出國報告（出國類別：其他）

出席「東亞暨太平洋地區長期生態研究網生態資訊管理研討會」報告書

服務機關：行政院農業委員會林業試驗所
姓名職稱：鄭美如 助理研究員
陳建文 研究助理
派赴國家：韓國
出國期間：96年10月16日～96年10月22日
報告日期：97年1月14日

目　　次

一、目的

長期生態研究已成爲全球之議題，在累積了大量且龐雜的生態資料後，如何將這些資料整合與分享，是近十年來長期生態研究最重要的議題。美國長期生態研究網的生態學家、生態資訊管理員與資訊技術學家經過多年的努力，已成功的發展出生態學資訊共通的標準語言 EML(Ecological Metadata Language)，可以讓龐雜的生態資訊相互交換分享，延續生態資料的生命與提升其使用價值，藉由分享，擴大研究尺度，對於生態研究有更全面更深入的視野。

林業試驗所於 2004 年起積極投入生態資訊整合與服務的研究，成立生態資訊研究小組，除赴美考察生態管理具代表性之研究站外，並派員赴美學習最新技術，同時與美國長期生態資訊管理人員持續作各項技術交流與研討。

2005 年 7 月於北京召開第一次東亞暨太平洋地區長期生態研究網（East Asia and Pacific Regional Network ILTER, EAP-ILTER）資訊管理研討會，同意以 EML 爲其資料交換之標準。2006 年 2 月於台北舉行之第二次生態資訊管理研討會便由林試所團隊與數位美國資深資管員詳實介紹 EML 結構、編輯、資料結構與生態資訊管理系統。

東亞暨太平洋地區長期生態研究網已於 2006 年的年會上組成資訊管理工作小組，成員共 16 名，由林試所林朝欽博士擔任主席，並預定於 2007 年 7 月在高雄六龜舉辦東亞暨太平洋地區長期生態研究資訊管理委員會會議與生態資訊管理研習會（EAP-ILTER Information Management Committee Conference and Ecological Information Management Workshop），由本所生態資訊研究小組成員與美國長期生態研究網數位資深資訊管理員共同擔任講師，並期能藉由會議討論出區域性研究資訊之分享機制，建立生態資訊管理網絡，讓不同地區不同語言之生態資料有一致的標準與資訊分享平台，逐步達成會員網間的研究資訊交流與研究合作，擴

大研究領域與尺度。

2007 年 10 月在韓國舉辦的第三次生態資訊管理研討會爲由本所生態資訊研究小組規劃設計課程內容，研討會之資訊管理技術教學與指導由本所林朝欽博士及其團隊，並邀請美國資深生態資訊管理員數名共同完成。承接前兩次之研習會與 7 月在台灣舉辦之研習會內容與協議，本次會議將完整的建立區域性長期生態資訊網資訊交換與共同研究之模式，消除區域資料收集的壁壘，實際應用於現有的研究議題上，並與美國及其他亞太地區之會員網進行經驗之交流及資訊與資料交換平台架構之研商，以建構亞太地區生態資訊整合與分享之機制與平台基礎。綜合言之，參與本次會議之目的爲：

1. 養成我國生態資訊管理之專業人才，並成爲服務其他生態學者與資訊工作者的界面角色，加速生態研究資料之整合與分享。
2. 以「生態資訊研究小組」之專業技能，實現技術傳承與分享，並穩固台灣在東亞暨太平洋地區生態資訊管理之領導地位，成爲美國的長期合作夥伴。
3. 擴展林試所之研究成果與生態資訊服務網站至國內生態相關之研究機構與大學，使得長期之研究資料得以完整之保存，進而再利用及創造更珍貴的研究價值。

二、過程

日 期	起 訖 地 點	工 作 任 務
10/16(二)	台北→首爾	去程
10/17(三)	首爾	參加會議
10/18(四)	首爾	參加會議
10/19(五)	首爾、珍島	參加會議與路程
10/20(六)	珍島、木浦	試驗地參訪
10/21(日)	木浦、首爾	試驗地參訪與路程
10/22(一)	首爾→台北	返程

三、成果與心得

本次在首爾的國民大學(Kookmin University)所舉辦的第三次東亞暨太平洋地區長期生態研究網生態資訊管理研討會是與長期生態研究、都市林及都市生態系國際研討會一同舉行，會議始於2007年10月16日，一連舉行四天室內的研討會，10月20~21 日兩天爲野外試驗地的參訪，除第一天爲都市林相關研究的研討外，其餘三天均爲長期生態研究與生態資訊管理相關之議題，會議議程如附錄一。此次會議中，林試所金恆鑣所長與本所生態資訊管理團隊，包含林朝欽博士、陳建文先生、邱振財先生與鄭美如小姐等五人與會，而生態資訊管理團隊包含與美國Dr. John Porter共同發表的報告有三篇(附錄二)，二張海報展示(附錄三、四)。

