

**MEETING THE CHALLENGES OF AN INTERNATIONAL, GRASSROOTS  
ORGANIZATION OF SITES DEPLOYING SENSOR NETWORKS: THE GLOBAL  
LAKE ECOLOGICAL OBSERVATORY NETWORK (GLEON)**

**B. J. Benson<sup>1</sup>, L. Winslow<sup>1</sup>, P. Arzberger<sup>2</sup>, C. C. Carey<sup>3</sup>, T. Fountain<sup>4</sup>, P. C. Hanson<sup>1</sup>, T. K. Kratz<sup>5</sup>, S. Tilak<sup>4</sup>**

<sup>1</sup>Center for Limnology, University of Wisconsin-Madison, 680 N. Park Street, Madison, WI 53706 USA; <sup>2</sup>California Institute for Telecommunications and Information Technology, and the Center for Research on Biological Systems, University of California-San Diego, La Jolla, CA 92093-0043 USA; <sup>3</sup> Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853 USA; <sup>4</sup>San Diego Supercomputer Center, University of California-San Diego, La Jolla, CA 92093-0505 USA; <sup>5</sup>Trout Lake Station, Center for Limnology, University of Wisconsin-Madison, 10810 County Highway N, Boulder Junction, WI 54568 USA

**Abstract**

Realizing the full potential of embedded sensor networks to generate new scientific knowledge requires the sharing of data and expertise and the interdisciplinary collaboration of environmental scientists and information technologists and managers. The Global Lake Ecological Observatory Network (GLEON) is a grassroots network of limnologists, ecologists, information technology experts, and engineers with a common goal of building a scalable, persistent network of lake ecology observatories. GLEON's technological and organizational innovations provide models for how a grassroots organization can function to catalyze science based on environmental observing networks. Evolving solutions within GLEON to technological and organizational challenges include ways for sharing expertise, the development and deployment of software to enable effective management and sharing of sensor network data, the generation and documentation of GLEON operating principles and procedures, and the training of students in the new technology and large-scale scientific collaboration.

**Keywords:** sensor networks, cyberinfrastructure, ecoinformatics, environmental observatories, GLEON, lakes

**1. Introduction**

Many of the environmental challenges being addressed by scientists in today's world are regional or global in scope such as the effects of climate change, land use change, invasive species or human population growth and distribution. Often the study of these complex issues involves controls and interactions at multiple temporal and spatial scales. Embedded sensor networks are making new contributions to environmental sciences by extending the scales of spatial and temporal measurement (Estrin et al. 2003, Porter et al. 2005). Fully realizing the potential of this technology to generate new scientific knowledge will require the sharing of data and expertise and the interdisciplinary collaboration of environmental scientists and information technologists and managers.

A significant challenge then is finding effective ways to support these collaborations and scientific investigations that span regions and the globe. The Global Lake Ecological Observatory Network (GLEON) is a grassroots network of limnologists, ecologists, information technology (IT) experts, and engineers with a common goal of building a scalable, persistent

network of lake ecology observatories ([gleon.org](http://gleon.org); Kratz et al. 2006). Lakes, in particular, are a key ecosystem under stress in this changing world. The stated mission of GLEON is to facilitate interaction and collaboration among an international, multidisciplinary community of researchers focused on understanding, predicting, and communicating the impact of natural and anthropogenic influences on lake ecosystems by developing, deploying, and using networks of emerging observational system technologies and associated cyberinfrastructure. The inaugural GLEON meeting occurred in March 2005. As of April 2008, there were 298 individuals affiliated with GLEON, representing 31 countries. There have been six GLEON meetings attended by an international group of lake scientists and professionals from technical fields involved in sensor technology and information systems. Resources for building the GLEON community and its capacity for scientific collaborations have recently been significantly augmented by an NSF Research Coordination Network (RCN) grant. Technology development that is benefiting both GLEON and the Coral Reef Observatory Network (CREON) has been fostered through grants from the Gordon and Betty Moore Foundation.

We address here some evolving solutions to technological and organizational challenges faced by an international, grassroots organization of sites deploying sensor networks such as GLEON. These solutions include ways for sharing expertise, the development and deployment of software to enable effective management and sharing of sensor network data, the generation and documentation of GLEON operating principles and procedures, and the training of students in the new technology and large-scale scientific collaboration.

