GBIF Participant posters

Covering a wide range of topics, the posters presented in this collection were displayed by GBIF Participants at the 18th GBIF Governing Board Meeting and associated events, in Argentina in October 2011.

To broaden the use of the posters, and to ensure the opportunity for the Participants not present in Argentina to see the amazing posters, we have collected them in an electronic "booklet". The posters are provided in PDF format. If you have problems accessing the files, or if you wish to receive the files in a higher resolution, please contact us at info@gbif.org.

The authors hold the copyright to their respective posters. If you wish to use the poster, or if you are interested in more information about the topics presented, please contact the relevant person directly; an overview of contact details for each poster is shown in the table below.

The posters are also available electronically from the GBIF website www.gbif.org/orc.

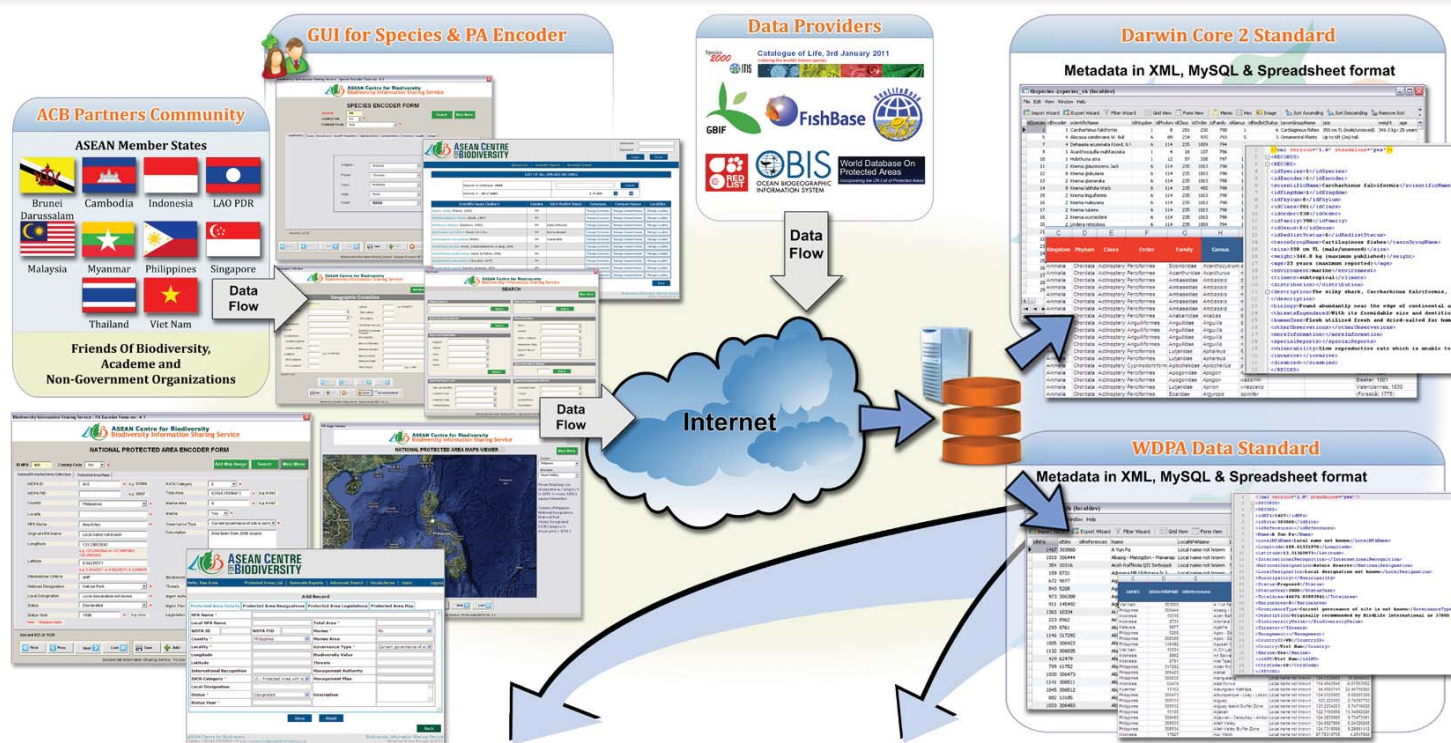| Participant / Node	Poster title	Contact person	Contact email
ACB	Asean Centre for Biodiversity: Data sharing, interoperability and products	Dr. Sheila G. Vergara	sgvergara@aseanbiodiversity.org
DanBIF	*DanBIF: Danish Biodiversity Information Facility*	Lotte Endsleff Isabel Calabuig	lendsleff@snm.ku.dk icalabuig@snm.ku.dk
EOL	*ENCYCLOPEDIA OF LIFE (EOL) version 2: A Vision for a Global Collaborative Project*	Katja Schulz	SchulzK@si.edu
GBIF Argentina	*The Collections Digitization Project in MACN*	Dr. Martín Ramírez	ramirez@macn.gov.ar
GBIF Cuba	*Cuba GBIF Mentoring*	Dr. Francisco Cejas Rodríguez	franciscocejas@yahoo.es
GBIF Finland	*The Process of Digitising Natural History Collection Specimens at Digitarium*	Hannu Saarenmaa	hannu.saarenmaa@uef.fi
GBIF Germany: Bacteria & Achaea	*Open up the treasure chest of prokaryotic diversity for next-generation data analysis*	Jörg Overmann	joerg.overmann@dsmz.de
GBIF Germany: Diversity Workbench	*The Diversity Workbench - a virtual research environment and archive structure for biodiversity data*	Dagmar Triebel	triebel@bsm.mwn.de
GBIF Germany: Evertebrata II	*The GBIF-Germany Node for Molluscs, Spiders and Myriapods*	Gerhard Haszprunar	haszi@zsm.mwn.de
GBIF Germany: Fungi & Lichens	*The GBIF-Germany Node for Fungi and Lichens: Work stations and data flow*	Dagmar Triebe	triebel@bsm.mwn.de
GBIF Germany: Insects	*Out the drawer and on the web - Large scale metadata capture and publication of insect collections*	Alexander S. Kroupa	alexander.kroupa@mfn-berlin.de
GBIF Germany: Plants, Algae and Protists	*Towards the second decade - New prospects for the botanical community within GBIF Germany*	Sabine von Mering	s.vonmering@bgbm.org
GBIF Germany: Vertebrates	*Enlarging Information on Vertebrate Collection data*	Renate van den Elzen	R.Elzen.ZFMK@Uni-Bonn.De
GBIF Germany: DNA-Bank-Network	*DNA Bank Network - How GBIF technologies enable sample referencing and data integration for biological DNA repositories*	DNA Bank Network	contact@dnabank-network.org
GBIF Germany: Project OpenUp!	*GBIF Meets the Arts – Natural History in a Cultural Context*	Project OpenUp!	OpenUpCoord@bgbm.org
GBIF-D	*GBIF-D enables free access to Biodiversity Data: How to contribute?*	Maren Gleisberg	m.gleisberg@bgbm.org
GROMS	*Mobilizing Migratory Species Data – integrating Global Register of Migratory Species and GBIF data*	Dr. Klaus Riede	k.riede.zfmk@uni-bonn.de
NLBIF (1)	*Towards a Taxonomic Backbone for Europe*	Yde de Jong	yjong@uva.nl
NLBIF (2)	*BioVeL Developing a Biodiversity Virtual e-Laboratory*	Yde de Jong	yjong@uva.nl
TanBIF (1)	*National BioPortal: The getaway for Tanzania Biodiversity Data and Information*	Ms. Hulda Gideon	hgideon@costech.or.tz
TanBIF (2)	*TanBIF-GIS Tool For Ecological Niche Modeling, Prediction and Assessment of Biodiversity*	Ms. Hulda Gideon	hgideon@costech.or.tz
University of Kansas (1)	*An Online Biodiversity Informatics Training Curriculum: Marshalling the Collective Knowledge of the GBIF Community*	Town Peterson	town@ku.edu
University of Kansas (2)	*GBIF-enabled Knowledge of Birds of the World: A Challenge to Fill Gaps and Complete Knowledge*	Town Peterson	town@ku.edu
WII	*Wildlife Institute of India – Atlas of Living Australia Mentoring Project*	Dr. Gautam Talukdar	gautam@wii.gov.in

