BIODIVERSITY INFORMATION RETRIEVAL ACROSS NETWORKED DATA SETS

By

Dr Sarinder Kaur

Institute of Biological Sciences
Faculty of Science
University of Malaya, Kuala Lumpur

ISKO UK Conference 2009

Sir Robert May (1998)

"*... The next century will be the '**A**ge of Biology', just as [the twentieth century was] the age of physics and astronomy. Specifically, those countries who best know how to correlate, analyze and communicate biological information will be in the leading position to achieve economic and scientific advances.*"

Biodiversity information:

disperse
unreported
inaccessible

BIODIVERSITY INFORMATION: MALAYSIAN PERSPECTIVE

Identified as one of the world's twelve mega-diversity areas with extreme rich biological resources

over 15,000 species of higher plants,
300 species of mammals,
254 species of breeding birds,
198 species of amphibians,
379 species of reptiles,
over 150,000 species of the invertebrates,
over 4,000 species of marine fishes
449 species of freshwater fishes in Malaysia

(Earth Trends, 2003)

Current status

Some institutions have information documented in the form of databases, while other institutions are looking into it

Databases are in heterogeneous formats.

Few databases are on the Web, while many are available offline

Some of these databases are well-structured, others are largely project /species specific and/or unstructured

These databases exist independently

There is no meta-data (data- dictionary)

NO FRAMEWORK TO LINK THE SCATTERED DATA SO AS TO FACILITATE EXCHANGE OF DATA AMONGST THE DIFFERENT DATABASES AVAILABLE IN THE COUNTRY.

Need to overcome this setback….

HOW??

Dissemination and retrieval of information in a networked environment....

COMMUNICATION of biodiversity information (*biodiversity informatics*)

LINK all the disperse information clusters in remote and distributed medium

THE SOLUTION

Rationale for indigenous DATABASE INTEGRATION SYSTEM

?

Implementation issues existing systems
No control

INDIGENOUS framework for integrating DISTRIBUTED HETEROGENUOUS *relational biodiversity databases*

to facilitate biodiversity data sharing among the scientific contemporaries especially in Malaysia.

Philosophy:

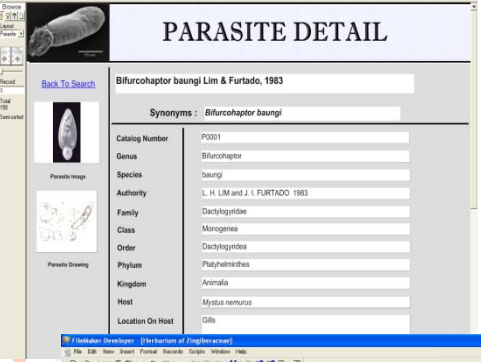
Database KEPT & MAINTAINED by DATA OWNER

SIMPLICITY is stressed to offer a USABLE system for the scientific community in Malaysia and future integration with other global networks.

The **FOCUS**

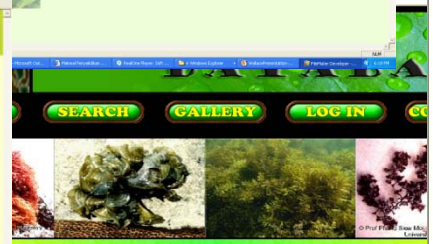
- DEVELOPMENT OF A NEW DATABASE INTEGRATION SYSTEM
- DEVELOPMENT OF BIODIVERSITY DATABASES ACCORDING TO REQUIREMENTS SET BY PARTNERS

*R*ELATIONAL *B*IODIVERSITY *D*ATABASES

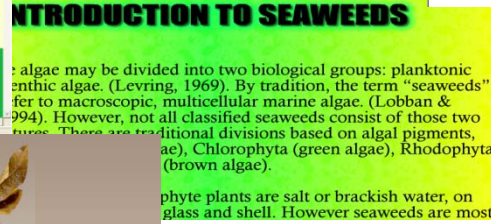

PARASITE DETAIL
Back To Search
Bifurcohaptor baungi Lim & Furtado, 1983
Synonyms : Bifurcohaptor baungi
Catalog Number P0001
Genus Bifurcohaptor
Species baungi
Authority L. H. LIM and J. I. FURTADO 1983
Family Dactylogyridae
Class Monogenea
Order Dactylogyridea
Phylum Platyhelminthes
Kingdom Animalia
Host Mystus nemurus
Location On Host Gills

