajracharya@icimod.org

**Abstract:** The HKKH Partnership Project aims at consolidating institutional capacity for systemic planning and management of protected area ecosystem over three national parks in the Hindu Kush-Himalayan region. The project is being implemented at multiple scales involving stakeholders from local, national and regional partners in the region, and domain experts from the region and beyond. The project initiated the concept of **Integrated Web Portal (IWP)** to promote interdisciplinary collaboration, communication and dissemination among many partners.

IWP has been developed with capability to integrate a wide variety of data and information resources such as GIS data, satellite images, interactive maps, model data and documents; all of which are served with associated metadata. This paper attempts to give an overall software architecture of IWP and its components. It highlights the **knowledgebase system** which allows systematic cataloguing of metadata in a standardized format for information resource description. To further integrate the domain knowledge, a novel approach using Topic Maps standard is also presented. Here, two Topic Maps ontologies are designed - one for describing the domain data and the other for describing the spatial data. The merging of the two ontologies provides a semantically rich access point for all domain data and relevant spatial data. Such an approach can be useful for several other environmental domains and can be further integrated with similar applications thus providing interoperability among heterogeneous data sources.

**Keywords:** Portal, Content Management, Geospatial Metadata, Knowledge Management, Topic Maps, Ontology, Glacier and Glacial lakes, Spatial Data Infrastructure.

## **Introduction**

The unprecedented growth of internet based application has provided integrated and innovative solutions to collect, manage and disseminate information and knowledge. The integrated web portal has been designed and developed in the framework of the "HKKH Partnership Project" with an aim to provide a virtual platform to build, share and disseminate information and knowledge resources in a decentralized and distributed fashion as well as to provide a gateway for project's information resources to the wider audience.

Given system specification and requirements of IWP, an approach was taken to develop software components that are independent, where automation and integration between portal components were possible. The project information are organized in different components and categories such as spatial data, image data, bibliography, model etc.

The information managed in IWP fulfilled the project requirements. However, when a user needs domain specific resources, these are often fragmented across different components and categories of IWP. Integrating different domain knowledge with its relevant resources would make this more efficient in searching and retrieving relevant domain resources. The ISO Topic Maps standard (ISO/IEC 13250:2003) approach makes this possible by providing the ability to store, together with data and complex metadata that represents the meaning of the data stored through a one page user interface.

As a case example, SATOPI project's case is taken for demonstrating how this approach can be useful for further enhancement of IWP. The SATOPI project make use of domain knowledge that are particularly useful for glaciologist or other research scientists to have easy navigational experience to retrieve glacier and glacial lakes related information along with geospatial resources from a single page.

## **What is HKKH Partnership Project?**

The HKKH Partnership Project aims at consolidating institutional capacity for systemic planning and management of protected area ecosystem over three national parks in the Hindu Kush-Himalayan region. These three national parks are Sagarmatha National Park in Nepal, Central Karakoram National Park (CKNP) in Pakistan and Qomolangma Nature Preserve (QNP). The project is being implemented at multiple scales involving stakeholders from local, national and regional partners in the region, and domain experts from the region and beyond. The project implementation involved various applied research and development works that required integration of diverse data and information, interpretation of various forms of information and knowledge, e.g. system dynamic models, simulation results, historical data bases, technical reports and so on.

## **IWP Architecture**

The integrated web portal adopts state-of-the-art web-based tools and technologies for improving access, interactive use and exchange of data, information and knowledge. IWP has been built with the capability to integrate a wide variety of data and information resources such as project documents, bibliographic information, GIS and model data, interactive maps, satellite images and research data and more importantly all of these data and information are associated with metadata. Access to these information resources by multiple users i.e. technicians, researchers, scientists, decision makers etc. enhances collaborative learning and sharing culture among the stakeholders of all three national parks in Nepal, Pakistan and China of the HKKH Partnership Project.

The major components of IWP as shown in Figure 1 are thematic and general content management, forum and knowledge management, access and menu management, together with a workflow system in place.

IWP content management system (CMS) consists of three criteria for structuring information resources: firstly, by defining *information type*; secondly, by *Identifying rules* e.g. selecting content hierarchy within the information type, and lastly, *applying the standards for consistency* e.g. applying style sheet for common visual identity.

