

## **INTEGRATION OF BIODIVERSITY DATABASES IN TAIWAN AND LINKAGE TO GLOBAL DATABASES**

***Kwang-Tsao Shao***<sup>1\*</sup>, ***Ching-I Peng***<sup>1</sup>, ***Eric Yen***<sup>2</sup>, ***Kun-Chi Lai***<sup>1</sup>, ***Ming-Chih Wang***<sup>1</sup>,  
***Jack Lin***<sup>1</sup>, ***Han Lee***<sup>1</sup>, ***Yang Alan***<sup>3</sup>, and ***Shin-Yu Chen***<sup>3</sup>

<sup>1</sup>*Research Center for Biodiversity, Academia Sinica, Taiwan, R.O.C. No. 128, Academia Rd Section 2, Nankang 115, Taipei, Taiwan.*

*Email: zoskt@gate.sinica.edu.tw*

<sup>2</sup>*Computing Centre, Academia Sinica, Taipei, Taiwan*

<sup>3</sup>*Institute of Physics, Academia Sinica, Taipei, Taiwan*

### **ABSTRACT**

*The biodiversity databases in Taiwan were dispersed to various institutions and colleges with limited amount of data by 2001. The Natural Resources and Ecology GIS Database sponsored by the Council of Agriculture, which is part of the National Geographic Information System planned by the Ministry of Interior, was the most well established biodiversity database in Taiwan. This database, however, mainly collected the distribution data of terrestrial animals and plants within the Taiwan area. In 2001, GBIF was formed, and Taiwan joined as an Associate Participant, starting the establishment and integration of animal and plant species databases; therefore, TaiBIF was able to co-operate with GBIF. The information of Catalog of Life, specimens, and alien species were integrated by the Darwin core standard. These metadata standards allowed the biodiversity information of Taiwan to connect with global databases. Presently, more than 10 institutes and museums that collect the specimen data and distribution information of Taiwan are integrated, and they can be queried via the Union Catalog of National Digital Archives Program, TaiBIF, and other international nodes of GBIF. In TaiBIF, we have developed a platform and an exchange mechanism, which use the species and GIS distribution as the primary keys. The integrated databases allow users to search for various species web pages by the name and distribution of the species using GIS techniques and display species and specimen basic information and distribution data.*

**Keywords:** Biodiversity, Information Integration, TaiBIF, TaiBNET, GBIF

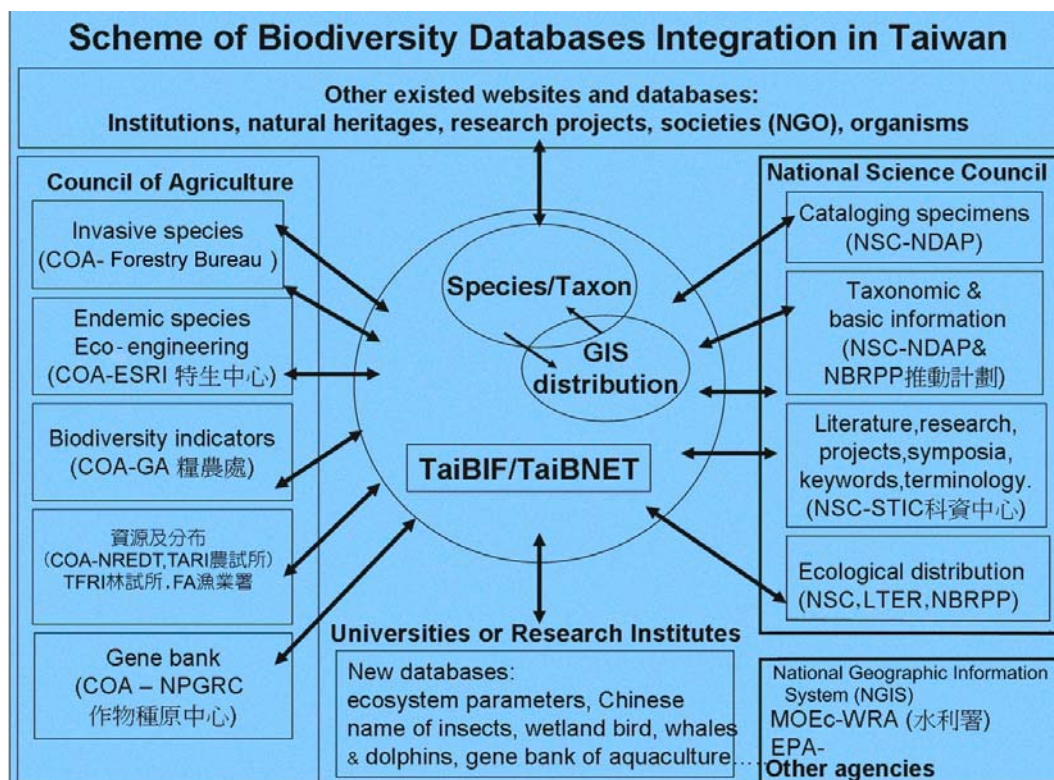
## **1 INTRODUCTION**

Although the land area of Taiwan is not large, it possesses extraordinarily abundant biodiversity resources and many endemic species. To manage and utilize these resources properly, a complete

database and a national website for biodiversity must be built to enable the public to easily retrieve all the relevant information. Furthermore, the establishment of a biodiversity database is fundamental to the promotion of conservation, education, and biodiversity research.