ASEAN CENTRE FOR BIODIVERSITY
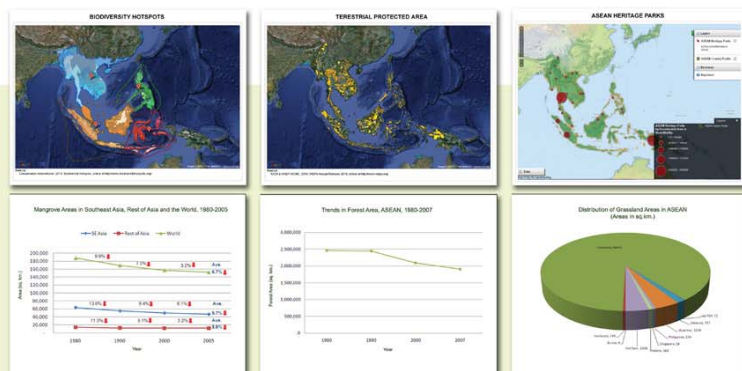
DATA SHARING, INTEROPERABILITY and PRODUCTS

The biodiversity information sharing service (BISS) of the ASEAN Centre for Biodiversity supports collaboration among the ASEAN Member States (AMS). Based on data and information shared by the AMS through their clearing house mechanisms and national reports, BISS contains an inventory of biodiversity resources of the ASEAN region and constantly develops a growing collection of knowledge products.

To further enhance current information holdings, interfaces for the species (species encoder) and protected areas (protected areas encoder) databases were developed. These are in the final stages of testing and will be made available to all AMS in the second quarter of 2011.
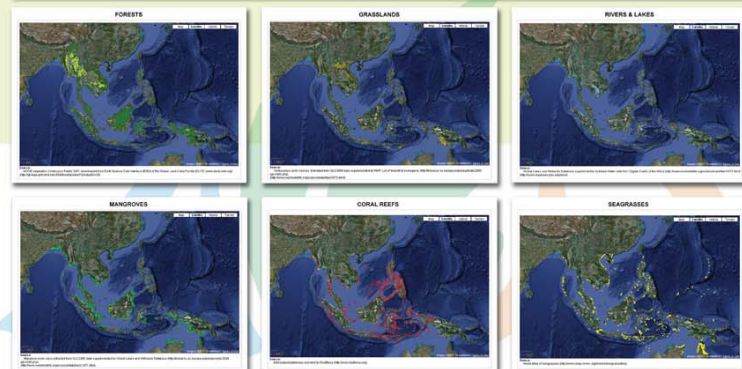
Having access to these interfaces will ease data entry and future sharing, promote a common biodiversity information platform for the region, provide the necessary database frameworks for AMS that have not developed their own, lay the foundation for regional analysis that will support future biodiversity related decisions at the national and regional levels, and contribute to having information available for AMS reports to Multilateral Environmental Agreements.
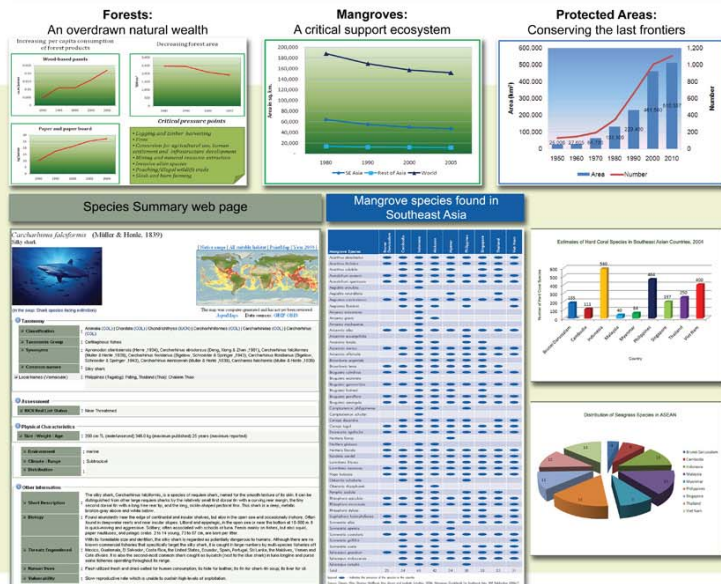
GUI for Species & PA Encoder

Data Providers

Catalogue of Life, 3rd January 2011

GBIF · FishBase · RED LIST · OBIS Ocean Biographic Information System · World Database On Protected Areas

Darwin Core 2 Standard

Metadata in XML, MySQL & Spreadsheet format

ACB Partners Community

ASEAN Member States

Brunei Darussalam · Cambodia · Indonesia · LAO PDR · Malaysia · Myanmar · Philippines · Singapore · Thailand · Viet Nam

Friends Of Biodiversity, Academe and Non-Government Organizations

Data Flow

Data Flow

Internet

Data Flow

Data Flow

WDPA Data Standard

Metadata in XML, MySQL & Spreadsheet format

REGIONAL SUMMARIES

BIODIVERSITY HOTSPOTS

TERESTRIAL PROTECTED AREA

ASEAN HERITAGE PARKS

Mangrove Areas in Southeast Asia, Rest of Asia and the World, 1980-2005

Trends in Forest Area, ASEAN, 1980-2007

Distribution of Grassland Areas in ASEAN (Areas in sq.km.)

ECOSYSTEMS STATUS and TRENDS

Forests: An overdrawn natural wealth

Mangroves: A critical support ecosystem

Protected Areas: Conserving the last frontiers

Increasing per capita consumption of forest products

Wood-based panels · Decreasing forest area

Paper and paper board

Critical pressure points
- Logging and timber harvesting
- Conversion for agricultural use, human settlement and infrastructure development
- Mining and mineral resource extraction
- Invasive alien species
- Flooding, illegal wildlife trade
- Slash-and-burn farming

Species Summary web page

Mangrove species found in Southeast Asia

Carcharhinus falciformis (Müller & Henle, 1839)

Silky shark

Estimates of Hard Coral Species in Southeast Asian Countries, 2001

Distribution of Seagrass Species in ASEAN

The species database takes off from the taxonomic tree defined by the Catalogue of Life (COL). Where information is missing from COL, ACB accesses taxonomic information from the IUCN Red List, World Register of Marine Species (WORMS), FishBase, OBIS, SealifeBase, Living Natural Treasures, Germplasm Resource Information Network (GRIN) and World Agroforestry Centre.

ECOSYSTEMS OCCURENCE

FORESTS · GRASSLANDS · RIVERS & LAKES

MANGROVES · CORAL REEFS · SEAGRASSES

KEY BIODIVERSITY AREA · IMPORTANT BIRD AREAS · MARINE PROTECTED AREA

The loss and transformation of habitats are key causes of biodiversity loss in the ASEAN region. BISS provides map-based information on the location and extents of ecosystems in the region.

With support from APBON and other donors, BISS will provide year-based maps to demonstrate habitat trends and thus inform policy preparation aimed at the conservation of critical habitats and ecosystems. The above ecosystem maps are available from http://bim.aseanbiodiversity.org/biss/

Knowing the availability and status of species per country is important to defining and refining conservation strategies for each AMS. With information drawn from the Key Biodiversity Area maps and similar map-based information, each AMS will be able to prioritize areas for conservation.
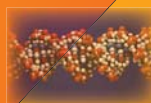
DanBIF

Danish Biodiversity Information Facility
National Node of GBIF – Global Biodiversity Information Facility

❀ **DanBIF – Danish Biodiversity Information Facility** (www.danbif.dk) – represents and organises Denmark's – incl. the Faroe Islands and Greenland – activities and obligations in connection with the Danish membership of GBIF.