10月17日上午主要介紹國際長期生態研究與生態資訊管理之概論，由美國長期生態研究網總部的Deana Pennington博士介紹有關網絡技術在科學上的應用，接著由美國長期生態研究網資深生態資訊管理員John Porter博士介紹東亞暨太平洋地區的生態資訊管理訓練，而由林朝欽博士介紹使用生態後設資料語言進行生態資訊管理。下午則由東亞各會員網介紹其長期生態研究現況及其資訊管理，包含了韓國、蒙古、馬來西亞、泰國與台灣等會員網的介紹，另由國際長期生態研究網主席金恆鑣所長介紹國際長期生態研究網的現況與願景，最後由國際強生態研究網資訊管理委員會主席Kristin Vanderbilt博士主持，進行生態資訊管理廣泛的討論，藉由這個討論，來自各國的科學家與生態資訊管理員可以進行經驗的分享與意見交流，並對於未來資訊分享與方式有更進一步的共識。

10月18日大會將會議場所移至商學院的電腦教室，針對生態資訊管理之技術與相關議題做一完整的介紹，包含了生態資訊學、EML與Morpho(生態後設資料語言編輯系統)、後設資料主題。下午藉由所有與會者各自建立的研究資料工作流程，分享了研究主題內容與研究方法，並透過美國設計科學工作流程與應用系統(Kepler Workflow System)的Dr. Deana Pennington的指導，修正了各個工作流程，讓與會者更清楚的了解科學工作流程的觀念。接著進行分組研討合作，各組成員需不同國籍且須包含科學家與生態資訊管理員，以實際的研究進行

工作流程的細部討論與規劃，由科學家提供實際的研究資訊與資料內容，再與資管員進行資料管理與分析流程的討論，進而建構一實際可行的科學工作流程。這樣的工作分組方式，不僅可以達到國際交流的目的，更可以讓科學家與資管員進行面對面的有效溝通，縮短資料分析應用的時程，同時提升資料管理效率與資料分享的可能，及再利用的價值。美國的John Porter博士、Kristin Vanderbilt博士，與陳建文先生，更就如何利用Kepler Workflow發展物件辨識分析流程進行討論。該程序計畫針對生態監測數位影像，進行影像內移動物件的分析及計數，未來可用於捕捉生物的活動事件及活動頻率。

10月19日，由美國與台灣的資管員協助其他與會人員以分組的方式進行實作與研討。生態資訊系統的建置小組，由林朝欽博士與邱振財先生指導馬來西亞的資訊管理員建置Metacat系統；科學工作流程實做小組由Deana Pennington博士與陳建文先生協助來自日本與韓國的資管員與科學家，利用Kepler Workflow System，以實際的研究資料建立可操作的科學工作流程，由John Porter博士、Kristin Vanderbilt博士及鄭美如小姐協助大陸、蒙古與日本的與會者建置研究後設資料建置，再利用本所生態資訊研究小組所開發的資料檢核與分析系統進行資料檢核與基本資料分析，最後以Kepler Workflow System完成研究所需的資料分析與得出研究結果。透過分組實作與討論的方式進行，不僅可以讓與會者更爲快速的了解生態資訊管理的意義與實際操作流程，更能因過程的溝通與問題的解決，讓彼此更加熟悉，建立長期的交流與合作關係，是非常有效且附加價值高的研討方式。

與會人員於研討會場前廣場合影，照片中人員包括本所邱振財先生、中國科學院李峰博士、韓國金恩植博士、本所金恆鑣所長、林朝欽博士、鄭美如小姐，以及日本Eiichi Maida先生

由本所林朝欽博士與John Porter 博士共同主持『International Workshop on Ecological Information Management』一節

來自各國的生態科學家與資訊管理員齊聚一堂吸收新知

由美國與台灣的資訊管理員引導其他國家的科學家
建立後設資料及利用科學工作流程進行資料分析

海岸生態系長期生態研究站參訪(一)

海岸生態系長期生態研究站參訪野外參訪(二)

森林生態系野外參訪

四、建議

1. 建立一個符合區域性多語系之需求的實際可運作的亞太區域的生態資訊管理系統，讓資訊交流平面化，讓具跨國界與跨學門（如結合 IT）的特性的生態研究合作更有效率。
2. 訓練種子教師以協助台灣各學術及研究單位建置完善的生態資訊管理系統，讓研究從前端的資料蒐集到最後的資料分析以最少的人力應用資訊技術提高效率，進而迅速整合台灣之生態資訊庫。
3. 建構國際生態資訊管理人才合作制度，除定期派員前往歐美等先進國家進行新興生態資訊學之研習外，亦需同時派遣種子教師至亞洲其他國家協助建立生態資訊管理系統與培訓管理人員，以早日實踐全球生態資訊管理平台的願景。此外，邀請國際學者專家來訪，也是一項增進我方與國際學術交流的有效策略。