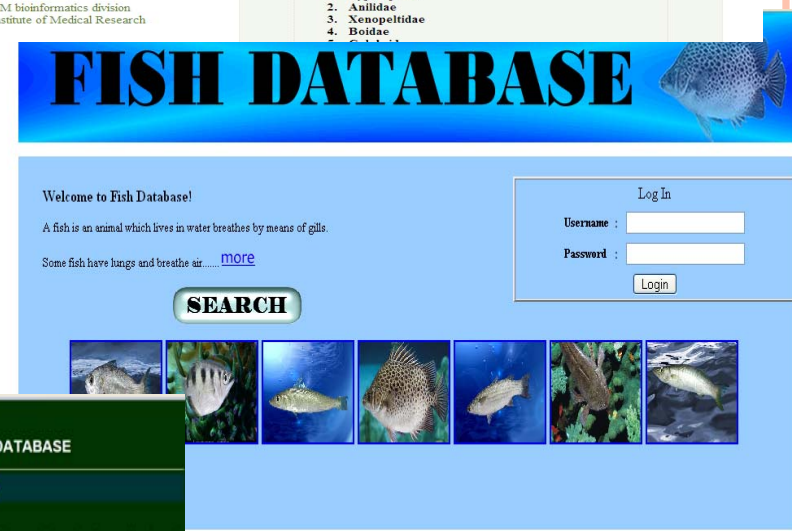

MALAYSIAN INDIGENOUS FERN

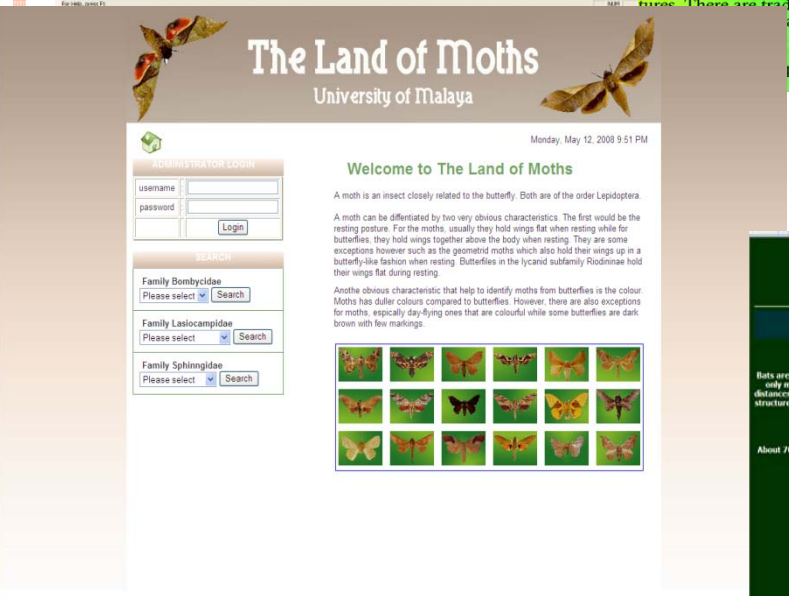

Herbarium of Zingiberaceae
Alpinia aquatica

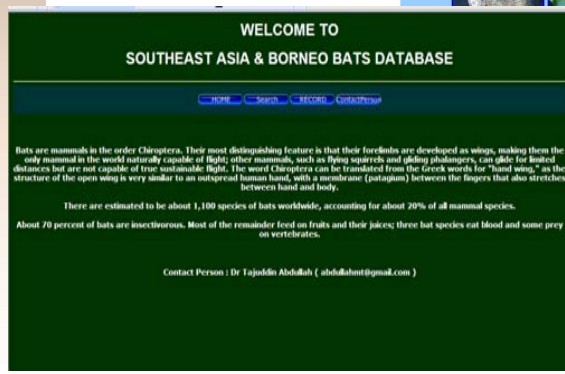

SEARCH
GALLERY
LOG IN


Snakes of Peninsular Malaysia
- archival records
Welcome, Guest. Please login to access administrator authority.
Username
Password
login
Menu
Home
Search record
Classification
Other Information
Picture Gallery
Reference
Acknowledgement
Related links
UM bioinformatics division
Institute of Medical Research
Introduction to Snakes of Peninsular Malaysia - Archival Records Database
This database of archive records on the Snakes of Peninsular Malaysia was developed as part of a broader Biodiversity Database Information System in the Institute of Biological Sciences, University of Malaya. This is a collaborative project between the Institute for Medical Research and the University of Malaya. The database contains more than 500 records of 100 species of snakes collected in the Malay Peninsula and curated in the Biomedical Museum of the IMR.
For more information
Snakes of Peninsular Malaysia
In Malaysia, there are about 129 species of snakes belonging to 8 families.
1. Typhlopidae
2. Anilidae
3. Xenopeltidae
4. Boidae