❀ DanBIF consists of a network of members and other interested parties (museums, research institutions, public administration, companies etc.), a governing board and a secretariat placed at the Natural History Museum of Denmark, next to the international GBIF secretariat.

❀ DanBIF is co-funded by the Danish Council for Independent Research (Natural Sciences of the Ministry of Science, Technology and Innovation), and by The Faculty of Science, University of Copenhagen.

Contact to data owners and users in Denmark

The DanBIF network is constantly growing and currently includes 62 institutions, organisations etc., with more than 180 individuals from Denmark, the Faroe Islands, and Greenland. DanBIF members are appointed by The Danish Council for Independent Research and include representatives of research institutions involved in biodiversity research. 'Other interested parties' includes a wide spectrum of data owners and others. Personal contact to the DanBIF community and investigative activity towards engagement in GBIF of new data owners are ongoing.

The DanBIF portal and other avenues of dissemination

❀ The enthusiasm and involvement of data owners and users are a prerequisite for DanBIF's activities. Therefore, DanBIF prioritises detailed and widespread dissemination about DanBIF and GBIF to different segments of users, primarily through the bilingual DanBIF portal (www.danbif.dk). Many services are available through the portal. Other dissemination initiatives include newsletters, articles, exhibitions, radio, TV, lectures, email information service, etc.

❀ DanBIF has organised a number of meetings and conferences about topical biodiversity themes and co-organised an international conference about invasive species.

❀ A GBIF participant node is expected to constitute a 'meeting point' for information and activities related to biodiversity. The DanBIF secretariat hosts and administrates a number of external portals and web services. For most of these, the DanBIF secretariat functions as web editor.

Mobilisation of biodiversity data housed in Denmark includes three overlapping activities:

❀ **Processing of existing data sets**
DanBIF mobilises existing electronic data on physical specimens or observations of organisms. In contrast, the huge task of doing the practical work of transferring data to an electronic format (digitizing) rests with the data owners. The DanBIF secretariat has the knowledge and experience to function as a consultant for digitisation projects.

❀ **Identifying and acquiring 'hidden' data sets**
The first step in data acquisition consists in registering metadata, e.g. which major taxonomic group the data set concerns and the overall geographical coverage and time period for collecting or observing the species. The metadata follow the dataset into the GBIF portal and enable the data user to evaluate the dataset for specific tasks. DanBIF has developed a catalogue of biodiversity collections and databases housed in Denmark: *Collections Metadata Inventory*.

❀ **Developing user-friendly IT tools for upload of collections and observations into DanBIF's database**
The vast majority of existing knowledge about biodiversity is still not available electronically but exists as specimens in natural history collections, paper-based literature etc. DanBIF has an important task in providing tools and know-how to the data owners for them to digitize such data. DanBIF has developed e.g. the *DanBIF Bird Registry* for data digitisation and presentation by the Natural History Museum of Denmark's bird section.

DanBIF objectives

❀ **Make biodiversity data housed in Denmark available for search through the GBIF and DanBIF portals, with emphasis on data concerning Danish species.** The majority of DanBIF's resources are channelled into this objective.

By medio 2011 Danish (incl. Faroese) data owners have made more than 5.3 million data records searchable through GBIF, of which 90% concern occurrences in Denmark.

❀ **Adapt DanBIF's IT infrastructure according to GBIF development and data owner requirements.** As the GBIF network and data volume grows, the organisation of GBIF becomes increasingly decentralised. DanBIF adapts its IT infrastructure for data access, storage and sharing accordingly.

❀ **Increase awareness of DanBIF and GBIF, to inspire up- and download of GBIF data by Danish researchers and organisations.** DanBIF endeavours to establish contact to all relevant data providers and users, e.g., basic and applied research institutions, public administration, private companies and civil scientist societies, who possess biodiversity data and/or are involved in biodiversity research.

❀ **Develop DanBIF into a permanent infrastructure.** DanBIF has proven its value as a necessary and dependable link between GBIF and Danish biodiversity stakeholders. DanBIF has established expertise and resources which are vital for a sustained Danish contribution to and use of GBIF. Like other international research infrastructures, GBIF requires stable national structures to ensure access to the worldwide information bases.

DanBIF's Multimedia Server
The Multimedia Server (www.multimedia.danbif.dk) presents collections of biodiversity photographs, videos, sound files and literature. Inclusion of images often significantly increases the usefulness of data about a particular specimen. GBIF's data portal shows links to such images but the image files themselves must be housed by the data owner. However, constant, online access to the pictures via GBIF does not work if the files are stored on the researchers' personal computers. Therefore DanBIF has developed and maintains a Multimedia Server complying with GBIF guidelines. The user interface is accessed through a web browser. This system makes it possible to store multimedia as well as literature files in a safe and sustainable way and to publish the files together with curatorial and georeference information.

DanBIF's Data Hosting Centre
Together with the GBIF secretariat, DanBIF is building a Data Hosting Centre for use by Danish data owners, but also for orphaned data sets from other countries and organisations. This will provide contributors of data with safe storage and advanced data curation tools. The hosted data sets will be accessible for search in the GBIF context. DanBIF has contributed to the construction and implementation of the centre. The development work will be expanded with, i.a., tools for data cleaning and other IT curatorial tools, including documentation and manuals.

Data about species - www.allearter.dk
Occurrence data is one form of data. Equally important are qualitative data about species: taxonomy, distribution, habits, evolution and relationships. The point of reference for all these disciplines is the species name. A list of all species known from Denmark is being built and maintained at *www.allearter.dk*. This is a DanBIF project working to list the names of all Danish animals, plants, fungi etc. along with other taxonomic attributes, and to make this information easily and freely available for download at one web site.

Inconsistencies regarding scientific and Danish names as well as taxonomy are the rule rather than the exception. With the help of experts, allearter.dk will form a common platform to promote a Danish unified use of names in literature, municipal records, collections, etc. The number of species in allearter.dk now exceeds 31,000. Most groups are covered - and more are on the way. The total number will be up to 35,000 known species (there are estimated 35-40,000 species occurring in Denmark). The list also contains information on protected species, e.g. red list species and species protected by international conventions, and on alien and invasive species.

DanBIF was established July 1. 2001 through a contract of co-funding between the Danish Natural Science Research Council and the Faculty of Science, University of Copenhagen, Denmark.

The DanBIF board includes a director, a chairman and one representative of the biological sciences from each of the major Danish universities (Aarhus, Copenhagen, Odense).

DanBIF is hosted by the Natural History Museum of Denmark, Zoological Museum, University of Copenhagen. The DanBIF Secretariat includes a Nodes Manager, an IT Manager and a Scientific Communications Officer & web editor.

Contact: Danish Biodiversity Information Facility
Natural History Museum of Denmark
University of Copenhagen
Universitetsparken 15
DK-2100 Copenhagen OE
DENMARK

www.DanBIF.dk

ENCYCLOPEDIA OF LIFE (EOL) version 2:
A Vision for a Global Collaborative Project

Katja Schulz, Jennifer Hammock, and Cynthia Parr

National Museum of Natural History, Smithsonian Institution, Washington, D.C., USA

BROAD PARTICIPATION & EXPERT VETTING

EOL integrates text, name data, images, videos, and maps from multiple sources. Content providers range from natural history museums, government agencies, and international research consortia to members of popular sites like Flickr and Wikipedia. Information from unreviewed sources is clearly marked. Expert curators ensure the quality of the core collection by authenticating materials submitted by diverse projects and contributors.