五、附錄

(一)議程表

The 3rd International Workshop on ILTER Ecological Information Management in the East Asia-Pacific Region

- Promotion of ILTER Network Activities on Ecological Information Management -

16-21 October 2007
Institute of Forest Science
Kookmin University
Seoul 136-702, Korea

Organized by the Institute of Forest Science, Kookmin University

Sponsored by Environmental Management Corporation, Korea Forest Service, Korea National LTER Research Center, Korea Science and Engineering Foundation, Korean Association of Biological Sciences, Ministry of Environment Republic of Korea, and National Assembly Climate Change Forum (in alphabetic order)

Under the auspices of the Ecological Society of Korea, Korea LTER Committee, and Research Center for Ecological Design and Management of Urban Forests

PROGRAM

PART I: International Conference on LTER, Urban Forests, and Urban Ecosystems (114 Building Number 7, Kookmin University)

DAY 1: Tuesday 16 October 2007

08:30 - 09:30 Registration

09:30 - 10:00 Introduction and Inaugural Speeches (Chair: Eun-Shik Kim, Kookmin University)

- Eun-Shik KIM (Director, Institute of Forest Science, Kookmin University)
- Jae Chun CHOE (President of the Ecological Society of Korea)

- Byung-Sun IHM (President of Mokpo National University)
- Woo Hyeok BYUN (Korea University)
- Jong-Geel JE (Member, National Assembly, Republic of Korea)
- Man-Yong SHIN (Dean, College of Forest Sciences, Kookmin University)

10:00 - 10:20 Break and Photo Session

10:20 - 12:30 Session 1: Climate Change, LTER, and Urban Ecosystems (Chair: Yow Han Son, Korea University)

- John PORTER, Baltimore Ecosystem Study (BES) Urban Long-Term Ecological Research (University of Virginia)
- Kristin VANDERBILT, Urban Ecology at the Central Arizona-Phoenix Long Term Ecological Research Site (University of New Mexico)
- Mira LEE, Overview of Urban Forest Policies in Korea (Korea Forest Service)
- Kwang Bin CHOI, Policies on parks and green spaces of Seoul (Seoul Metropolitan Government)
- Yow Han Son, Development and management of urban forests: A new project funded by the Korea Forest Service (Korea University)

12:30 - 14:30 Lunch

14:30 - 16:40 Session: LTER and Urban Forests in the East Asia-Pacific Region (Chair: Keizo Tabata, Kyoto University)

- Woo Kyun LEE, Assessment of ecosystem vulnerability to climate change with CEVSA model in Korea (Korea University)
- Deana PENNINGTON, Structure and Function of Cross-Disciplinary Collaborations and the Flow of Information (US LTER Network, University of New Mexico)
- Feng LI, Ecological Evaluation, Planning Approach of Urban Green Space and Urban Forest in China (Chinese Academy of Sciences)
- Keizo TABATA, Tree Population Dynamics of Large-scale Mature Urban Forest in Kyoto city, Japan (Kyoto University)
- Hen-biau KING, Environment, Forest Ecosystems and Forest Research in Taiwan (Taiwan Forestry Research Institute)

16:40 - 17:00 Break

17:00 - 17:30 General Discussion on Ecological Management of Urban Forests and Urban Ecosystem (Chair: Hen-biau KING, Taiwan Forest Research Institute)

- Eun-Shik KIM (Kookmin University)
- Woo Kyun LEE (Korea University)
- Feng LI (Chinese Academy of Sciences)
- Keizo Tabata (Kyoto University)

17:30 End of Day 1

PART II: International Conference on ILTER and Ecological Information Management (114 Building Number 7, Kookmin University)

DAY 2: Wednesday 17 October 2007

09:90 - 10:00 Registration

10:00 - 12:00 Session: ILTER and Ecological Information Management (Chair: Deana PENNINGTON, University of New Mexico)

- Deana PENNINGTON, Advances in Cyber-Enabled Science: An Overview (University of New Mexico)
- John PORTER, Ecological Information Management Training in the East-Asia-Pacific Region (University of Virginia)
- Chau Chin LIN, Ecological Information Management Using Ecological Metadata Language (Taiwan Forestry Research Institute)
- Jae Chun CHOE, Korea National Long-term Ecological Research (KNLTER) (Ehwa Womans University)
- Victor AMOROSO, Status and Activities of the Philippine Long-Term Ecological and Biodiversity Research (Central Mindanao University, Musuan, Bukidnon, Philippines)

12:00 - 14:30 Lunch and Poster Session

14:30 - 17:10 Session: LTER and Ecological Information Management in the East Asia-Pacific Region (Chair: John PORTER, University of Virginia)