The biodiversity databases in Taiwan were dispersed to various institutions and colleges with limited amount of data by 2001. The information of Catalog of Life, specimens, and alien species were not integrated until a five-year program, National Digital Archives Program, was established during 2002-2006. Meanwhile, a Biodiversity Action Plan was enacted by Executive Yuan, and it is requested by the National Science Council, Academia Sinica, and various institutions to establish and integrate databases under a national portal for biodiversity, Taiwan Biodiversity Information Facility (TaiBIF) with correspondence to the Global Biodiversity Information Facility (GBIF) (Edwards, Lane, & Nielsen, 2000; Bisby, 2000).

The Department of Life Sciences of the National Science Council appointed the Academia Sinica to gather taxonomists to construct the Catalog of Local Experts and the Catalog of Life (excluding alien species) for Taiwan biodiversity from 2002 to 2004. The catalogs were placed under the Taiwan Biodiversity National Information Network (TaiBNET) <http://taibnet.sinica.edu.tw>. Presently, more than 10 institutes and museums that collect the specimen data (animal and plant) and distribution information of Taiwan are integrated using the Darwin Core metadata standard. They can be queried via the Union Catalog of National Digital Archives Program (NDAP), TaiBIF, and TaiBNET. Figure 1 shows how these databases were integrated based on the species names and GIS distribution as the primary key. The integrated databases allow users to search for web pages of various species by the name and distribution of the species by using GIS techniques.



**Figure 1.** The schema used to establish and integrate databases of biodiversity using the species and GIS distribution as the primary key.

## 2 TAIBIF - TAIWAN PORTAL OF GBIF

TaiBIF stands for Taiwan Biodiversity Information Facility, which is in charge of integrating Taiwan's biodiversity information, including the Catalog of Life (a list of Taiwanese native species) and local experts, the illustrations of species, the introduction of native species and invasive species, Taiwan's terrestrial and marine organisms, biodiversity literature, geographical and environmental information, information about relevant institutions, organizations, projects, observation spots and publications, etc. Furthermore, TaiBIF uses DiGIR (Distributed Generic Information Retrieval) with the Darwin Core as its metadata standard to achieve data exchange. Both Chinese and English websites of TaiBIF are online for public use (Figure 2) <http://taibif.org.tw>.



Figure 2. Website of TaiBIF (in Chinese & English)

There are three ways of connecting and sharing data from local databases to global databases/networks (Figure 3). (1) The linkage of a local database to global species databases, such as the Fish Database of Taiwan links to FishBase by taxon name (Shao, Peng, Yen, Lai, Wang, Lin, et al, 2005). (2) The connection through national nodes like TaiBIF to GBIF (Huang, Wu, Lai, Shao, Peng, & Yen, 2004). (3) The linkage from a local database to a regional database such as to Asia-Oceania data in Species 2000 AO (Buckridge & Gordon, 2000).