EOL content can be reused

EOL is building an open access collection of biodiversity information. All content can be redistributed and reused. Data can be downloaded through the EOL API:

eol.org/api

EOL-GBIF PARTNERSHIP

1. Occurrence Maps

GBIF provides specimen-level geographic occurrence data to EOL in the form of taxon-specific map visualizations.

2. Standards Development

The new EOL data transfer schema is based on elements from the Darwin Core standard. EOL is also supporting efforts of the National Biodiversity Institute Costa Rica (INBio) to develop an extended Darwin Core architecture for importing species information from Plinian Core based systems.

EOL GLOBAL PARTNERS

EOL is a global collaborative project. Internationalization of interfaces and content are important priorities.

Biodiversity institutions & initiatives on five continents are establishing EOL sister sites serving the specific needs of particular regions. Providing information in local languages, these sites may use EOL branding, software, and content. EOL global partners also share information about the local flora & fauna with the rest of the world through the main EOL site.
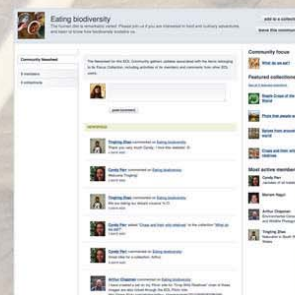
A PLATFORM FOR COLLABORATION & INTERACTION

The recent release of EOL version 2 has introduced the concepts of virtual collections and communities to the project. These features greatly increase the interactivity of the site and provide diverse opportunities for collaborative knowledge discovery, content creation, and data management.

EOL **collections** are annotated lists of items such as taxa, text or media objects (images, videos, sounds), EOL member profiles, communities or other collections. Participants can use collections to organize EOL content from their personal point of view.
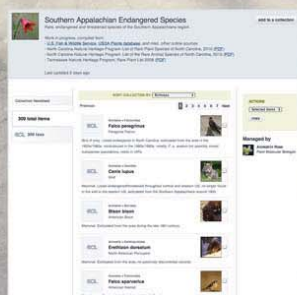
EOL **communities** provide a virtual place for like-minded people to congregate and interact. Members can showcase and collaborate on collections, post comments and view an activity stream of events of interest to the community.

SETTING PRIORITIES FOR CONTENT ACQUISITION

To provide rich pages for taxa of greatest interest to our audience, our content acquisition strategy is guided by a hotlist of high impact species:

- species of conservation concern
- invasive species
- agents of disease
- species of agricultural significance
- keystone species
- biological indicators
- charismatic species
- model organisms

What is a rich taxon page?

We have developed a set of page richness metrics to track progress toward our content goals. These metrics take into account the following factors:

content breadth
    number of topics & media types
content depth
    number of words, media files
content source diversity
    number of content providers
content effectiveness
    audience, ratings, readability

CURRENT EOL CONTENT
(September 2011)

Pages with content .... >750,000

Pages with text .... >710,000

Pages with images .... >125,000

Pages with vetted content ..... >670,000

Pages with rich content .... >50,000

Pages with literature links ... >1,150,000
(courtesy Biodiversity Heritage Library, BHL)

Content Partners 180

Image Credits: World map courtesy NOAA/NGDC
All photos by Flickr members, released under creative commons licenses (*creativecommons.org*)

pygmy leatherjacket by Doug Anderson   BY-NC
Farmer tending to his rice fields by clip works   BY-NC-SA
surfer chicks by Dave Young   BY
mushroom 3 by Dan Bergstrom   BY-NC
Green on green! by Tarique Sani   BY-NC-SA

The Collections Digitization Project in MACN

Zermoglio P.F., M.L. Barone, M.F. D'Jonsiles, M.S. De Majo, M.D. Ezcurra, V.D. Fadel, M.E. Guala, M.C. Luna, M.F. Marfil, O. Trucco, M. Viglino, D. Rodríguez & M.J. Ramírez.

Box 1. Aurora is a database application designed and maintained in MACN, for capturing and maitaining biological Collection specimen data. It is based in MS-Access and MS-Visual Basic, and allows several users to work simultaneously.
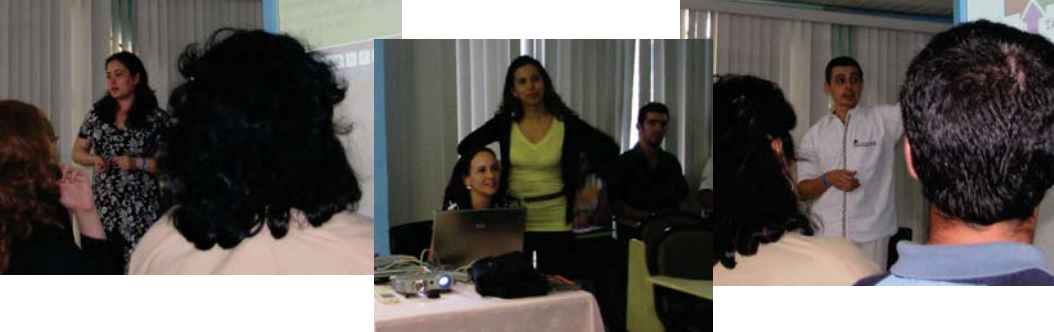

It has different screens where information about collectors, taxonomy and geographical location can be settled and organized. Each screen allows the user to modify certain information that could be related with many registers, like collector's names, and so makes it easier for him to handle large sets of data at a time. A main screen for each register is displayed, and gathers all the information referring to the specimens. It includes fields for identification data (as catalogue number and location in the collection), collector and date, taxonomy, geographical location and it's georreferenciation. Data fields in Aurora are mapped to corresponding terms in DarwinCore, and are currently published to GBIF, SNDB and IABIN data portals using TAPIR. We plan to publish our data through IPT in the near future.

The MACN's Collections Digitization Project began in 2008 as the first initiative of the kind for the institution. Its main purpose is to convert registers of specimens held in the museum from paper to digital format and to make the information available for the rest of the scientific community. For achieving that task, every year interns from Buenos Aires University are contracted and are assigned to a certain collection, in which they work entering data using the application Aurora (Box 1).

The organization of the project includes a general coordinator, an informatics responsible, a processes coordinator and interns from the Facultad de Ciencias Exactas y Naturales, Buenos Aires University. Interns are assigned to collections depending on their academic background and interests. Their work consists mainly of digitazing the specimen information contained either in individual labels or in accession books. They count with the continuous advice and guidance of the curators from each collection.

The type of information that is digitized includes taxonomic, collector, date, geographical and biological data, and is ordered and entered into the database following the international standards. Revisions are made periodically to detect and correct mistakes that could have been committed during the transcription process.

Interns also dedicate time to georeferencing localities, so that specimens can be placed precisely in space. The protocols used for this are a modification of those from MaNIS-HerpNET, with tutorials developed by members of the project. These documents are periodically reviewed to accommodate new necessities that appear during this task. Google Earth is used for drawing polygons containing the desired localities, and the point-radius method is used to calculate the extension and error of the measurements, using the MaNIS georeferencing calculator. The checking of administrative divisions is made using DIVA-GIS.

Communication among all members of this project is actively promoted, as it is seen as an efficient way of enhancing individual capabilities and improving group performance. In this respect, web pages, containing protocols and other useful information, and e-mail groups were developed and are continuously updated.

The Argentine Museum of Natural Sciences holds some of the largest and more diverse biological collections in Argentina, counting more than 1.5 million specimens. Its 21 biological collections are consulted daily by scientists and naturalists from all over the world, whose knowledge contribute to the consolidation of its immeasurable value.