- Ariuntsetseg, L., LTER Activities and Initiatives of Ecological Data Management in Mongolia (National University of Mongolia)
- Rahim, Omar Ali Abdul, LTER Network Initiative in Malaysia (ICT Unit, Forest Research Institute Malaysia)
- Meei-ru Jeng, The Activities and Information Management System of Taiwan Ecological Research Network (Taiwan Forestry Research Institute)
- Hen-biau KING, Recent Development and Future Perspective of the International Long Term Ecological Research (ILTER) (Taiwan Forestry Research Institute)
- Kristin VANDERBILT, Information Management Challenges and Opportunities in the International LTER Network (University of New Mexico)
- Yongyut Trisurat, Long-Term Ecological Research in Thailand (Kasestsart University)

17:10 - 17:30 Break

17:30 - 18:00 Promotion of ILTER Information Management and General Discussion (Chair: Kristin VANDERBILT, University of New Mexico)

- Deana PENNINGTON
- John PORTER
- Chau Chin LIN
- Hen-biau KING

18:00 End of Day 2

PART III: International Workshop on Ecological Information Management (505 College of Economy and Business, Kookmin University)

DAY 3: Thursday 18 October 2007

08:30 - 09:00 Introduction to Eco-informatics (John PORTER)

09:00 - 09:30 Introduction to EML and MORPHO (Kristin VANDERBILY)

09:30 - 10:00 Metadata driven tools (1) (Chau Chin LIN)

10:00 - 10:15 Coffee break

10:15 - 11:15 Metadata driven tools (2) (Deana PENNINGTON)

11:15 - 12:00 Review of Morning Session and Discussion

12:00 - 13:30 Lunch

13:30 - 17:00 Collaborative Conceptual Modeling (Deana PENNINGTON)

18:00 End of Day 3

DAY 4: Friday 19 October 2007

08:30 - 12:00 Group Work

12:00 - 13:30 Lunch

13:30 - 15:00 Review of the Workshop and Discussion

15:00 End of Workshop

PART IV: Field Tour (optional)

15:00 Departure for Field Visit (Hampyeong Bay and Evergreen Forests in South Western Part of Korea)

DAYS 5-6: Saturday 20 – Sunday 21 October 2007

- Sun-Kee HONG, Long-term Ecological Research of Coastal Ecosystem in Korea (Mokpo National University, Korea)
- Byung Sun IHM (Mokpo National University)

(二)報告內容

第一篇 全文

Ecological Information Management Training in the East-Asia-Pacific Region

John Porter, Chau-Chin Lin, Meei-ru Jeng, and Sheng Shan Lu

ABSTRACT

Information management is a necessary component of long-term research. Without long-term data there can be no long-term research, and without active management there is no long-term data. Ecological information management is a rapidly evolving field that has as its aim the management, archiving, distribution and integration of ecological data. The East-Asia Pacific (EAP) region of the International Long-Term Ecological Research Network (ILTER) has recently engaged in a series of workshops focusing on augmenting the management of ecological data in the EAP region with the aim of developing systems and infrastructure to support regional analyses. Here we outline the activities undertaken since 2005 to achieve this aim and discuss future opportunities.

INTRODUCTION

Successful management of information is crucial to the success of long-term research. Without long-term data, there is no long-term research, and without active management and archiving, there is no long-term data! The past 15 years have seen dramatic changes in the way ecological information is managed and shared. The amount of ecological data that is shared has rapidly increased, with one data-discovery search engine (http://knb.ecoinformatics.org) listing more than 15,000 different ecological data sets in 2007.

The East-Asia Pacific (EAP) Region of the International Long-Term Ecological Research (ILTER) Network consists of groups that operate Long-Term Ecological Research (LTER) sites. Formal members of the EAP include Australia, China, Mongolia and Taiwan, but candidate networks are forming in

Japan, Malaysia, Philippines, Thailand and Vietnam. The goal of the network is to help improve interactions between individual LTER groups with the aim of producing improved ecological understanding at regional and global scales.

A 2005 meeting of the EAP group identified three goals for information management:

- Identify regional requirements for information management systems
- Build member networks' capacity to manage information to support data sharing and synthesis
- Initiate regional science projects that build on the information management infrastructure

There were several challenges that faced the EAP in achieving these goals. First, there were large differences in the resources available to support Long-Term Ecological Research among the regional LTER groups, and those difference extended to the resources available for information management as well. Second, ecoinformatics (the management of ecological information) is a relatively new field with few opportunities for formal academic training. Finally, it is a rapidly evolving field that has seen major breakthroughs in standards, tools and technologies the last 5 years. To overcome these challenges and achieve the goals, the EAP identified a series of training workshops that would lead to a better-integrated ecoinformatics framework for the region that could serve as a model for ILTER information management in other regions.

