Research note

Forest Dynamics Plot Database of Taiwan: A Content Management Approach

Chau-Chin Lin,[1,2] Yu-Hwang Wang,[1] Sheng-Shan Lu[1]

【Summary】

This study adopted the content management system, Drupal Ecological Information Management System, to develop a database and a website to host 16 forest dynamics plots of Taiwan. These plots established since 1990 are from 1 to 25 ha in size. We edited the detailed metadata of the 16 plots based on Ecological Metadata Language to link all related information including the dataset, project, researcher, publications, and plot profile. The linked information is presented in a webpage format. Users can easily browse the information through designed pages such as 116 publications which have been collected with bibliographical pages. In order to use the dataset of plots, we also created a search interface for users to query data with survey tags or species names. The results showed that the Drupal Ecological Information Management System is a potential tool for forest dynamics plots data management.

Key words: metadata, EML, data Management, LTER.

Lin CC, Wang YH, Lu SS. 2015. Forest dynamics plot database of Taiwan: a content management approach. Taiwan J For Sci 30(4):281-8.

[1] Forest Protection Division, Taiwan Forestry Research Institute, 53 Nanhai Rd., Taipei 10066, Taiwan. 林業試驗所森林保護組，10066台北市南海路53號。

[2] Corresponding author, e-mail:chin@tfri.gov.tw 通訊作者。

Received July 2015, Accepted August 2015. 2015年7月送審 2015年8月通過。

研究簡報

台灣森林動態樣區資料庫：內容管理系統之應用

林朝欽[1,2] 王豫煌[1] 陸聲山[1]

摘 要

本研究取用生態內容管理系統發展資料庫及網站作為台灣16個森林動態樣區之資訊管理，此16個森林動態樣區是自1990年以來陸續建立的，樣區面積由1~25 ha不等。我們首先依生態後設資料語言編輯16個樣區的詳細描述包括：樣區數據集、計畫內容、研究人員、出版文獻、樣區資訊，並將這些資訊串接後以網頁型態呈現。使用者可以透過網頁瀏覽器查詢所需的資訊，例如我們所蒐集到的116篇研究報告及其所關連的所有資訊。除此，為提供使用者查詢數據集內容，我們也建立一個搜尋界面可以快速查找樣區調查樹種、量測值。本研究結果顯示生態內容管理系統是一個有潛力的工具用以管理森林動態樣區資訊。

關鍵詞：後設資料、生態後設資料語言、數據管理、長期生態研究。

林朝欽、王豫煌、陸聲山。2015。台灣森林動態樣區資料庫：內容管理系統之應用。台灣林業科學30(4):281-8。

The study of forest dynamic plots began in the 1970s. The first 50-ha forest dynamics plot was established on Barro Colorado Island (BCI) in Panama by the Smithsonian Tropical Research Institute in the 1980s (Ashton et al. 1999). After more than 3 decades of development, currently there are over 60 plots around the world networked by The Center for Tropical Forest Science - Forest Global Earth Observatories (CTFS-ForestGEO) (http://www.forestgeo.si.edu/). It is a global network of forest research dedicated to the study of tropical and temperate forest function and diversity across the Americas, Africa, Asia, and Europe, including 3 sites in Taiwan. The CTFS-ForestGEO monitor the growth and survival of approximately 6 million trees and 10,000 species.

In Taiwan, there are 16 plots that range in size from 1 to 25 ha which have been successively established since 1990 (Yang et al. 2008). A great deal of plot data has accumulated throughout censuses at each site. The wealth of these datasets presents challenges of information management for researchers. For example, to fulfill the database potential, the datasets need to meet several requirements. First, the data need to be stored in a way suitable for long term survival. Second, the data need be readily accessible. Third, the data need to be supported by clear descriptions of the context in which the study was undertaken (Magnuson 1990, Le Due et al. 2007, McIntosh et al. 2007). A database system designed by the CTFS attempts to meet those requirements. However, the system is not an open-source system. It also lacks management of all components related to the plot project such as researchers, publications and contact information (Lin et al. 2009). In order to better document the information related to the 16 plots in Taiwan, we attempted to adopt an alternative way to manage the plot information.