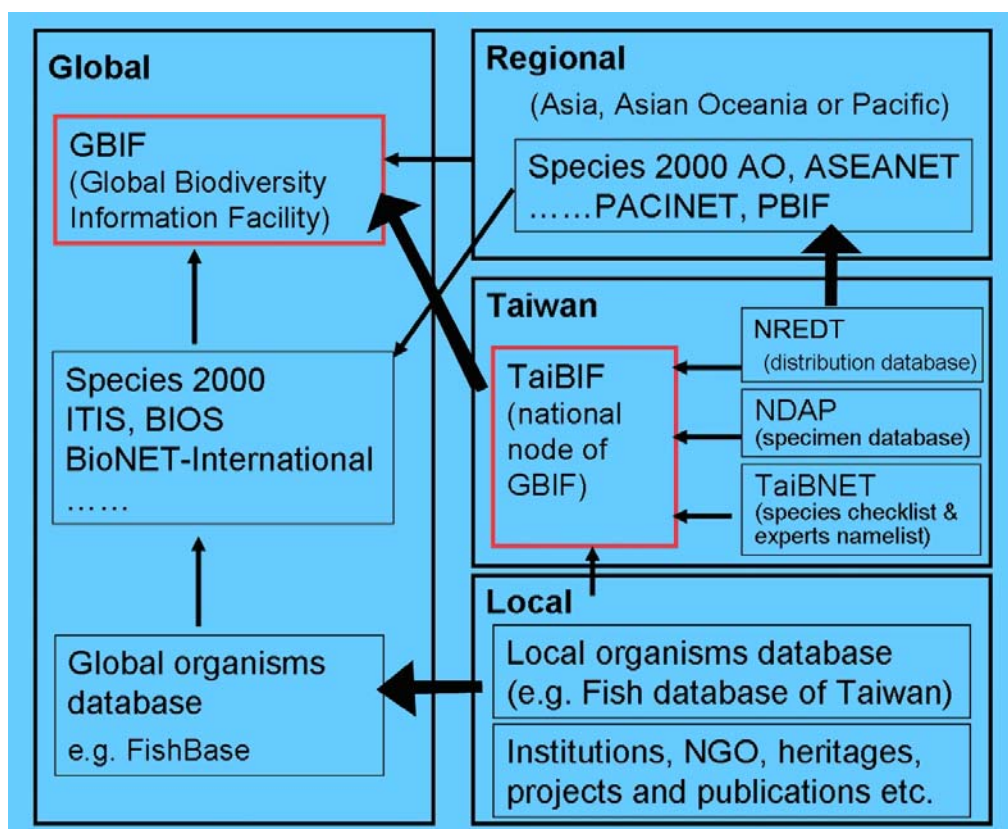


Figure 3. Integration of Taiwan Biodiversity National Information Network with global databases.

### 3 CONTENT OF TAIBIF

#### 3.1 Catalogs of Life and Local Experts on Biodiversity

TaiBIF is derived from the Catalog of Life and the Catalog of Local Experts on Biodiversity established by TaiBNET. The TaiBNET database has recorded more than 46,000 indigenous species and 650 local experts on biological diversity. It provides a search engine for data query based on an up-to-date database. The information includes classification hierarchy, Chinese name, author and year, and authorized citation and can be retrieved via species name, classification hierarchy, and full text search.

#### 3.2 Dynamic Species Information Linkage

Furthermore, users can click on the species hyperlink with the text of the scientific name and link to domestic or global databases as NDAP species data, Species 2000, and Discover Life to retrieve the details, such as the descriptions of characteristics, pictures, habitats, specimen records, relevant literature, geographical distribution, etc.



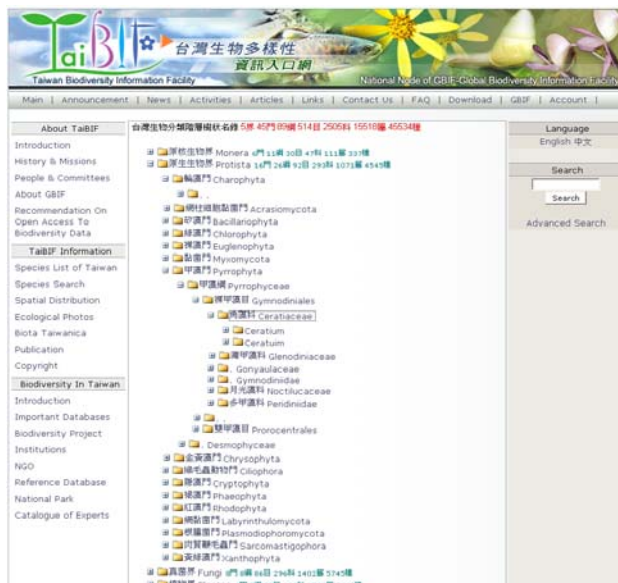


Figure 4. Use AJAX (Asynchronous JavaScript and XML) to develop the classification hierarchy of Taiwan native species.



Figure 5. *Pygoplites diacanthus* species integrated query

### 3.3 Integration of Specimen, Observation and Habitat Databases

TaiBIF provides a search engine for specimens, observation records, and distribution of native species. TaiBIF visitors can easily retrieve information about the distribution, habitat, specimen records, observation records, and ecological function of the species. The specimen data are the research results of NDAP. The observation data are collected by various experts from field studies and EIA reports sponsored by NSC, COA, or the Ministry of the Interior (MOI).

### **3.4 Ecological Image System**

The system provides services including catalog service, map service, photograph query, photography upload, and geography tags.