BEIJING 2005

The first of a series of workshops was held in Beijing during July 2005. With attendees representing EAP LTER groups from Australia, China, Indonesia, Japan, Korea, Malaysia, Philippines, Taiwan, Thailand, and Vietnam. Additional participants from outside the EAP region came from Mexico, South Africa and United States. The workshop had several focus areas. The first was to share information on status of Information Management (IM) efforts and infrastructure in different EAP ILTER groups. This provided a context for a series of presentations that provided an overview of methodologies for managing ecological information.

These presentations included more detailed explanations of the role of ecological metadata in data discovery, access and use. Metadata or "data

about data" provides a way to describe the diverse forms ecological data can take, the scientific context in which measurements were taken, and provides the content for searching for needed data. Ecological Metadata Language (EML) provides a structured way of representing ecological metadata so that many of the steps required to utilize data can be automated.

Following the workshop the first EAP Information Managers Committee meeting was held. It decided on a strategy for achieving EAP goals that involved a series of workshops hosted by different EAP members.

GUEISHAN TRAINING CENTER, TAIWAN 2006

The first follow-up workshop was held at the Gueishan Training Center in Taiwan during February 2006. The workshop focused on developing information management strategies and training on the tools needed to support them (Lin et al. 2006). It began with overviews of different systems developed by EAP participants and in the U.S. LTER Network. These were followed by instructional sessions on how to use relational database management systems (RDBMS) as a component of ecological information management. There was also a discussion of ways to archive data to assure its long-term preservation.

The primary focus of this workshop was on Ecological Metadata Language (EML), specifically how to create EML documents. A wide variety of approaches for creating EML were demonstrated and tested by the participants. These included text editors, eXtensible Markup Language (XML) editors, and spreadsheets coupled to post-processing programs. However, the principal focus went to the "Morpho" editor, an XML editor specifically designed for creating EML documents. Workshop participants each used Morpho to create one or more complete metadata documents for data from their LTER program.

KAOHSIUNG TAIWAN 2007

A second workshop was held in Kaohsiung in Taiwan during July 2007. Whereas the previous workshop focused on the creation of EML documents, this workshop focused on the use of EML metadata and their associated data. Workshop participants saw demonstrations of tools that automatically read

EML metadata and then produce quality assurance analyses using the “R” statistical programming language (Lin et al. 2007). Using a web interface automatic checks confirm that the columns of data are of the right type (e.g., numbers vs. character strings) and that they fall within the specified numerical range for each variable. There were also presentations on advanced techniques for improving data quality. However, the principal focus of the workshop was on how to use the KEPLER program.

KEPLER is a tool for developing scientific workflows through the use of a graphical user interface. Researchers can connect to a variety of data sources, including structured query language (SQL) databases and EML-documents and their related data. These data sources can then be connected to a variety of components capable of doing a wide variety of analyses from simple statistical summaries to advanced genetic algorithms. These components can then be connected to additional components or display tools to provide a full analysis. Videoconferences with software developers and other experts helped resolve any questions the participants had about how to use the software. The workshop concluded with presentations by the participants of a variety of work flows that they had developed using KEPLER. This workshop helped set the stage for a subsequent workshop to be held in Korea in October 2007.

OTHER ACTIVITIES

In addition to the workshops, there were a number of formal and informal activities that provided training and facilitated communications among workshop participants. These activities included exchanges of personnel across EAP members for “hands-on” training. For example, an information manager from Japan spent several weeks working with the Taiwan Ecological Research Network, and five information managers from Taiwan each spent several months working at U.S. LTER sites (Porter 2007). There were also additional meetings of the EAP-ILTER Network that made adjustments to strategies and continued to define and modify objectives for EAP information management activities.

Many of the EAP information managers also attended the LTER All-Scientists’ Meeting at Estes Park, Colorado, USA during September 2006. There they participated in the annual U.S. LTER information management committee meeting and in numerous other sessions at the conference. A special session

focused specifically on International LTER Information Management, including representatives from other ILTER regions including Europe, Africa and South America.

ACHIEVEMENTS

The workshops have resulted in a number of achievements. Most important is the greatly improved communication regarding issues confronting long-term ecological science and ecological information management within the EAP region. The developing community of ecological information management experts is increasingly sharing expertise and resources, making the tasks of each individual group easier. It has also led to the development of new tools and approaches to ecological information management, that again have been shared. One example of this is a tool developed by the Taiwan Ecological Research Network that extracts geographic information about data collection locations and automatically generates an online web map showing the locations. This tool has been incorporated into ecological information sharing sites in the U.S. and elsewhere.