Web pages of the project:

http://www.macn.gov.ar/

http://www.macn.gov.ar/investigacion/proyectos/colecciones/coleccionesfw/pro_colecciones.php

https://sites.google.com/site/coleccionesmacn/

Cuba GBIF Mentoring
supported by GBIF Colombian, Costa Rica and Spanish Nodes

www.gbif.org

...free and open access to biodiversity data

The main objective of this Mentoring is to support Cuban researchers, both biodiversity informatics specialists and curators of biological collections, on the managing, integration, sharing and use of data and metadata, based in the knowledge of the GBIF Colombian, Costa Rica and Spanish Nodes, and so generate the conditions needed to strengthen the Cuban biodiversity information network.
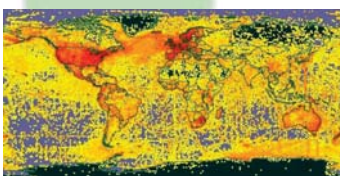
The workshop program followed these main themes:

• **IMPORTANCE, USES AND BENEFITS OF THE EXCHANGE OF PRIMARY BIODIVERSITY DATA. BIODIVERSITY INFORMATION NETWORKS.** Ángela Suarez (Colombia) and Francisco Pando (Spain)
• **MANAGEMENT AND INTEGRATION OF BIODIVERSITY INFORMATION.** Ángela Suarez (Colombia), Francisco Pando (Spain)
• **BIOINFORMATICS.** María Mora (Costa Rica) and Arelys Sotillo (Cuba)
• **PUBLICATION OF DATA AND QUALITY ASSESSMENT / BIODIVERSITY INFORMATICS.** Francisco Pando (Spain) and Mónica l. Vera (Colombia)
• **INTRODUCTION TO THE DOCUMENTATION OF METADATA IN BIOLOGICAL DIVERSITY.** Mónica l. Vera (Colombia)
• **FUNCTIONALITY OF THE CUBAN GBIF NODE CONFIGURATION.** Including:
   1. Presentations on the experiences of Cuban leading institutions in the integration of information on biodiversity, by rapporteurs institutions ANC, JBN, BIOECO, IES
   2. Discussion on lessons learned and priorities, based on the demand of current research and the needs of policy in Cuba - Francisco Cejas (Cuba)
   3. Commitments and agreements between Cuban institutions and the administration of the GBIF Cuba - Francisco Cejas (Cuba)

40 representatives (both curators and biodiversity informatics specialists) from 18 Cuban Institutions related with living collections or observation/ecological monitoring, completed a one week (from 8 to 12 November 2010) Mentoring Workshop on the strategies and tools developed by the GBIF network, and trained in the managing, integration, sharing and use of biodiversity information (specially information from biological collections) and metadata documentation.

Participants are now in capacity to understand and use currents standards, tools, initiatives and networks on the subject, according to the national and international requirements , and declared themselves ready to support the development of the national information infrastructure for collecting, unify, and disseminate data on species, complementing the studies of biodiversity and ecosystems, and try to ensure the participation of their institutions and national biological collections through the Cuban node of the GBIF.

Contact: Dr Francisco Cejas Rodríguez
GBIF Cuban Node Manager
cejas@ecologia.cu

Results of the mentoring have been presented since June 2010:

• Ecology and Systematic Institute, Scientific and Technique Forum: "BIOCUBA, proposal for a Cuban biological diversity Web site", June, 2010
• Workshop on Microbes Collections: Conference on GBIF Cuba, November, 2010
• Day of Science in the Tropical Fruits Research Institute, Conference on GBIF Cuba, Feb - 2011
• Workshop for Latin America Online Species Catalogs: Experiences in the creation of formats and databases for the Cuban Biological Collections, March - 2011
• Environment World Day in the Liofilizados and Hormones Laboratories (LIORAD): Conference on GBIF Cuba, June - 2011

About GBIF

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access.
http://www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use.
http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together mobilise scientific biodiversity data.

GBIF invites you to download species occurrence data freely and openly from http://data.gbif.org

GBIF invites you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.

Finally, in association with Francis Strobbe, from the Royal Belgian Institute of Natural Sciences, the following objectives will be envisaged: To set up and maintain a credible, hierarchically structured, annotated **electronic catalogue of the species living in Cuba**, including references to inter alia publications, databases and consulted experts per taxon, among others. This will allow researchers and all other potential users to find all species recorded in Cuba, classified in a comprehensive consensus taxonomic system.

The Process of Digitising Natural History Collection Specimens at Digitarium

Juha Lehtonen, Susanne Heiska, Mika Pajari, Riitta Tegelberg and Hannu Saarenmaa

Joensuu, Finland, www.digitarium.fi

Digitarium is the digitisation centre of the Finnish Museum of Natural History and the University of Eastern Finland. It aims at industrialising digitisation through an efficient process. Different steps of the workflow can be carried out in remote locations, and by the best available experts.
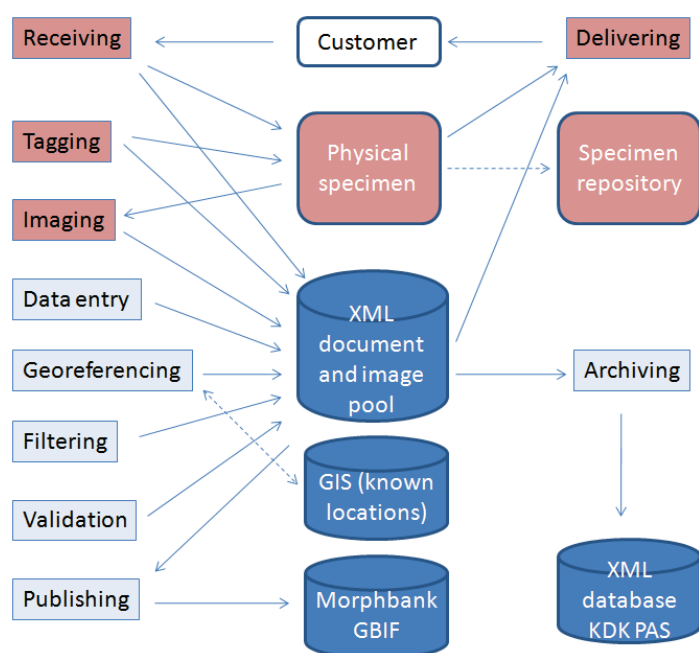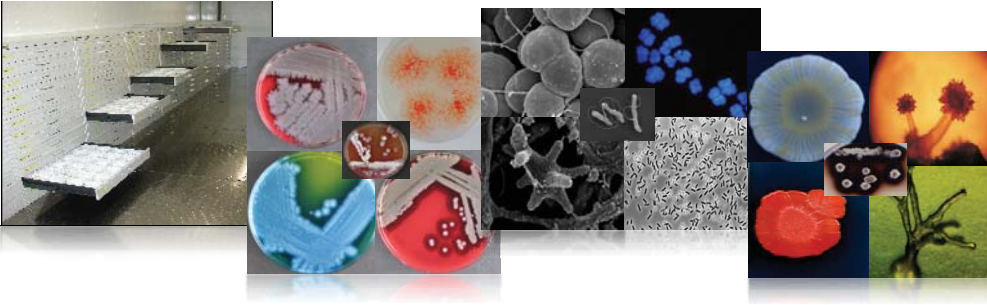