## **4 DATA PROVIDERS OF TAIBIF**

TaiBIF not only collects the local species lists and their specimens, distributions, photographs, and habitats but also cooperates with three major organizations and two universities in Taiwan for different purposes as following:

1. Science and Technology Information Center, National Applied Research Laboratories (STIC):  
Links the biodiversity database of documents and workshops between TaiBIF and STIC.
2. Council of Agriculture and National Taiwan University: For inquiries about 'Natural resources and Ecology GIS database in Taiwan' and 'The Biodiversity Information of Taiwan R.O.C.' for related laws, regulations, policy, and the international trend.
3. Ministry of Education and National PingTung University of Science and Technology: For inquiries about education information and linkage with Taiwan Forestry Research Institute's "100 of the World's Worst Invasive Alien Species and Endangered Species Database," Endemic Species Research Institute's Eco-engineering Database, and National Plant Genetic Resources Center's Germplasm Bank.

In addition, TaiBIF also supplies the information of national related organizations, publications, national parks, activities and news.

## **5 CONCLUSIONS AND FUTURE DEVELOPMENTS OF TAIBIF**

More than 1,000 hits are made each day on homepage of TaiBIF. TaiBIF will be greatly improved once 150,000 specimen records become available in 2006. The future developments of TaiBIF are as follows:

1. Continue to collect new biodiversity data and photos and integrate other sources into TaiBIF with the function for the new data.
2. Cooperate with government agencies and various institutions to facilitate the establishment, exchange, and integration of biodiversity information, including national geographic information system, fauna and flora of Taiwan, eco-engineering, germplasm bank, resource distribution,

ocean bank, alien invasive species, agriculture, ecological observation grid, environmental impact assessment, science education, specimen archives, references, cryobanking, and barcode of life.

3. Introduce advanced information technology to TaiBIF, such as data grid, web service, data mining tools, etc. and construct TaiBIF as a knowledge management platform to provide more comprehensive and powerful services.
4. Apply international standards, such as Ecological Metadata Language (EML), into TaiBIF and facilitate international cooperation.
5. Participate in international organizations such as GBIF, GTI (Riede, 2003), OBIS, Barcode of Life Data Systems (BOLD), CoMLD, Discover Life, etc. and improve international cooperation.

After completing the steps mentioned above, TaiBIF is expected to become a data sharing platform, an information generating engine, a bridge between Taiwan and the world, and a foundation for future biodiversity studies. The future tasks in 2006-2007 include:

1. Provide over 150,000 digitalized specimens to GBIF.
2. Collect data over 5,000 native species in specimen and ecological photos.
3. Increase species records to 47,000 and establish introduced species checklist.
4. Make Biota Taiwanica (English edition) internet accessible.
5. Establish a domestic reference of databases including the retrieval and integration of books, documents, compact disks, and databases.
6. Integration of ecological distribution of the plant species, animal and marine life.

Biodiversity informatics is an emerging field, which will provide integrated services of distributed multi-model, multi-type, and multi-disciplinary content resources, to foster new research paradigms and to derive new knowledge from them. A new infrastructure supporting better data collection, analysis, query and access, management, resource discovery, and dissemination is necessary to meet the requirements of biodiversity researchers, data curators, museums, the public, etc. TaiBIF is an organization which collects and integrates biodiversity data of various institutes in Taiwan. Based on the experiences of TaiBIF and related works, a web-based content management system (CMS) will be the most viable solution to reach the goals stated previously.



## **6 REFERENCES**

Buckeridge, J. St J. S., & Gordon D. (2000) Species 2000 New Zealand: Outcomes of the February Symposium. *XVIII (New) International Congress on Zoology*, Athens, Greece.

Edwards J. L., Lane M.A., & Nielsen E.S. (2000) Interoperability of Biodiversity Databases: Biodiversity Information on Every Desktop. *Science* 29 289(5488), pp 2312 – 2314.

Bisby, F.A. (2000) The Quiet Revolution: Biodiversity Informatics and the Internet. *Science* 29 289(5488), pp 2309 – 2312.

Huang, J.C., Wu, H.H., Lai, K.C., Shao, K.T., Peng, C.I., & Yen, E. (2004) Taiwan Biodiversity Information Facilities Project: Developing information infrastructure and integrating databases in support of TaiBIF. *Greater China GIS Conference 2004*, Hong Kong, China.

Shao, K.T., Peng, C.I., Yen, E., Lai, K.C., Wang, M.C., Lin, J., Lee, H., & Chen, Y.X. (2005) Integration of Biodiversity Database in Taiwan and Linkage to Global Databases. *2005 International Workshop on Integrated Biodiversity & Natural Specimens Databases & Forum of Species 2000 Asia-Oceania*. Taichung, Taiwan.

Olson, R. J. & McCord, R.A. (2002) Archiving ecological data and information. In: Michener, W. K. & Brunt, J.W., (Eds.) *Ecological Data: Design, Management and Processing*. Oxford, UK: Blackwell Science.

Riede K. (2003): Biodiversity Informatics in Germany: ongoing projects and their possible contribution to the Global Taxonomy Initiative (GTI). *Global Taxonomy Initiative in Asia. National Institute for Environmental Studies* (pp. 294-300). Japan.