In the 1990s, webpages were nothing more than simple text files nestled in folders on a server somewhere on the internet. A website in those days was a collection of files in a particular folder (Byron et al. 2012). When websites grew in size, it was obvious that this approach did not scale well. The search for solutions to this problem went to software focused on a particular task or application using scripts and Common Gateway Interface (CGI) programs. This approach first used special tags in each HyperText Markup Language (HTML) file. Second, the new approach changed to using databases to store pieces of similar content. The approach slowly and surely emerged to manage these different kinds of content and features using a single, consistent user interface which is called a Content Management System (CMS) (Byron et al. 2012). The CMS changed software focused on a particular task or application. Instead of creating and saving the same information continually throughout a website, a CMS allows web developers to write it once and have the website work it out.

Major benefits of the CMS were highlighted by Gupta et al. (2001): enriched information sharing and collaboration; improved data security; standardization and lower web publishing costs; and 're-usability' of the content for multiple media. Drupal is one of the next-generation CMSs (Byron et al. 2012). It allows one to create and organize many kinds of content, provides user management tools for both the maintainers of and the visitors to a site, and gives one access to thousands of third-party plug-ins that appear as new features. Drupal's modular architecture and open-source nature has made it a popular Hypertext Preprocessor (PHP) application framework and content management system for hundreds of thousands of web developers around the world (James and Noble 2012).

Developed in partnership among the US Long Term Ecological Research (LTER) Network, Univ. of New Mexico, Univ. of Puerto Rico, Univ. of Wisconsin, and Palantir.net, the Drupal Ecological Information Management System (DEIMS) is an installation profile for storing, editing, and sharing data and information about biological and ecological research (San Gil et al. 2010). The DEIMS provides user-friendly forms to describe all contextual information about one's data. The DEIMS automatically produces Ecological Metadata Language (EML) to share research records with other networks and metadata clearinghouses such as the Knowledge Network for Biocomplexity (KNB). Figure 1 shows the architecture of the DEIMS with EML templates embedded. The DEIMS was adopted for the Luquillo Forest Dynamics Plot (LFDP) of US LTER Luquillo (San Gil et al. 2010). This shows that the DEIMS is a suitable system for use in Taiwan.

The first version of the DEIMS has been deployed in 5 US LTER sites since 2010. Currently the DEIMS in its second version is free for all and available at https://www.drupal.org/project/deims. It is an attractive option to the International LTER (ILTER) Network because Drupal automatically supports content and user interface translation. LTER Europe has adapted it for managing site-related information for its 20 member countries (Vanderbilt et al. 2015).

Based on a survey of categories used to classify and serve information through 26 US LTER websites, the most often recurring categories of information include: scientific data, publications, research projects, personnel directories, sites, facilities, maps, outreach activities and informatics which are the design sources of the custom content type of the DEIMS (San Gil et al. 2010). Using survey results, 6 custom content types have been de-

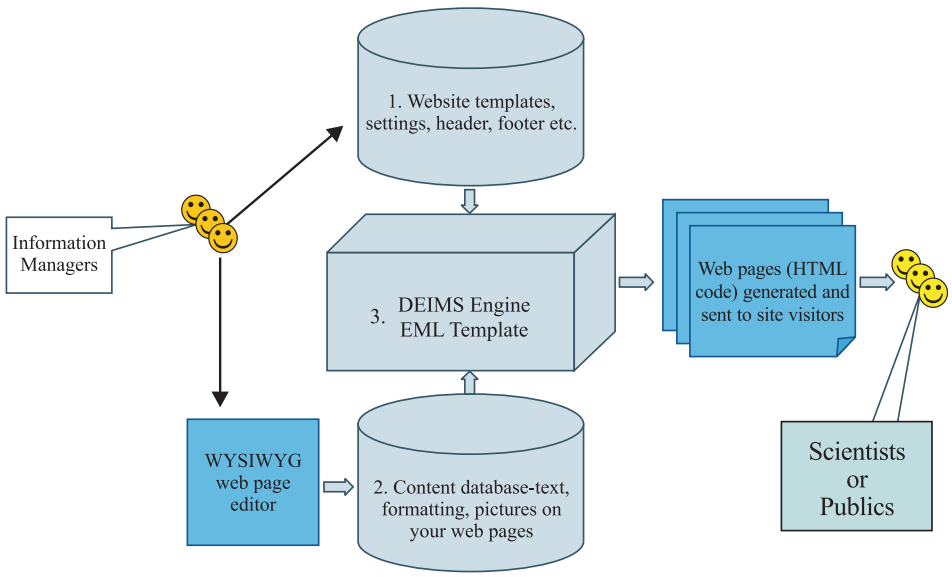

Fig. 1. The Drupal Ecological Information Management System (DEIMS) architecture which adds Ecological Metadata Language (EML) template to the profiles to allow users to create EML documents based on its standard.