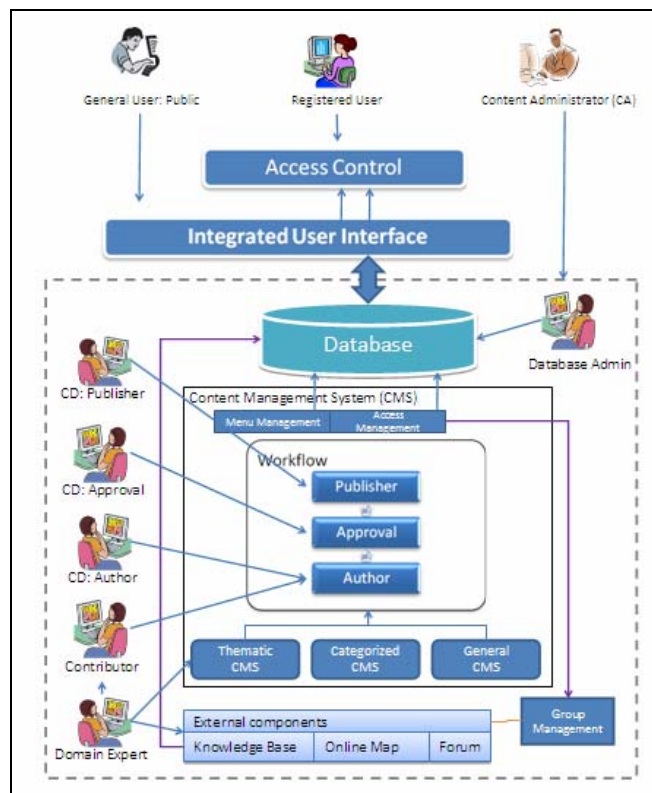


Figure 1. System Architecture showing how users can interact with IWP components

During structuring process of the information resources it is essential for administrator or the content developer to understand the *Content Workflow Cycle*. There are three stages of content workflow in IWP system: *Author*, *Approve* and *Publish* as illustrated in the Figure 1. There are various *information types* a content developer can author, approve and publish. Depending on rights assigned, a content developer can author, approve and/or publish limited or all types of contents.

Once the content is authored, then it goes for approval to those who have the approval rights.. Any updates or edit if required can be made by the same person as well. Once the content is approved, it then has to be published according to the workflow process. A content developer with publishing rights can review the content and use his/her discretion to publish the content iteratively. The content can be tagged with certain life (time period) and will automatically be archived. The archived contents can be searched

for future purpose. Thus, IWP CMS provides a complete workflow cycle that includes authoring, approving, publishing and archiving of contents.

## Knowledgebase Module

The knowledgebase system in IWP allows systematic cataloguing of metadata in a standardized format for information resource description. Here, the metadata record describes the typical attributes for spatial as well as non-spatial data cataloguing.

The system is based on GeoNetwork open source platform which is highly interoperable online metadata management system which uses ISO 19115 and Dublin Core ISO 15836. In addition, it uses modified Dublin Core for non-spatial recording such as bibliography. Figure 2 shows the list of spatial data catalogue.