The process is pictured at left. Only the steps in red colour require handling of the physical specimens. All specimens are tagged with a two-dimensional barcode and unique URI (web address), and imaged with cameras at 450 dpi or higher. Data entry is made from the images. There is no collection management database, but all data is handled in versioned XML documents following the Darwin Core standard. The process and workflow is driven by a custom-built digitisation workbench tool (three snapshots below). For publishing the data, the Morphbank image database and GBIF IPT are being used (below left).
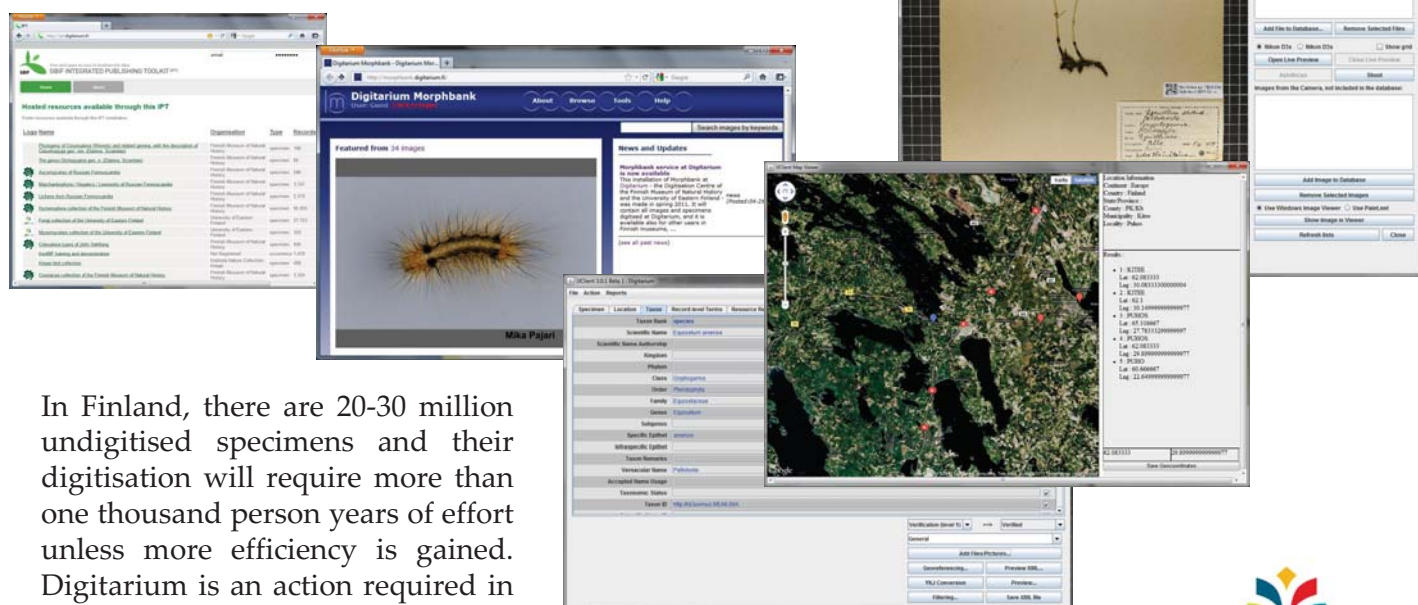
In Finland, there are 20-30 million undigitised specimens and their digitisation will require more than one thousand person years of effort unless more efficiency is gained. Digitarium is an action required in the national digitisation strategy, which is guiding the work.

UNIVERSITY OF EASTERN FINLAND

Open up the treasure chest of prokaryotic diversity for next-generation data analysis

C. Söhngen[1], B. Bunk[1] and J. Overmann[1,*]

[1]German Collection of Microorganisms and Cell Cultures (DSMZ), Braunschweig, Germany,
www.dsmz.de
*corresponding author: joerg.overmann@dsmz.de

GBIF.DE
BACTERIA & ARCHAEA

www.gbif.de/prokaryotes

Unlocking the richness of bacterial biodiversity

Thus usually invisible to the human eye bacteria and archaea are omnipresent in soils, water and even our bodies. Nevertheless, in comparison to their great species variety, the information about their individual functions and interaction with higher taxa and their importance in the ecological system is still poorly understood.

Apart from recording information about taxonomic classification, geographic origin and their availability in culture collections, further data on prokaryotes would help to understand scientific key issues. Centralized and freely available information enables advanced analysis and research regarding issues like biological interactions of bacterial species or bacteria with higher taxa, geographical coincidence or potential metabolic properties taking genomic and metagenomic data into account.

In this accord the GBIF data set of the DSMZ Prokarya Collection Catalogue currently contains 16,638 indexed records belonging to a total of 7,475 species, available via http://data.gbif.org.

About GBIF-D:

GBIF was established in 2001 and Germany is one of the founding nations. Currently GBIF Germany (GBIF-D, www.gbif.de) is funded as a 3-year project by the Federal Ministry of Education and Research (BMBF, 01 LI 1001 A-F).

Eight institutional nodes form the backbone of GBIF-D, each responsible for a range of taxonomic groups:

- BGBM Berlin-Dahlem responsible for Coordination
- DSMZ Braunschweig responsible for Bacteria & Archaea
- BGBM Berlin-Dahlem responsible for Plants & Protists
- BSM Munich responsible for Fungi & Lichens
- MfN Berlin responsible for Insects
- ZSM Munich responsible for Invertebrates II
- Senckenberg Frankfurt responsible for Invertebrates III
- ZFMK Bonn responsible for Vertebrates
- MfN Berlin responsible for Fossils

For any questions concerning GBIF, GBIF-D, access to and provision of data please feel free to contact gbif-d@bgbm.org

Fig. 1. *Corallococcus coralloides*
Fig. 2. *Chondromyces apiculatus*

Two exemplary photos of fruiting bodies formed by myxobacteria. The order of myxobacteria are common in soils and animal dung and have an extraordinary life cycle[1]. Their cells move by gliding over the substrate surface and under conditions of starvation the swarm forms fruiting bodies (Fig. 1.+ 2.) via cell aggregation [1]. The myxobacteria have very large genomes (up to 13 Mbp), compared to other bacteria [2], which indicates multiple potential metabolic activities and their importance as producers of secondary metabolites [1].
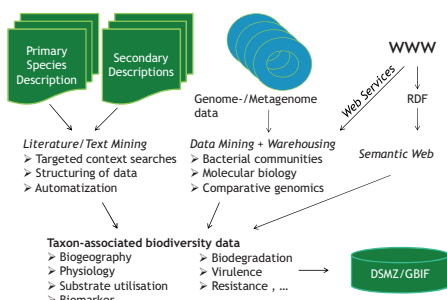

Fig. 3. Sketch of mobilization concept of prokaryotic diversity data with the focus on automatic data extraction.

Digitization and mobilization case example: Myxobacteria

Currently the DSMZ provides 2,800 different myxobacteria cultures for international research and in its function as GBIF data provider offers information about these strains via the GBIF Data Portal (DSMZ Prokarya Collection Catalogue, http://data.gbif.org/datasets/resource/1862).

Besides the cultures themselves, a comprehensive archive of index cards is part of the collection of myxobacteria which contain descriptions on growth habits, appearance of fruiting bodies, details of geographic origin and data regarding the habitat of the strains. Furthermore, the material contains images of these cultures, showing their growth on media and formation of their fruiting bodies. In order to preserve this rich source of taxon associated information it has to be digitized. However, to enable data analysis and comparative studies on this source of information the digitization is only the first step of processing. The data in text format has to be structured and made evaluable by text and data mining methods which are integrated into a data mobilization concept (Fig.3). This subsequent data processing provides the basis to gain insights about cross-species relationships, distinct roles in the ecological patterns, potential sources of drugs or even pathological threats.

About GBIF:

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access.
www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use.
http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together to mobilise scientific biodiversity data.