signed and defaulted within the DEIMS since the first version. The 6 custom content types were improved and tested in 5 US LTER sites and LTER Europe and were proven to be user friendly (San Gil 2013). We adopted the second version of the DEIMS to develop a database and website to host 16 forest plots of Taiwan. The DEIMS expands Drupal for instance with 6 custom content types for managing ecological research data including:

1. person - allows the creation of a directory - with contact info for researchers, etc;
2. dataSet - describs the basic information about a dataset (title, abstract, etc.);
3. dataFileStructure - gives details about a data-containing spreadsheet or rendered view (details of dataFileStructure are shown in Table 1);
4. researchSite - gives details about a location where research is conducted (latitude, etc);

Table 1. Detailed content of a dataset (data file structure) in each Ecological Metadata Language (EML) document

Category	Items	Sub-items
System Information	Title, Dataset ID, Short Name, Abstract of dataset	
Dataset	Basic information	
	Data Source	Variables
		Name, Label, Definition, Type
	Personnel	
	Method	
	Temporal	
	Publication related	
Physical Path	File system, URL	

5. variable - gives details about a variable used in the research (a date, a code-set, etc.); and
6. researchProject - gives details about an umbrella research project, if any.

Based on the above 6 types of content, we then interviewed the 16 plots' principle investigators (PIs) to consult about the project that they conducted in the past including research team members to collect any information that they were willing to provide.

The result of this adaptation is shown in Figure 2 and 3 which are the homepage of these 16 plots with all 6 collaborative agencies and details of the plot information including the major researchers in different regions of Taiwan. We also edited all 16

Fig. 2. Homepage of 16 plots with all 6 collaborative agencies.

Fig. 3. Information of each plot (using the north as an example).

plots' EML documents which describe all the detailed information of the project including the abstract, associated research project, data file structure (Table 1), personnel involved (owner of the data, contact, data manager, and associated researchers), geotemproal information, and methodology. General users can access all these metadata without raw data. In addition, 116 publications from 1990 to 2013 were collected and edited with the title, publication type, year of publication, journal name, volume and paging, and keywords to allow searching by suers (Fig. 4). In order to use datasets of the plots, we also created a search interface with tags or Chinese species name, or by year of the census, or family of the species. The search not only searches 1 plot but also crosses all 16 plots. Totally, there are 960,000 estimated records for the 16 plots. Since the data volume is large, we improved the DEIMS by implementing the search engine, Apache Solr, that enhances data discoverability (Fig. 5).

Results of testing of the DEIMS showed that the DEIMS is a potential tool that allows forest dynamics plot research to document and distribute data using a common model. The promising grassroots innovation and leveraging of existing open-source software can benefit information management solutions like forest dynamics plot research. Therefore, we concluded that the DEIMS can be used as an optimal system to manage quality forest dynamics plot information. In addition, integrating Solr, an open-source search engine, into the DEIMS augments big data searches.

ACKNOWLEDGEMENTS

The authors thank all principal investigators for the 16 plots' useful information related to the projects that conducted by Professor Chang-Fu Hsieh at Taiwan University, Prof. Kuoh-Cheng Yang at Providence University, Mr. Jei-Ker Lin of National Chung-Hsing Univ., Prof. Tsung-Hsin Hsieh of National

