A Conceptual Model of Ecological Observation Service Supporting Data Life Cycle

Taesang Huh, Jae-Hyuck Kwak, Sangwan Kim, Eunkyu Byun,
Geunchul Park, Soonwook Hwang, Hoe-Kyung Jung*
Korea Institute of Science and Technology Information, Korea
*Paichai University, Korea
E-mail : tshuh@kisti.re.kr, jhkwak@kisti.re.kr, sangwan@kisti.re.kr, ekbyun@kisti.re.kr,
gcpark@kisti.re.kr, hwang@kisti.re.kr, *hkjung@pcu.ac.kr

1. Introduction

As ecology is becoming more complex and rapidly changing into a more open, collaborative and data intensive science, many ecologists are asking for implementation of the ecological observation system that is capable of acquiring, managing, analyzing and searching the large size of various types of data such as images, videos, audios, records and coordinate information of an investigation equipment, and providing statistics and analysis services.[1][2] Ecological visualization service is mostly based on geographic information system (GIS) since the topographical position in data characteristic of ecological observation data is very important. The amount of data is still growing due to the advent of new investigation technology like various sensors, high performance radar and intelligent drone, and due to discovery of new ecospecies. In this paper, we discuss the conceptual model of scalable Ecological observation service allowing various ecologists and eco-communities to have a good collaboration research and share Korean ecological data with international ecological working groups.

2. Related Works

Ecological observation service needs reliable and available cyber infrastructure supporting data-intensive science for analysis and searching of massive amount of data for long-term experiments caused by the emergence of new investigation technology, the discovery of new ecospecies and environment change. Recently, well-known cloud storage solutions are also a more scalable and efficient way to store data with the ability to add more capacity, performance and objects whenever a new node is added to the cloud. The other important requirement is to design data management workflow, as data life cycle, including the planning, collection, assurance, description, preservation, discovery, integration, and analysis of data.[2][3] It is moreover needed to share data with international legacy ecosystems which are several ecosystems outside of Korea such as National Ecological Observatory Network (NEON), Environmental Change Network (ECN) and a few systems in Long-term Ecological Research for years.[4][5]

3. Ecological Observation Systems

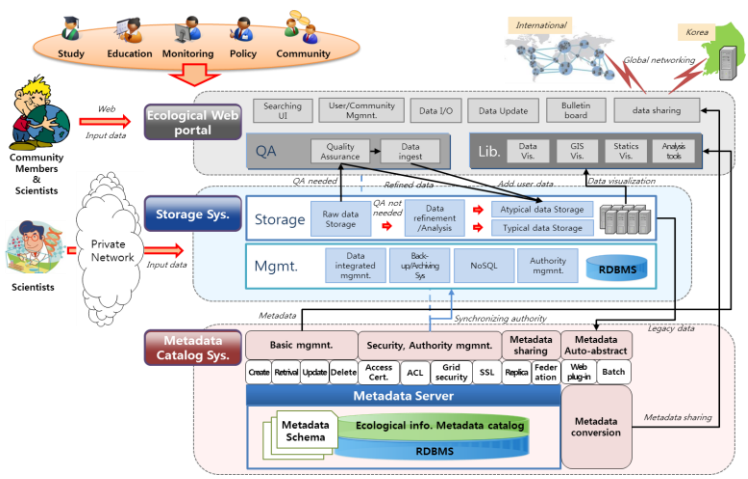

Figure 1. Ecological Observation Systems Infrastructure

* Corresponding Author

The our conceptual design approach is to primarily consider the raw input data and changed data format trough data analysis, data refinement, data service and sharing data with international and domestic communities. In order to give a fast searching and analysis of data, these systems need to generate manually or automatically metadata for raw data. As shown in Figure 1, ecological observation systems Infrastructure, input data came through public web or private network get into ecological observation systems and are refined by QA (Quality Assurance) process. After that, the data is ready for being analyzed by ecologists. Ecological observation systems has mainly three sub-systems consisted of web portal, storage system and metadata catalog system.

The goal of ecological web portal is to provide ecological data and information through dynamic web applications. Various mechanisms for searching for data are provided from metadata web application programs using metadata catalog system and many different type of information (video, audio, statistics, text-based abstract and analyzed reports, and others) based on geographic information system. It helps the scientists' analysis through analysis tools and libraries. Finally, it shows effectively some data with diverse graphical interfaces based on GIS.

Storage system starts from raw input data which isn't refined and is closely connected with its metadata. Raw data without quality assurance needs objective data refinement and additional accurate data ingest process. Data is decided to go into atypical or typical data storage in this workflow. Storage system have to manage metadata of data based on RDBMS and be consisted of scalable storages like cloud storages due to long-term data accumulation.

Metadata catalog system supports not only raw metadata acquired from user input but also the new metadata standards representing an expansion to previously existing metadata and new high-level, or data abstract elements, and spatial data element were added for compliance. Metadata conversion component generates EML (Ecological Metadata Language) to systematically share metadata with international ecological communities. [6]

4. Conclusion

In this paper, we propose and design the model of ecological observation service to support data life cycle and allow ecologists' research by acquiring, managing, analyzing, searching ecological data and sharing ecological data globally. It is consisted of three main subsystems such as ecological web portal system proving various mechanisms for searching for data and analysis tools, Storage system controlling atypical and typical data working along with other subsystems and metadata catalog system having general management functions with powerful security and converting the modes of data expression.

As a result, we expect that ecological observation service will enable to create general knowledge through long-term, inter disciplinary research, synthesis of information, and development of theory and inform the ecological research groups by creating well designed and useful databases. Furthermore, it will provide policymakers to decide ecological policy by decision supporting service.

5. References

[1] C. L. Redman, J. M. Grove, and L. H. Kuby, "Integrating social science into the long-term ecological research (LTER) network: social dimensions of ecological change and ecological dimensions of social change", Ecosystems, Vol. 7, No. 2, (2004), pp. 161-171.

[2] C. Strasser, R. Cook, W. Michener, A. Budden and R. Koskela, "DataONE: promoting data stewardship through best practices", *Proceedings of the Environmental Information Management Conference 2011*, University of California, USA, (2011), pp. 126-131.

[3] D. L. Henshaw and G. Spycher, "Evolution of ecological metadata structures at the H.J. Andrews Experimental Forest Long-Term Ecological Research (LTER) site", North American science symposium: toward a unified framework for inventorying and monitoring forest ecosystem resources, Guadalajara, Mexico, (1998), pp. 445-449.

[4] G. Spycher, J. B. Cushing, D. L. Henshaw, S. G. Stafford, N. Nadkarni, "Solving problems for validation, federation, and migration of ecological databases", Global networks for environmental information, Proceedings of Eco-Informa '96, Lake Buena Vista, Environmental Research Institute of Michigan (ERIM), U.S.A, pp. 695-700.

[5] J. H. Porter and J. T. Callahan, "Circumventing a Dilemma: Historical approaches to data sharing in ecological research", Environmental Information Management and Analysis: Ecosystem to Global Scales, Taylor & Francis Ltd, London, U.K., (1994), pp. 193-202.

[6] Ecological Metadata Language (EML), https://knb.ecoinformatics.org/#external//emlparser/docs/index.html, (2014)