NEXT STEPS

The workshop in Seoul Korea in October 2007 will take the first steps in trying to use the data, tools and approaches to look at regional ecological issues. It will give information managers and ecological researchers the chance to use the metadata, specialized software and analytical tools they learned about at the previous workshops in a real-world setting. From that experience we expect to be able to better evaluate where additional training is needed and if there are adjustments in approach that need to be made. These needs can then be addressed in additional meetings and workshops.

LITERATURE CITED

Lin, C. C., J. H. Porter, and C. W. Hsiao. 2007. Automating analysis of sensor data using Ecological Metadata Language (EML). Pages 5 *in* Conference on Coastal Environmental Sensor Networks. Center for Coastal Environmental Sensing Networks, University of Massachusetts, Boston, Boston, MA.

Lin, C.-C., J. H. Porter, and S.-S. Lu. 2006. A Metadata-Based Framework for Multilingual Ecological Information Management. Taiwan Journal of Forest Science **21**:1-6.

Porter, J. H. 2007. Information Management Interactions with Taiwan. *in* LTER Network News.

第二篇 摘要

The activities and information management system of Taiwan Ecological Research Network

Chau-Chin Lin, Meei-ru Jeng, John Porter, Teng-Chiu Lin, and Sheng-Shan Lu

The activities of the Taiwan Ecological Research Network (TERN) followed the original goals of 1990s to create the network and specific topics of each site. Recently, based on the past data collected, TERN focuses on cross-regional site comparisons including typhoon/hurricane and ecosystem dynamics, global weathering and climate change, ecology and management of forest plantation, and regional flux research topics. In addition to research activities, TERN has put efforts on ecological information management since 2004. Based on the progress of these efforts, we have hosted sever international ecological information management workshops to share the experiences with member networks of EAP-ILTER. In the near future, TERN will continuously host some science conferences.

In terms of ecological information management system of TERN, we have collaborated with US LTER and adapted Ecological Metadata Language (EML) as the standard to develop our information management system. This system permits sharing of functionality across multiple systems. TERN IM team also devoted to develop a data checking tool based on EML documents stored on our metadata database. This tool allows researchers to access EML, upload data, check data quality, and then run R code on the server. Furthermore, the system links to scientific workflow system called Kepler as the automatically integrating tool for researchers who focus on ecological modeling.

Keywords: Ecological Metadata Language, Kepler, Ecology, Taiwan

(三)壁報展示摘要

Poster 1

Automating Analysis of Sensor Data Using Ecological Metadata Language

Chau-Chin Lin,[1)] John H Porter,[2)] Chi-Wen Hsiao,[1)] Sheng-Shan Lu,[1)] Meei-Ru Jeng[1,3)]

【Abstract】

Most of the traditional methods used by ecologists to collect field data are time and manpower intensive. The development of senor networks, however, allows sensors to gather data in the field and deliver them to the laboratory automatically. But the shift toward this data collection paradigm in turn creates new challenges for indexing, navigating, documenting and analyzing data. Furthermore, standard mechanisms for analysis and quality assurance checking of senor collected data are lacking. Therefore, an automated data quality checking and analyzing system based on metadata has tremendous advantages for ecologists when they use sensor collected data. To address the management and quality assurance problem of sensor data, we have developed a tool based on the Ecological Metadata Language (EML). This tool allows researchers to access EML, upload data, check data quality, and then run R code on the server. Researchers can use it for data manipulation, calculation, and graphical display online without the need to install any statistical software.

---

1) Taiwan Forestry Research Institute. 53 Nan-Hai Rd., Taipei 10066, Taiwan.

2) Department of Environmental Sciences, University of Virginia. 291 McCormick Road, Charlottesville, VA 22904-4123, USA.

3) Corresponding author, e-mail:beer@tfri.gov.tw

Poster 2

Metadata-driven Dataset Analysis Using Scientific Workflow for Real-time Meteorological Data in Botanical Garden

Chien-Wen Chen
Division of Forest Biology, Taiwan Forestry Research Institute

Implementation of wireless sensor network in Taipei Botanical Garden(TPBG) offers real-time biological and meteorological data. With utilization of 802.11 wireless protocol, information from wireless sensor network of TPBG can be investigated anytime and anyplace where Internet is available. Meteorological station logs air temperature, relative humidity, atmospheric pressure, wind speed, and wind direction. Data is received as electrical file, saved in the field workstation temporally, then transported to FTP server located in research building of TFRI. Wi-Fi connection is built using directional antenna set on the top of tower to communicate with the local area network of Taiwan Forestry Research Institute(TFRI). Another dataset file is duplicated and deposited in Storage Resource Broker(SRB) server simultaneously for duplication backup.