GBIF invites you to download species occurrence data freely and openly from http://data.gbif.org

GBIF & GBIF-D invite you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.
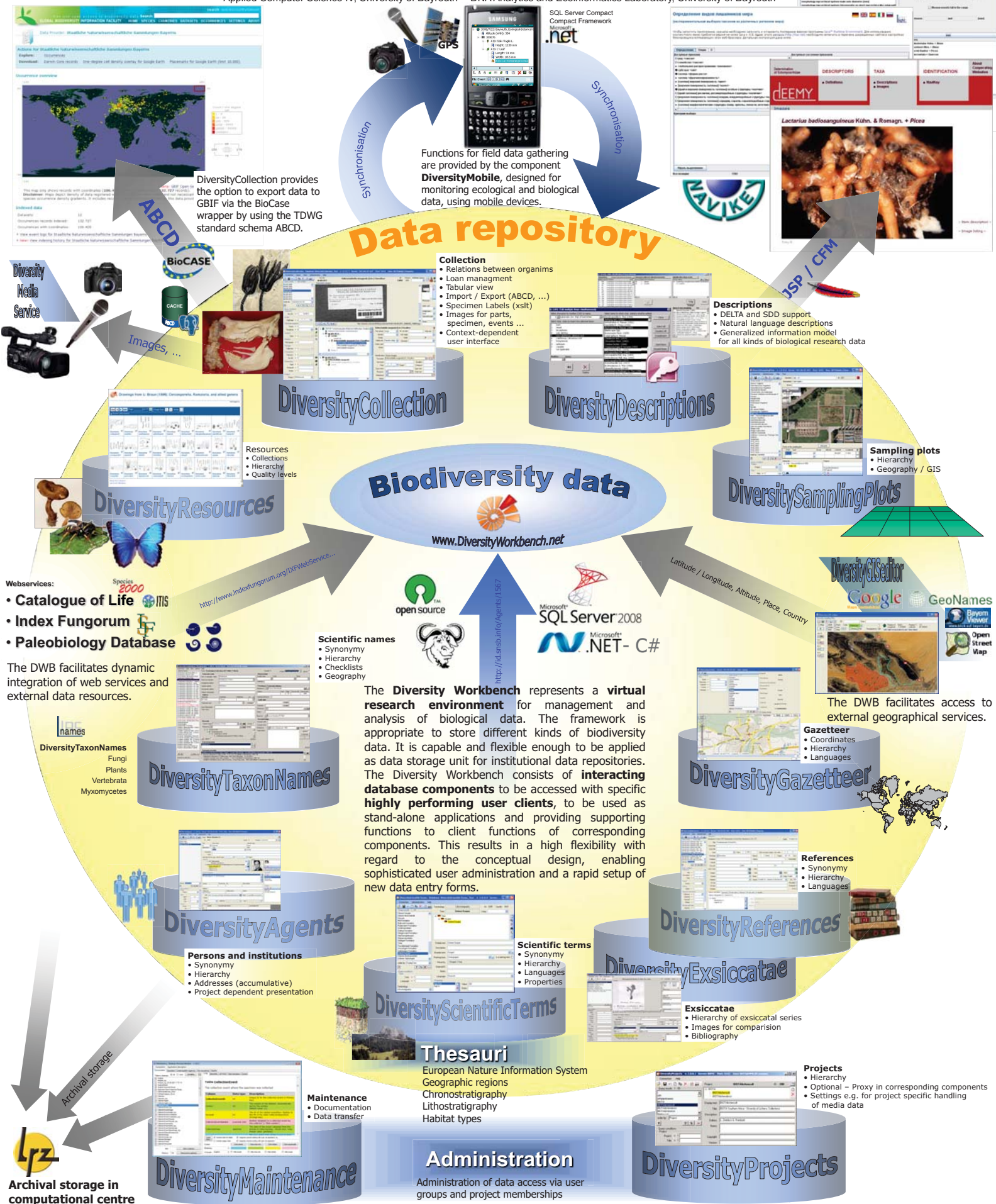
Literature:
[1] Reichenbach, H. Myxobacteria, producers of novel bioactive substances. J. Ind. Microbiol. Biotechnol 27, 149-156 (2001).
[2] Schneiker, S. et. al. Complete genome sequence of the myxobacterium Sorangium cellulosum. Nat. Biotechnol. 25, 1281-1289 (2007).

The Diversity Workbench
a virtual research environment and archive structure for biodiversity data

Dagmar Triebel[1,2], Markus Weiss[1], Tobias Schneider[4], Wolfgang Reichert[1], Gregor Hagedorn[3], Gerhard Rambold[5]

[1] IT-Center of the Bavarian Natural History Collections, Munich   [2] Botanische Staatssammlung, Munich   [3] Julius Kühn Institut, Berlin
[4] Applied Computer Science IV, University of Bayreuth   [5] DNA Analytics and Ecoinformatics Laboratory, University of Bayreuth

SQL Server Compact
Compact Framework
Microsoft
.net

GPS

Synchronisation

Synchronisation

Functions for field data gathering are provided by the component **DiversityMobile**, designed for monitoring ecological and biological data, using mobile devices.

DiversityCollection provides the option to export data to GBIF via the BioCase wrapper by using the TDWG standard schema ABCD.

ABCD

BioCASE

Diversity Media Service

Images, ...

Data repository

Collection
• Relations between organims
• Loan managment
• Tabular view
• Import / Export (ABCD, ...)
• Specimen Labels (xslt)
• Images for parts, specimen, events ...
• Context-dependent user interface

Descriptions
• DELTA and SDD support
• Natural language descriptions
• Generalized information model for all kinds of biological research data

JSP / CFM

DiversityCollection

DiversityDescriptions

Resources
• Collections
• Hierarchy
• Quality levels

DiversityResources

Sampling plots
• Hierarchy
• Geography / GIS

DiversitySamplingPlots

Biodiversity data

www.DiversityWorkbench.net

DiversityGISEditor

Google   GeoNames
Bayern Viewer
Open Street Map

Webservices:
• **Catalogue of Life**   Species 2000   ITIS
• **Index Fungorum**
• **Paleobiology Database**

http://www.indexfungorum.org/IXFWebService...

open source

GNU

http://id.snsb.info/Agents/i.567

Microsoft SQL Server 2008
Microsoft .NET- C#

Latitude / Longitude, Altitude, Place, Country

The DWB facilitates dynamic integration of web services and external data resources.

names

DiversityTaxonNames
Fungi
Plants
Vertebrata
Myxomycetes

Scientific names
• Synonymy
• Hierarchy
• Checklists
• Geography

The **Diversity Workbench** represents a **virtual research environment** for management and analysis of biological data. The framework is appropriate to store different kinds of biodiversity data. It is capable and flexible enough to be applied as data storage unit for institutional data repositories. The Diversity Workbench consists of **interacting database components** to be accessed with specific **highly performing user clients**, to be used as stand-alone applications and providing supporting functions to client functions of corresponding components. This results in a high flexibility with regard to the conceptual design, enabling sophisticated user administration and a rapid setup of new data entry forms.

The DWB facilitates access to external geographical services.

Gazetteer
• Coordinates
• Hierarchy
• Languages

DiversityTaxonNames

DiversityGazetteer

References
• Synonymy
• Hierarchy
• Languages

DiversityAgents

Persons and institutions
• Synonymy
• Hierarchy
• Addresses (accumulative)
• Project dependent presentation

Scientific terms
• Synonymy
• Hierarchy
• Languages
• Properties

DiversityScientificTerms

DiversityReferences

DiversityExsiccatae

Exsiccatae
• Hierarchy of exsiccatal series
• Images for comparision
• Bibliography

Thesauri
European Nature Information System
Geographic regions
Chronostratigraphy
Lithostratigraphy
Habitat types

Projects
• Hierarchy
• Optional – Proxy in corresponding components
• Settings e.g. for project specific handling of media data

Archival storage

Maintenance
• Documentation
• Data transfer

Archival storage in computational centre

DiversityMaintenance

Administration
Administration of data access via user groups and project memberships

DiversityProjects

The GBIF-Germany Node for molluscs, spiders and myriapods

Projects, work stations and data flow

Gerhard Haszprunar, Roland R. Melzer, Jörg Spelda
Zoologische Staatssammlung München
Münchhausenstr. 21,
81247 München
Germany

haszi@zsm.mwn.de
www.gbif.de

Fig. 1. Resources with records hosted by Germany — Occurrence overview: The GBIF-D node invertebrates II contributes with records on Mollusca, Chelicerata and Myriapoda.