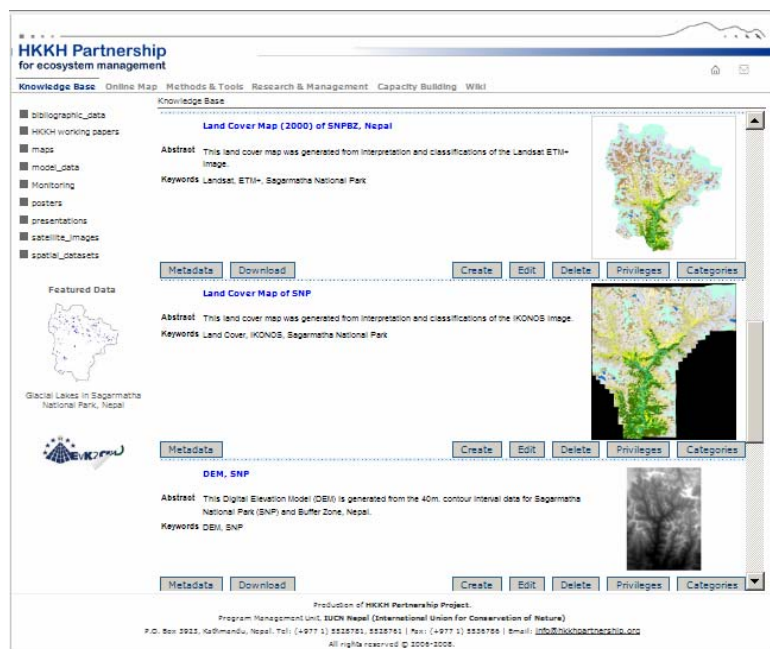


Figure 2: Spatial data catalogue page

Other important part in this system is categorization and keyword attribute assigned to the metadata. These metadata categories for the project includes bibliography data, project working papers, maps, system dynamic model data, monitoring, raster and vector datasets.

Besides cataloguing of information on data, GeoNetwork is also capable of providing distribution services and facilities to exchange information with other catalogues as a part of the network. As HKKHPartnership Project area is located at three national parks of different countries, GeoNetwork helps in organizing geospatial information by creating network of catalogues between park offices which allows users to find information in a single interface.

## Why Topic Maps

Topic Maps will be useful for further enhancement of IWP by improving in disseminating the resources. As described in previous sections, the knowledgebase is a catalogue of information about data i.e. metadata and it allows this metadata to be organized in a content-category. Similar is the case with CMS system which adopts hierarchical approach to organize the resources. The limitation to this sort of approach is that the user has to navigate through different layer components and its categories. The solution would be to construct a master index layer (e.g. back-of-book index page) where project relevant subjects are listed in terms of index with defined association between them. Thus, the user will be able to retrieve information correctly on a given subject through a single page user interface. Topic Maps are such paradigm, where each subject is a topic.

Topic Maps technology is an ISO standard (ISO/IEC 13250:2003) for the semantic structuring of knowledge networks and for the interchange of knowledge. It is especially useful in the project scenario for several reasons:

1) Structure to provide contextual information:

The building blocks of Topic Maps are topics, its occurrences and associations between topics. A topic represents a subject and can contain occurrences that relate to this subject (a reference to knowledgebase metadata record, a reference to an image describing the subject etc.). In addition, topics can be associated to each other. The subject is a concept and concepts are connected by network of associations. This provides a structure that corresponds to the way we store knowledge and a way to navigate between topics in Topic Map. It also let users ask questions about the knowledge within the Topic Map, for example, where a certain glacier lake is located, or what is common to a group of glaciers and so on. This gives meaning to the data that is kept in the Topic Map, and thus, Topic Maps technology provides a semantic layer over data and facilitates its accessibility.

2) Easy to pull together information from a variety of sources:

As mentioned in above, each topic can contain occurrences that can either be data or references to information resources that are related to the concerned topic. These references can point to any resource such as text, graphs, model, image, video, geospatial data, Earth Observation (EO) products, or simply metadata records in the knowledgebase. This makes Topic Maps an excellent choice for creating a complex metadata layer over information resources.

3) Hidden topics of interest are made explicit:

The Topic Maps standard provides the ability to merge Topic Maps. This means that when a Topic Map that describes the geospatial data products (where, for example, each product is represented by a different topic, and those topics are associated to a mission, or to a period, etc.), and a Topic Map that describes glaciers and glacial lakes are given, these two Topic Maps can be merged in order to achieve an integration of the resources. It means that this approach gives holistic view to the resources for any given topic.

## Approach to integrate glacier and glacial lakes resources

The International Centre for Integrated Mountain Development (ICIMOD) has collected data over 3252 glaciers and 2323 glacial lakes in Nepal. The inventory data gives information on glaciers, glacial lakes, areas where glacier lake outburst flood (GLOF) events have occurred and lakes that could pose a potential threat of GLOF. These data include different characteristics of lakes such as physical condition, spatial data, orientation or classification. In addition, geospatial data are also available over the region.