Dataset transported from TPBG into server can be the data source of scientific workflow. With the information from metadata document using the format of EML(Ecological Metadata Language), deposited real-time meteorological dataset is conducted to the scientific workflow software, Kepler. After finishing data ingestion process, Kepler will parse data attributes, process basic statistics, then produce graphs for the results of data analysis. Scientific workflow help researchers to process repeating scientific data acquisition, calculation, analysis, and presentation efficiently. Graphic user interface of Kepler will be advantageous to researchers for composing scientific workflow. Meteorological data can be the reference data for processing the cultivation of plants in the garden.

(四)壁報展示

Automating Analysis of Sensor Data Using Ecological Metadata Language

Chau-Chin Lin,[1] John H Porter,[2] Chi-Wen Hsiao,[1] Sheng-Shan Lu,[1] Meei-Ru Jeng[1]

Introduction

Traditionally, ecologists worldwide, use a variety of methods to collect field data. Most of these methods are time and manpower intensive. The development of senor networks, however, allows sensors to gather data in the field and deliver them to the laboratory automatically. Already a variety of ecological important data such generic meteorological measurements, soil and water temperatures and acoustical records are being collected by senor networks.

But the shift toward this data collection paradigm in turn creates new challenges for data management problems which include documenting, quality checking, navigating, and analyzing data.

Framework of system

To address the management and quality assurance problem of sensor data, we have developed a tool based on the Ecological Metadata Language (EML). This tool allows researchers to automate basic quality assurance processes and correct the mistakes checked from the system.

The system is building on the EML2R from the Processing Techniques for Automated Harmonization (PTAH) project, a web-based interface for creating "R" programs was developed. This interface not only extends the capabilities of the transformation but also has become a prototype of a server-side system that allows researchers to access EML, upload data, and then check data type, data range, and missing data definition based on the data table defined in EML document.

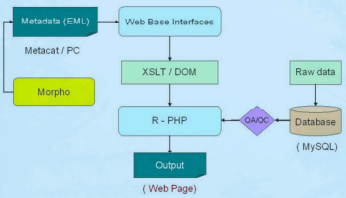


Fig. 1. The framework of automating analysis system

After the checking, our analysis tool provides the function to let users correct the errors online and save back to the system automatically without accessing their original EML documents from Morpho. Researchers can continue to run "R" code on the server. There are two modules in our system for users to run "R" code, one is allowing user to add "R" code for custom analysis and the other is running statistical analysis by a graphic user interface. Since "R" provides a wide variety of statistical and graphical techniques and is highly extensible, researchers can use it for data quality control, data manipulation, calculation, and graphical display online without the need to have their own copy of "R" locally.

The framework of automating analysis system, as indicated in Figure 2, provides researchers the whole process to document, check, and analyze ecological data.

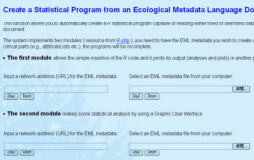


Fig. 2. The interface of the automating analysis system (http://metacat.tfri.gov.tw:8180/modules/index.php)

The current interface of our automating system has been used in TFRI, Taiwan Ecological Research Network (TERN), and Chinese Ecological Research Network (CERN) in various research topics, including forestry, ecology, hydrology, and soil science.

The automating analysis process is based on EML attached data table. Each data table should include entity name, and attributes. According to EML schema, each attribute should be the name of a field in a table (Fig. 3). Attributes are described by attribute definition, measurement scale. Each field should define one of 5 categories of measure scale including nominal, ordinal, interval, ratio, and time and the detail definition of the field. Based on these definitions, datasets can be checked automatically by reading EML document.

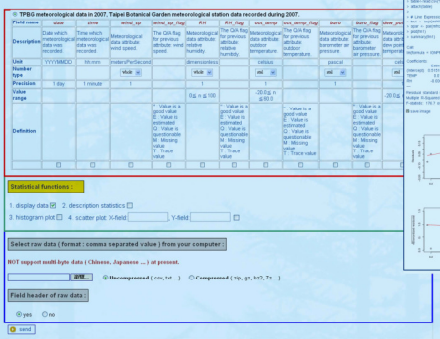


Fig. 3. Showing the metadata, selecting statistical function and uploading raw data

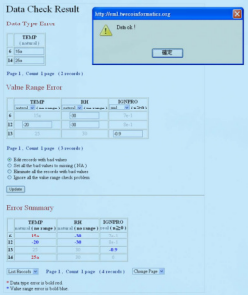


Fig. 4. Example of one dataset automating check result.

Figure 4 shows the result of a dataset checking. After dataset checking of type and range, researchers can go ahead to do some general statistical analysis or advance coding analysis based on R statistical functions (Fig. 5~7).

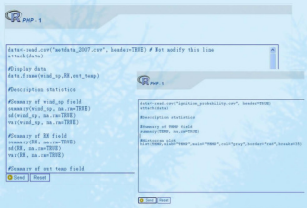


Fig. 5. Allow user editing R code

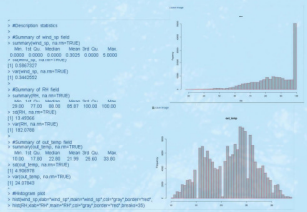


Fig. 6. Example of a simple descriptive statistical analysis and graph based on R project.