NTRODUCTION TO SEAWEEDS
algae may be divided into two biological groups: planktonic
enthic algae. (Levring, 1969). By tradition, the term "seaweeds"
fer to macroscopic, multicellular marine algae. (Lobban &
994). However, not all classified seaweeds consist of those two
ae), Chlorophyta (green algae), Rhodophyta
(brown algae).
phyte plants are salt or brackish water, on
glass and shell. However seaweeds are most


FISH DATABASE
Welcome to Fish Database!
A fish is an animal which lives in water breathes by means of gills.
Some fish have lungs and breathe air..... more
SEARCH
Log In
Username :
Password :
Login
Home
Forum
Glossary
Contact
LivingWeb
FishBase
Copyright ©Institute of Biological Sciences, Faculty of Science, University Malaya


The Land of Moths
University of Malaya
Monday, May 12, 2008 9:51 PM
ADMINISTRATOR LOGIN
username
password
Login
SEARCH
Family Bombycidae
Please select
Search
Family Lasiocampidae
Please select
Search
Family Sphinngidae
Please select
Search
Welcome to The Land of Moths
A moth is an insect closely related to the butterfly. Both are of the order Lepidoptera.
A moth can be diffentiated by two very obvious characteristics. The first would be the resting posture. For the moths, usually they hold wings flat when resting while for butterflies, they hold wings together above the body when resting. They are some exceptions however such as the geometrid moths which also hold their wings up in a butterfly-like fashion when resting. Butterflies in the lycanid subfamily Riodinnae hold their wings flat during resting.
Anothe obvious characteristic that help to identify moths from butterflies is the colour. Moths has duller colours compared to butterflies. However, there are also exceptions for moths, espically day-flying ones that are colourful while some butterflies are dark brown with few markings.


WELCOME TO
SOUTHEAST ASIA & BORNEO BATS DATABASE
HOME
Search
RECORD
ContactPerson
Bats are mammals in the order Chiroptera. Their most distinguishing feature is that their forelimbs are developed as wings, making them the only mammal in the world naturally capable of flight; other mammals, such as flying squirrels and gliding phalangers, can glide for limited distances but are not capable of true sustainable flight. The word Chiroptera can be translated from the Greek words for "hand wing," as the structure of the open wing is very similar to an outspread human hand, with a membrane (patagium) between the fingers that also stretches between hand and body.
There are estimated to be about 1,100 species of bats worldwide, accounting for about 20% of all mammal species.
About 70 percent of bats are insectivorous. Most of the remainder feed on fruits and their juices; three bat species eat blood and some prey on vertebrates.
Contact Person : Dr Tajuddin Abdullah (abdullahmt@gmail.com)

DATABASE INTEGRATION

PROPOSED DATABASE INTEGRATION SYSTEM

- *Links remote databases on networked environment.*
- *Supports heterogeneous data format.*
- *Supports heterogeneous database management systems.*
- *Links databases hosted in Windows and UNIX based platforms.*
- *Provides data security for database owners by allowing them to keep and maintain their own data and to choose information to be shared and linked.*