However, the information collected is kept across many different platforms in the form of database, raster and vector data, publications, scientific findings and literature. To be scientifically useful, all these information resources must be integrated with knowledge in order to improve our understanding of the GLOF phenomena. The SATOPI project uses these data resources by designing two separate ontologies – geospatial data and domain knowledge and develop user interface for easy access.

As mentioned, Topic Maps is a standard for knowledge representation and information integration. In Topic Maps, each topic represents a subject, symbol or representation as shown in Figure 2. Imja Tsho is a topic representing the glacial lake. So, Imja Tsho and other lakes are grouped together under the topic type “potentially dangerous lake”. Also, Imja Tsho is located in Nepal meaning that the topic “Imja Tsho” is associated to another topic “Nepal”. Similarly, “see also” relationship expresses an association between topics like “Dudh Pokhari” which is also considered as “potentially dangerous lake”.

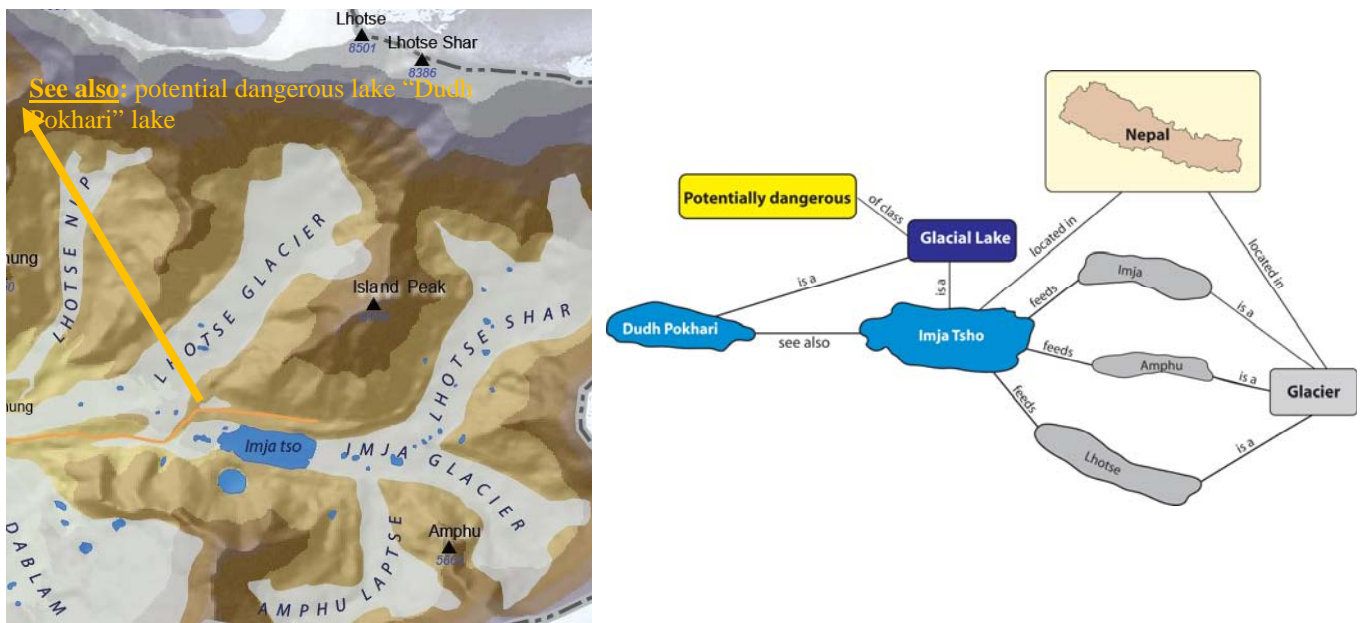


Figure 3: Topic Map example introducing the basic topic map concept

In addition, topics in Topic Map can hold other information called occurrences. These are something like page numbers provided in a back-of-book index page. Like

associations, occurrences are also typed. Therefore the topic “Imja Tsho” might have for example an occurrence of type “Article” (where the actual URL of the article is kept as the occurrence value).