## **2. Methods and Techniques**

### **2.1 Sharing expertise.**

The sharing of expertise is an important benefit of a research network such as GLEON. The technology associated with sensor networks is a relatively new and rapidly evolving field. Multiple solutions and approaches exist within the diverse collection of sites. Communication at meetings, ongoing working groups, and the GLEON website all represent channels for sharing expertise. In addition, GLEON has developed the Lake Information Database ([gleon.org/lakes](http://gleon.org/lakes)), a web-accessible database of information about GLEON lakes and the sensors deployed on them. The displayed information includes an overall lake description, values for lake characteristics such as lake area and nutrient concentrations, a list of measurements being taken and the sensors that are used for these measurements. GLEON members enter information into the database through a web-enabled application that has both an administrative and user interface. The user interface includes the opportunity to add new vocabulary for measurement types and sensors as well as guidance text on entering information. User additions to the controlled vocabulary are then vetted by a subgroup of the GLEON Steering Committee.

### **2.2 IT development and deployment.**

Many GLEON sites have acquired or will be acquiring sensor technology and know how to deploy sensors and download data to a repository on a local computer. However, it is often the case that this repository, often a text file archive, is not easily shared, queried or made accessible via the Internet. To eliminate these gaps, we undertook the creation of information management system software to allow scientists to access the data via the Internet. This development has been supported by a number of synergistic grants and collaborations. To date, the software has been installed at four GLEON sites: Lake Erken in Sweden, Lake Sunapee in New Hampshire, Lake Annie in Florida, and the North Temperate Lakes LTER in Wisconsin.

These sites vary in the number of sensors deployed, the extent of legacy data, and the extent to which an information management system was already in place for non-sensor data. A team of people involved in deployment package development traveled to each site to install the system, and, in some cases, assisted with instrumentation deployment.

There were multiple components of the installation process (i.e., install, document, train, test, and evaluate) beyond the actual installation of the software that automated the data flow from downloaded text files to an Internet accessible database. For documentation, an installation report was prepared ([gleon.org](http://gleon.org)) for each site installation, and a repository of required technologies was maintained on the GLEON website. Local staff were given an overview of the technology and trained to change the system configuration and troubleshoot problems. System operation was tested under continuous operation conditions, individual component shutdown, and system reboot. The unique installation process at each site was evaluated.

An important part of the installation process is a site preparation component that is ideally generated by site personnel prior to installation. Documentation is required of the physical instrumented buoy system, the sampling regime, data download frequency and storage location, the vocabulary used for measurement, the logical hierarchy that allows the physical system to be represented in the data structure, a description of any legacy sensor data, and security constraints. The local site situation can generate additional requirements. For example, Lake Sunapee, which has a very active lake association, needs solutions for making data available in near-real time to the public; these solutions will be co-designed and implemented through a recently awarded NSF CI-Team grant.

Early on, it was recognized that to facilitate data sharing within GLEON, member sites without information management infrastructure would need to be brought up to speed. The requirements included managing high-resolution sensor network data and making those data web accessible through inexpensive or free tools and software. In addition, an emphasis was placed on ease of use and robustness as many GLEON members lack IT support. The envisioned deployment software package (Figure 1) consists of both off-the-shelf and custom software. The database is MySQL v.5 Community Server, and the database administrative tool is MySQL Administrator. Logger debriefing uses the Campbell LoggerNet (or PC208w). The data model (Vega; Winslow et al. 2008) provides storage flexibility, accommodating reconfigurations of the sensor network without changes to the database schema. During deployments Open Source DataTurbine ([www.dataturbine.org](http://www.dataturbine.org); Tilak et al. 2007) was tested. DataTurbine is open-source streaming data middleware that provides reliable data transport, a framework for integrating heterogeneous instruments, and a suite of services for data management, routing, synchronization, monitoring, and visualization. Inca ([inca.sdsc.edu](http://inca.sdsc.edu)) was tested to monitor the software and data management infrastructure and allow remote monitoring and troubleshooting. More technical details can be found in the installation reports ([gleon.org](http://gleon.org)).