PROTOTYPE

Microsoft Windows

ASP

PHP

XML

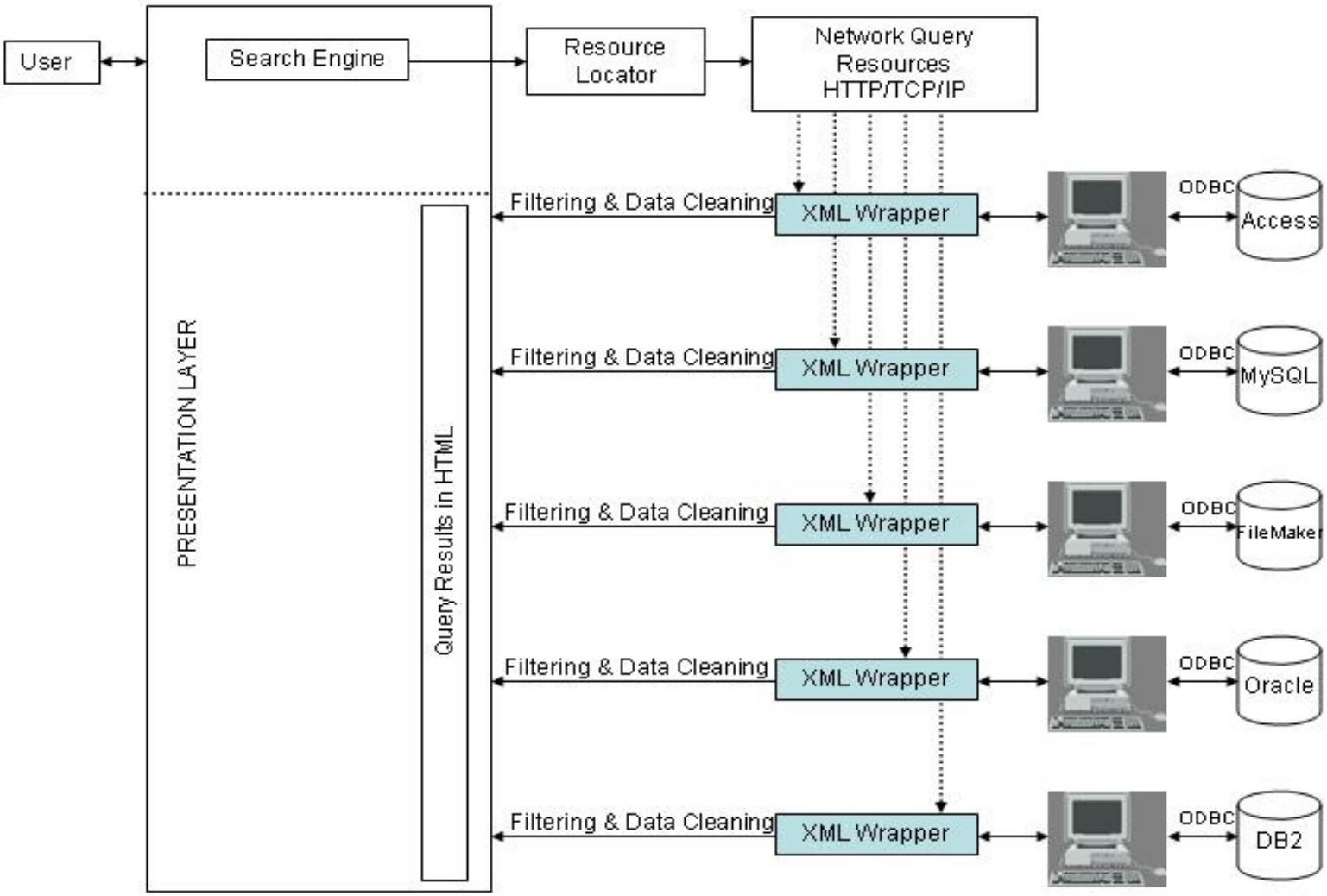
HTML

VBScript and Jscript

DBMS- FileMaker, mySQL, DB2..

Search Mechanism

ARCHITECTURE



Summary

a. Created data format & relational biodatabases
b. Developed indigenous integration system

c. In summary

Achieved objectives set out which are:-

- *Links remote databases on networked environment.*
- *Supports heterogeneous data format.*
- *Supports heterogeneous database management systems.*
- *Links databases hosted in Windows and UNIX based platforms.*
- *Provides data security for database owners by allowing them to keep and maintain their own data and to choose information to be shared and linked.*

FUTURE FOCUS

To further develop & enhance this biodiversity content architecture for our country

- e-library (Parasitofaunae Malaysiana)
- Data-security
- Use of Semantic technology
- Populate Relational Databases

RESEARCH PROJECTS

Parasites Informatics Network System (PINS)

In our effort to aid research in Parasitology in Malaysia we have developed PINS which is essentially a parasite taxonomic database with inputs from experts in the field who will man their respective database with technical assistance from us. We will be incorporating into this database related IT Alternatives to Taxonomy which are currently developing as well as ecological and biological data which will be linked semantically.

[*Sarinder Kaur & Lim LHS and other researchers*]

Image & text recognitions for taxonomy: retrieving, managing, enhancing and recognizing different types of images

This project is initiated to develop an image-recognition system to aid in species identifications (taxonomy). It is first necessary to develop a database to store images in different formats either from our research or retrieved from different sources. The retrieved images will probably have to be enhanced (if necessary) for image recognition. Our query to search and capture images from publications will be textual in nature since most of the images and illustrations are tagged textually. The system we are developing is for monogenean species, a parasitic flatworm.

[Arpah (PhD candidate), Sarinder Kaur & Lim LHS]

2-D to 3-D modeling of organisms for structural morphology

Visualisation in 3-D provides more information on the spatial relationships of objects or structural morphological features in the case of organisms. 3-D visualization is important in understanding structural and functional morphology and also important for taxonomic and identification purposes. The aim of this study is to construct 3-D models from 2-D images. In this study several softwares capable of doing this have been evaluated and we are currently using 3-d Max programme because of its advantage of being able of rendering irregular shapes using polygons.

[Teo Bee Guan (MSc candidate), Lim LHS, Sarinder Kaur]

Image recognition using DAISY

This project looks into using an available software on image recognition for taxonomy to aid in identifications. The group of interest is to use the software to image moths which have different shapes, sizes and colour patterns. This project involves setting up a database of selected groups of moths.

[Evelyn (MSc candidate), Sarinder Kaur & Lim LHS]

THANK YOU