These topics provide the conceptual skeleton of the topic map. These topics together are termed as Topic Maps Ontology. Ontologies are very useful when authoring Topic Maps as they help to identify borders of the domain of knowledge that it represents.

## Topic Navigation

The previous section describes that the ontology are a result of an intellectual process performed by Topic Map authors to model information from resources and domain experts. This provides a structure that corresponds to the way we grasp knowledge. Such a structure provides a way to navigate between topics in the Topic Map. This gives context and meaning to the data that is kept in the Topic Map.

The interface provides single access point to the available data, which allows interaction with the data using domain specific terminologies. This is done by authoring two Topic Maps: The first Topic Map Ontology is designed to include the above described data about glaciers and glacial lakes where a Topic Map is created by populating this ontology with the data from the existing resources or by connecting the resources to the Topic Map (shallow copy). In the same manner, a second Topic Map is created which represents the geospatial data. The two Topic Maps are merged and a user interface is provided in order to query the overall data using domain specific terminologies.

Search for: Glaciers, Lakes, Events, Photographs.

### Lakes Search Form [reset]

Search by name or identifier:

Localization:

Country:

Basin:

Sub-Basin:

Coordinates:

Types

Classification:

Activity:

Characteristics

Length:

Elevation:

Area:

Has Parent Glacier:  Any  Yes  No

Distance to Glacier:

Orientation:

Drainage Condition:

### Matching Lakes

Displaying 25 out of the 2316 matching lakes. [Display the full list.](#)

Identifier	Name
Gka_gl 51	
Gka_gl 57	
Gmar_gl 40	
Gmar_gl 50	
Gset_gl 1	
Gtri_gl 10	
Kar_gl 29	
Kar_gl 84	
Kar_gl 96	
Kdu_gl 91	
Kdu_gl 104	
Kdu_gl 105	
Kdu_gl 331	
Kdu_gl 368	
Kdu_gl 438	
Kin_gl 16	Panch Pokhari
Kmu_gl 29	
Kmu_gl 239	
Kmu_gl 260	

Figure 4: Topic type: “Lake” search page

The first page is the list of topic types; the respective topic types can be accessed by clicking on the respective topic names such as glacier, lakes, GLOF events etc. which displays a dedicated topic search page. The content within search page is further

subdivided by means of search criteria, appearing on the left side of the browser window and matching topics are displayed on right side of the page as shown in the Figure 4. These matching topics changes accordingly to the selection of search criteria.

After finding the desired topic, the user will be able to view the detail page as shown in Figure 5. This page is separated with building blocks of Topic Maps i.e. topics, names, association and occurrences along with its values i.e. it provides a concept of one page per topic.

Similarly, association block provides association of navigation from one topic to another topic. In this way, the front end can access the data without taking into account the complexity of location of different parts of the data available.

Search for: Glaciers, Lakes, Events, Photographs.

**Lake: Imja Cho (Kdu\_gr 350)**

**Description**

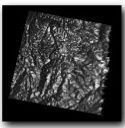
The Imja Cho lake is located in Nepal, in the Dudh Koshi sub-basin. The coordinates of its approximate center are 27°53'59" latitude North and 86°55'23" longitude East. It is oriented towards South-West.

The lake has a mean length of 410 meters and its width is **not available**. Its area is 48811.05 square meters and its elevation is 5023 meters.

The Imja Cho lake is classified as **Supraglacial Lake (S)** and its drainage condition is **Closed Lake (Cs)**.

The Imja Cho lake did not experience any known GLOF events in the past and is now considered as **potentially dangerous**.

The date at which the data has been collected is **not available**.



This is the ENVI-SAT image of Nepal and China. It is acquired on 9/27/2008 12:00:00 AM.

**Relations**

The Imja Cho lake is located in:

- Country: Nepal
- Basin: Koshi
- Sub-basin: Dudh Koshi

No merging information about the Imja Cho lake could be found.

No merging information about the Imja Cho lake could be found.

The Imja Cho lake is fed by:

- Glacier: Imja (Kdu\_gr 160, G086943E27899N) (Distance: 0 m) , Amphu (Kdu\_gr, 161, G086943E27880N), Lhotse (kdu\_gr 156, G08691E27919N)

No events related to the Imja Cho lake could be found.