### **2.3 Data sharing.**

The GLEON deployment package allows multiple destinations for the data stream, and in practice, the data from each of the deployment sites have been streamed to a central repository, in addition to the local repository. These data are then accessible through the GLEON web-site via custom query tools. The controlled vocabulary that is being developed for the Lake Information Database and the GLEON deployment package sets the stage for expanding data discovery and access beyond sites employing the deployment package. The controlled vocabulary includes measurements, sensors, and units.

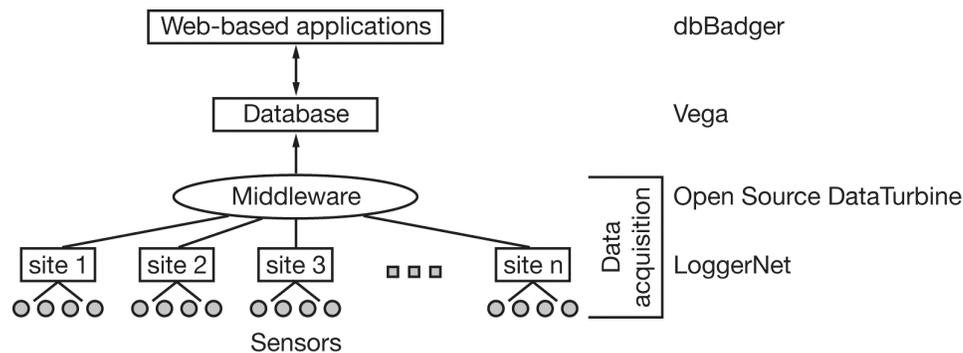


Figure 1. This diagram represents the data plane aspect of the system architecture for the GLEON deployment package. Data from sensors at distributed GLEON sites are routed to a centralized database as well as local databases through middleware. Web-based applications such as dbBadger allow users to query the database. The shared database is built on the Vega data model. Open Source DataTurbine has been tested at deployment sites for use as the middleware. System monitoring via Inca (not shown) pervades all the layers of the architecture and connects with components such as hardware (sensors and compute nodes), middleware, and databases.

### 3. Organizational issues

GLEON created operating principles and procedures to guide the growth and evolution of the network ([http://www.gleon.org/media/GLEON\\_OpPrincProc.pdf](http://www.gleon.org/media/GLEON_OpPrincProc.pdf)) as well as to have clear and transparent operating procedures. This document describes the organizational goals, values and principles, structure (including member and steering committee roles), and the policy on sharing of data. This document drew from the Pacific Rim Application and Grid Middleware Assembly (PRAGMA). This explicit statement of how to address these organizational issues increases organizational effectiveness, and the process of generating the structure and agreements has contributed to community cohesiveness.

### 4. Education

An important focus of the GLEON Research Coordination Network (RCN) is to inform, train, and mentor students while simultaneously preparing the next generation of scientists for large, collaborative, international, interdisciplinary science. The education of today's graduate students needs to prepare them to lead and collaborate within these larger, more complex research environments that are increasingly becoming a more extensive mode of conducting environmental science. Participation in GLEON provides several benefits to students: networking with other students and researchers across disciplines, learning new skill sets, the opportunity to experience a leadership role in an emerging organization, and traveling to different GLEON sites. Student activities have included conducting an informational meeting about GLEON for students attending the International Society for Limnology (SIL) meeting in Montreal in August 2007 and organizing the application process for students wanting to receive support to attend the GLEON VI meeting in Florida in February 2008 and the GLEON VII meeting in Sweden in September 2008. Fifteen students representing multiple disciplines and countries attended the GLEON VI meeting after a competitive application process. At GLEON VI, the students created the GLEON Student Travel-Funding Program, a cross-site collaborative effort for students to visit other GLEON sites and gain knowledge and experience within the network. This project is funded by the RCN.

## 5. Discussion

Grassroots networks such as GLEON can provide significant assistance to participating sites for implementing sensor networks and extending the technological development that supports sensor network deployment and information systems. Since its first meeting in 2005, GLEON has developed a web-accessible database of lake characteristics, measurements and sensor information for participating sites, developed and deployed a software suite that has enabled several sites to manage and share sensor data, conducted five more GLEON meetings at which extending technological developments that support the sites and network was a goal of the meetings. The community has benefited from relationships developed during the meetings, and there have been many instances of people with expertise from one GLEON site traveling to another site to assist in sensor deployment and various system design issues.