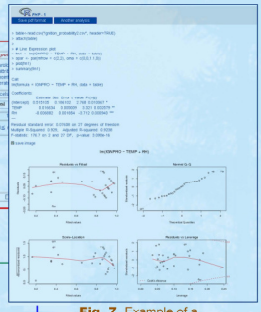


Fig. 7. Example of a linear regression analysis and graph based on R project.

Conclusions

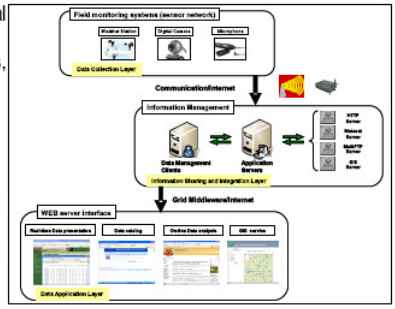
We have described the framework of our automating analysis system that enables data quality assurance and basic statistical analysis. The system can be used in any fields data which with EML documents. Our current work focuses on EML document attached dataset. We would like to move to automate analysis datasets that store on general database management system in the future.

1) Taiwan Forestry Research Institute. 53 Nan-Hai Rd., Taipei 10066, Taiwan.
2) Department of Environmental Sciences, University of Virginia. 291 McCormick Road, Charlottesville, VA 22904-4123, USA.

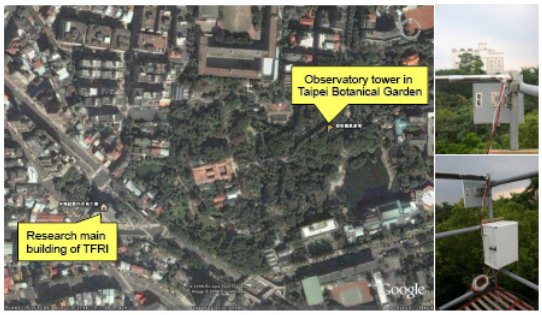
Metadata-driven Dataset Analysis Using Scientific Workflow for Real-time Meteorological Data in Botanical Garden

Chien-Wen Chen
Division of Forest Biology, Taiwan Forestry Research Institute

Wireless sensor network in ecological studies has advantages of minimizing the risk of data loss, real-time information acquisition. Taiwan Forestry Research Institute (TFRI) start to utilize wireless sensor network to collect ecological informations such as images, sounds, and environmental data. 3-tier ecological information services management was conceptualized out.



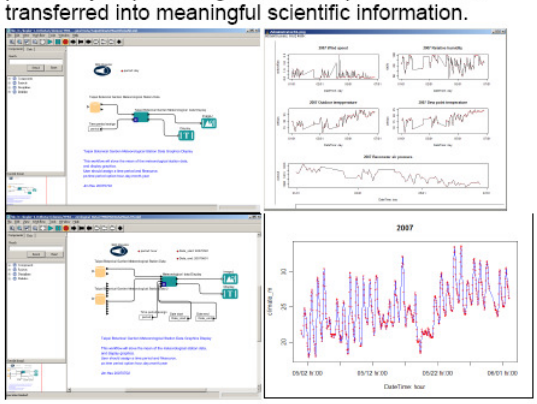
Wireless sensor network in Taipei Botanical Garden (TPBG) follows 802.11 network protocol. Observatory tower located in TPBG is ca. 263 m away from the main building of TFRI. By using Wi-Fi to connect toward Internet, it is not necessary to layout cables crossing the tree canopy. Directional panel antenna is set on the top of tower to transmit wireless signal toward TFRI's LAN.



Meteorological station is installed at the bottom of the tower. Measurements include: air temperature, related humidity, atmospheric pressure, wind speed, and wind direction.

Meteorological data is transfered back into FTP server located in TFRI continuously and automatically. Another backup dataset is deposited in SRB server simultaneously. With the incorporation of the Ecological Metadata Language (EML) metadata description, stored data can be retrieved, interpreted, rearranged, then used by other analytical softwares for further analyses and researches.

Data package of TPBG meteorological dataset is deposited in TFRI Ecological Informatics Working Group's Metacat server. On-line location of rawdata file is assigned in EML document as a FTP link. Attribute descriptions and rawdata physical link address can be parsed by Kepler to ingest and interpret data, then transferred into meaningful scientific information.



After finishing parsing, Kepler process basic statistics, then produce graphs for the results of data analysis. Scientific workflow help researchers to process repeating scientific data acquisition, calculation, analysis, and presentation efficiently. Graphic user interface of Kepler will be advantageous to researchers for composing scientific workflow.