**Resources**

Display 10 out of the 86 matching photographs.

Local ID	Device	Year	South	North	West	East	Width	Height	Abstract
3683	EnviSat	2008	27°17'00" N	28°25'33" N	86°22'07" E	87°38'35" E	1.14	1.27	This is the ENVI-SAT image of Nepal and China. It is acquired on 9/27/2008 12:00:00 AM.
3687	EnviSat	2008	27°22'29" N	28°29'08" N	86°08'08" E	87°09'15" E	1.11	1.05	This is the ENVI-SAT image of Nepal and China. It is acquired on 9/28/2008 12:00:00 AM.
3688	EnviSat	2008	27°17'09" N	28°25'36" N	86°21'47" E	87°38'10" E	1.14	1.27	This is the ENVI-SAT image of Nepal and China. It is acquired on 7/19/2008 12:00:00 AM.
3726	EnviSat	2008	27°19'10" N	28°25'50" N	86°08'35" E	87°09'42" E	1.11	1.05	This is the ENVI-SAT image of Nepal and China. It is acquired on 5/19/2008 12:00:00 AM.
3721	EnviSat	2008	27°17'56" N	28°24'30" N	86°07'08" E	87°10'12" E	1.11	1.05	This is the ENVI-SAT image of Nepal and China. It is acquired on 4/10/2008 12:00:00 AM.

Figure 5: Description page of Imja Tsho

Moreover, different parts of the data can be provided as different Topic Maps. For example, the geospatial data can be provided in a Topic Map which contains only relevant geospatial data and locations, but does not include glaciers and GLOFs related information. This information can be accessed through the second Topic Map. The ability to merge those two Topic Maps provides a second level of federation of data. In other future applications, one of those Topic Maps can be reused. For example, the same geospatial data can be used in different research in the same geographic area. The glaciers and GLOFs related Topic Map can be merged in another relevant application.

Both pages illustrated in the figures demonstrate the strength of using Topic Maps and design template consisting of building blocks of Topic Maps. This facilitates the navigational experience to the users with a sophisticated and rich user interface.

## Conclusion

IWP has been built with a customized content management system by establishing a simple and easy management of information resources through a workflow process and a common framework for information authoring, approving, publishing and archiving. IWP has proved useful to capture all relevant information during the entire course of implementation of the project and serves as a project's legacy even after its completion. Moreover, the innovative tools and technologies developed to realize IWP can easily be up-scaled to similar other initiatives in other mountain regions.

The paper demonstrates how Topic Maps technology can be used in order to integrate relevant geospatial data in an application domain. In other words, it provides a semantic layer over data and facilitates its accessibility to IWP knowledgebase and other domain specific resources by authoring Topic Map on application domains of HKKH Partnership Project.

Further, parts of the Topic Map can be model for different views of other application domain thus providing interoperability among heterogeneous data sources. This technology can be used as the electronic front door to discover and access geospatial data and resources through semantically rich integrated approach. The need of such integration is clearly endorsed by initiatives such as spatial data infrastructure (SDI).

## References

Bajracharya R. M., *Building Mountain Knowledge Hub in the Hindu Kush-Himalaya Region*; First International Topic Map User Conference, Oslo Congress Center, Oslo, Norway, 2007

Bajracharya S. R., Mool P. K., Shrestha B.R., *Impact of Climate Change on Himalayan Glaciers and Glacial Lakes. Case Studies on GLOF and Associated Hazards*, 2007

Pepper S., *Everything is a subject – the vision of subject-centric computing*; Topic Maps, 2008

Pinchuk R., R., Valentina B., S. Bajracharya, R. M. Bajracharya; *Integration of Earth Observation Data and Domain Knowledge Using Topic Maps*, TOWARDS eENVIRONMENT, 2009

Scharl A., *Towards the Geospatial Web: Media Platforms for Managing Geotagged Knowledge Repositories*, 2007

Ticheler J., *GeoNetwork - A Geographic Library on the web*; GEOconnexion International Magazine, 2008