The data that are being shared within GLEON are pivotal to new scientific understanding. The sensor network measurements have potential to provide new estimates of ecosystem rates, better calibration of models, and identification of key controls over ecosystem process across multiple scales. They have illuminated the impact of events in near real-time such as the role of typhoons in restructuring lake ecosystems in a remote lake in the mountains of Taiwan (Tsai et al. 2008). Work is underway that capitalizes on the interplay between high frequency data and the development and extension of models of lake ecosystem processes. Now is an exciting time in which to explore the ways in which the new scales of measurement provided by sensor networks can expand the questions and models that scientists address, and GLEON is actively engaged in this exploration.

The grassroots organizational paradigm has contributed to the successes of GLEON (Hanson 2007). The openness of the organization to innovation by individuals and the ability to capitalize on heterogeneity across the sites, technologies, and scientific approaches are strengths of the grassroots approach. Considerable flexibility to foster multiple solutions to problems and the awareness of the importance of building trust among participants have both promoted cohesiveness within the GLEON community. The Coral Reef Ecological Observatory Network and the National Phenology Network are other examples of broad scale networks using a grassroots organizational paradigm.

Leaders and technology developers within GLEON are aware of multiple new challenges to be addressed. Installation of the deployment package for sensor networks can be streamlined so that it will become possible for the installation team to do remote deployments with the assistance of local staff. Sharing of real-time data streams is now possible for those sites using the deployment package. A future goal is to extend the portal for sharing data in ways that allow participation by the heterogeneous collection of information management systems that exist across the GLEON sites. GLEON members have expressed an interest in sharing a wider scope of data beyond the sensor network data, such as spatial information layers. An ongoing challenge is to find the resources to support new cyberinfrastructure development within GLEON and to facilitate participation in GLEON meetings and activities. As GLEON continues to grow, we will undoubtedly face additional challenges related to organizational structure and perhaps need to address issues related to an optimal network size.

## 6. Conclusions

GLEON's technological and organizational innovations provide models for how a grassroots organization can function to catalyze science based on environmental observing

networks, provide assistance to participating sites in implementing sensor networks, extend the technological development that supports sensor network deployment and information systems, and develop tools to promote sharing of expertise and data across a research network. Interdisciplinary partnerships of lake scientists, engineers, computer scientists, educators, and information technology and management experts are required to make the vision for GLEON a reality.

### **Acknowledgments**

We gratefully acknowledge support from the U.S. National Science Foundation through grants DBI-0639229, NEON 0446802, NEON 0446017, DEB-0217533, OCI 0627026 and OCI 0722067 and from the Gordon and Betty Moore Foundation.

### **References**

- Estrin, D., W. Michener, G. Bonito, and workshop participants. 2003. Environmental cyberinfrastructure needs for distributed sensor networks: a report from a National Science Foundation sponsored workshop. Scripps Institution of Oceanography, La Jolla, CA. 12-14 August 2003.
- Hanson, P. C. 2007. A grassroots approach to sensor and science networks. *Frontiers in Ecology and the Environment* 5(7): 343–343.
- Kratz, T. K., P. Arzberger, B. J. Benson, C. Chiu, K. Chiu, L. Ding, T. Fountain, D. Hamilton, P. C. Hanson, Y. H. Hu, F. Lin, D. F. McMullen, S. Tilak, and C. Wu. 2006. Towards a Global Lake Ecological Observatory Network. *Publications of the Karelian Institute* 145:51-63.
- Porter, J., P. Arzberger, H. Braun, P. Bryant, S. Gage, T. Hansen, P. Hanson, F. Lin, C. Lin, T. K. Kratz, W. Michener, S. Shapiro, and T. Williams. 2005. Wireless sensor networks for ecology. *Bioscience* 55:561-572.
- Tilak S., P. Hubbard, M. Miller, and T. Fountain. 2007. The Ring Buffer Network Bus (RBNB) DataTurbine streaming data middleware for environmental observing systems. e-Science 10/12/2007, Bangalore, India.
- Tsai, J.W., T.K. Kratz, et al. 2008. Seasonal dynamics and regulation of lake metabolism in a subtropical humic lake. *Freshwater Biology*. *In press*
- Winslow, L.A. et al. 2008. Vega: A flexible data model for environmental time series data. Environmental Information Management Conference 2008. Albuquerque, NM.