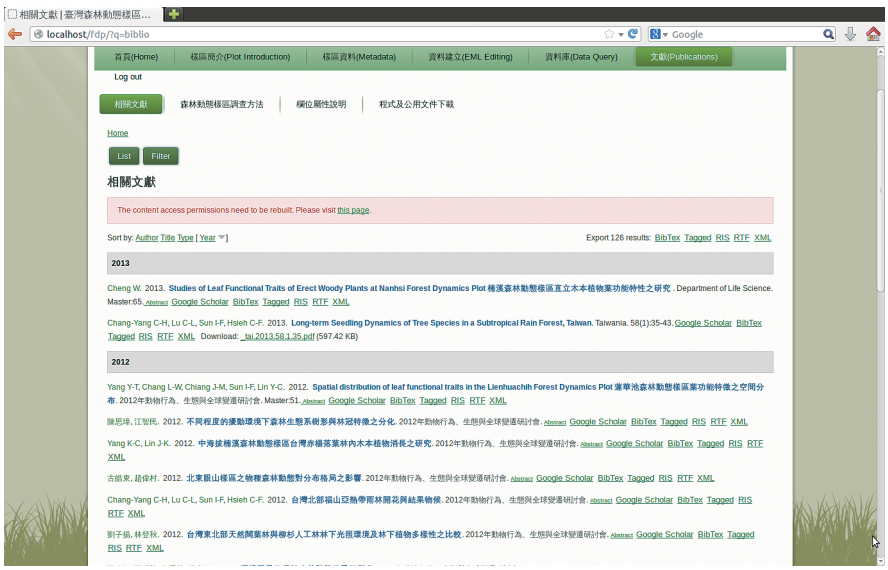

Fig. 4. Publications from 1990 to 2013 of plot data collected in the system which links to all related items such as authors, sites, and original papers that can be shared.

Fig. 5. Enhanced data search function by adding the search engine, SOLR, into the Drupal Ecological Information Management System (DEIMS) to improve the large volume of data archived.

Tainan Univ., and Prof. Wei-Chun Chao of Chia-Yi Univ.. Taiwan Forestry Bureau staff who provided background of plot projects they funded are also appreciated. Finally, we thank colleagues at the Taiwan Forestry Research Institute who provided raw data of 3 plots to test of the search function so the test of the DEIMS was fruitful.

LITERATURE CITED

Ashton PS, Boscolo M, Liu J, LaFrankie JV. 1999. A global programme in interdisciplinary forest research: the CTFS perspective. J Trop For Sci 11:180-204.

Byron A, Berry A, Bondt BD. 2012. Using Drupal. Sebastepol CA, USA. O'Reilly Media Inc. 467 p.

Gupta VK, Govindarajan S, Johnso T. 2001. Overview of content management approaches and strategies. Elec Markets 11(4):281-8.

James T, Noble M. 2012. Drupal 7 business solutions. Burmingham, UK. Packt Publishing. 407 p.

Le Due MG, Yang L, Marrs RH. 2007. A database application for long-term ecological field experiments. J Veg Sci 18(4):509-16.

Lin CC, Hsiao CW, Lu SS, Chiou WL, Chang LW, Jeng MR. 2009. Filling information management gaps of forest dynamics plot databases using Ecological Metadata Language and the scientific workflow system. Taiwan J For Sci 25(1):97-105.

Magnuson JJ. 1990. Long-term ecological research and the invisible present. Bioscience 40(7):495-501.

McIntosh ACS, Cushing JB, Nadkarni NM, Zeman L. 2007. Database design for ecologists: composing core entities with observations. Ecol Info 2(3):224-36.

San Gil I, White M, Melendez-Colom EC, Vanderbilt KL. 2010. Case studies fo ecological integrative information systems: the Luquillo and Sevilleta information managment systems. Metadata and Seman Res 108:18-35.

San Gil I. 2013. The New Drupal Ecological Information Management System (DEIMS). Databits, Spring 2013. Available at http://

databits.lternet.edu/spring-2013/new-drupa-lecological-information-management-system. Accessed 15 06 2015.

Vanderbilt KL, Lin CC, Lu SS, Kassim RA, He H, Guo X, et al. 2015. Fostering ecological data sharing: collaborations in the International Long Term Ecological Research Network. Ecosphere 6(10):art204.

Yang KC, Lin JK, Hsieh CF, Huang CL, Chang YM, Kuan LH, et al. 2008. Vegetation pattern and woody species composition of a broad-leaved forest at the upstream basin of Nantzuhsienhsi in mid-southern Taiwan. Taiwania 53(4):325-37.