Fig. 2. GLOMYRIS, the GLObal MYRiapod Information System provides, e.g., information on type material from the Verhoeff collection, media files on gonopod micropreparations, iconotypes from old literature, and unpublished manuscripts.

Fig. 3. Data retrieval for *Hypoaspis meliponarum* Vitzthum, 1930 (Acari) [from Systax portal].

Figs. 4 and 5. Work stations for the new data mobilisation research program. Microscope unit specialized for extended depth of field photos and box digitalization unit.

The node invertebrates II

Since 2002 the node invertebrates II of GBIF Germany [www.gbif.de/evertebrata2] has been established at the ZSM [www.zsm.mwn.de], representing the Mollusca (molluscs), Chelicerata (including spiders and mites) and Myriapoda (millipedes and centipedes) [www.biologie.uni-ulm.de/cgi-bin/query_all-pl?lang=d&pr=gbif-e2].

Databasing of type material

During the first funding period (German Federal Ministry of Education and Research, BMBF, grant 01LI0205), the node coordinated 7 subprojects databasing the type specimens housed in German zoological collections. In addition, records of non-type material, media files, taxon literature connections etc. have been included in the databases since then.

Edaphobase

Furthermore, the node invertebrates II is a cooperation partner of Edaphobase - GBIF-Database Soil Zoology [www.senckenberg.de/root/index.php?page_id=11082], and contributes georeferenced records of e.g. centipedes and millipedes. Including a geographic information system (GIS), Edaphobase provides cross-links between the species' distributional records and ecological parameters, and makes them accessible via the GBIF-portal.

Innovative data mobilisation

Since 2010 the node Evertebrata II is also funded by BMBF for the joint research project "GBIF-D, centres of excellence of innovative data mobilisation" in cooperation with the SNSB IT center (grant 01LI01001 B). Herewith the node's focus is on innovative imaging techniques, e.g. high resolution and extended field of depth pictures of samples ranging from minute mites or millipede gonopod preparations to the digitization of collection boxes each containing numerous specimens, and the mobilisation of these data. The second focus is on data on non-type material, observational and geo tagged/geo referenced data.

Fig. 6. Data flow diagram for the joint GBIF research program „centres of excellence of innovative data mobilisation".

About GBIF-D:

GBIF was established in 2001 and Germany is one of the founding nations. Currently GBIF Germany (GBIF-D, *www.gbif.de*) is funded as a 3-year project by the Federal Ministry of Education and Research (BMBF 01LI 1001 A-F).

Eight institutional nodes form the backbone of GBIF-D, each responsible for a range of taxonomic groups:

- BGBM Berlin-Dahlem responsible for Coordination
- DSMZ Braunschweig responsible for Bacteria & Archaea
- BGBM Berlin-Dahlem responsible for Plants & Protists
- BSM Munich responsible for Fungi & Lichens
- MfN Berlin responsible for Insects
- ZSM Munich responsible for Invertebrates II
- Senckenberg Frankfurt responsible for Invertebrates III
- ZFMK Bonn responsible for Vertebrates
- MfN Berlin responsible for Fossils

For any questions concerning GBIF, GBIF-D, access to and provision of data please feel free to contact *gbif-d@bgbm.org*

About GBIF:

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access.
www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use.
http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together to mobilise scientific biodiversity data.

GBIF invites you to download species occurrence data freely and openly from http://data.gbif.org

GBIF & GBIF-D invite you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.

Participating institutions

Zoologische Staatssammlung München, Museum für Naturkunde der Humboldt-Universität zu Berlin, Forschungsinstitut und Naturmuseum Senckenberg, Zoologisches Institut und Museum Hamburg, Haus der Natur Cismar, Naturkundemuseum Karlsruhe, Forschungsmuseum König Bonn, Landesanstalt für Umweltschutz Hof, Nationalpark bayerischer Wald.

References

- Berendsohn, W.G., van den Elzen, R., Häuser, C., Haszprunar, G., Jahn, R., Kiessling, W., Stackebrandt, E., Triebel, D. & Türkay, M. 2008: Global Biodiversity Information Facility (GBIF). Aufgaben der einzelnen GBIF-D Knoten und des GBIF-D Leitungsgremiums. In: GfBS News; Gesellschaft für Biologische Systematik. Newsletter 20/2008: 6-8.
- Bohn J, Dunlop J, Friedrichs A, Haszprunar G, Jäger P, Melzer R, Spelda J (2003) TYTEARG – TYpes of TErrestrial ARthropods of Germany. Sustainable use and conservation of biological diversity: a challenge for society. Berlin, 1.-4.12.2003. Symposium Report Part A: 299 [www.ptdlr-klimaundumwelt.de/de/265.php]
- Haszprunar G, Melzer RR (2003) GBIF-Germany, Evertebrata II (Mollusca, Chelicerata, Myriapoda). - a coordinated type programme. Sustainable use and conservation of biological diversity - a challenge for society. Berlin, 1.-4.12.2003. Symposium Report Part A: 298 [www.ptdlr-klimaundumwelt.de/de/265.php]
- Melzer RR, Friedrich S, Ritzerfeld M, Bohn J, Spelda J (2011): GloMyrIS and TyMunAC: Myriapoda and Acari databases of the GBIF-D node invertebrates II. *Spixiana* 34: 11-20.
- Knebelsberger, T., Schwabe, E. & Schrödl, M. 2005. Molluscan types of the Bavarian state collection of Zoology Munich (ZSM) Part 1: Polyplacophora, Scaphopoda, Cephalopoda. Spixiana 28: 97-108.
- Spelda, J., Unsöld, M., Pilz, C. & Melzer, R. R. 2005. The Global Myriapod Information System (GloMyrIS) of GBIF-Germany – an aid for scientific research. XIII International Congress of Myriapodology, Bergen, 25th-29th July 2005. De naturhistoriske samlinger, Bergen Museum, Universitet I Bergen: 48.
- Spelda, J. 2005. The GloMyrIS project of GBIF: Database structure and data exchange. Peckiana 4: 91-100.

copyright information under www.gbif-mycology.de

The GBIF-Germany Node for Fungi and Lichens

Work stations and data flow

BOTANISCHE STAATSSAMMLUNG MÜNCHEN

snsb it center

Federal Ministry of Education and Research

GBIF.DE

Dagmar Triebel, Tanja Weibulat, Dieter Neubacher, Wolfgang Reichert, Markus Weiss
Botanische Staatssammlung München and IT Center of the Bavarian Natural History Collections
Menzinger Straße 67
80638 München
Germany

triebel@bsm.mwn.de
www.gbif.de

We are roughly estimating that more than 20,000 fungal organisms are known from Germany and more than 100,000 to 150,000 different species (more than 5 million specimens) from all continents are hosted in German museums, herbaria and culture collections. Further valuable records are continuously gained in biodiversity and ecological research projects and through monitoring activities ordered e.g. by environmental agencies and national parks. The GBIF-Germany Node for Fungi and Lichens (www.gbif-mycology.de) aims to open this broad spectrum of scientifically, sociopolitically and economically relevant data in the field of mycological biodiversity to the public. The activities of the node are including all biological and taxonomical groups of fungi, myxomycetes (slime moulds) and lichens. The participating institutions are listed below.

Currently, the Mycological Node of GBIF Germany delivers more than 270,000 records of fungi, lichens and slime moulds to the international GBIF network, 200,000 of these are georeferenced (Fig. 1) and 30,000 are accompanied by images. Additionally, a number of information systems and portals with access to biodiversity databases are linked or even run as a GBIF node service.

GBIF-Germany "Fungi & Lichens" is part of the joint research project "GBIF-D, Kompetenzzentren innovativer Datenmobilisierung". Thus, it is setting up work stations for high resolution imaging (Fig. 2) and mobile work stations for efficient data capturing (Fig. 3).

In cooperation with the IT Center of the Bavarian Natural History Collections (www.snsb.info), IT workshops are organized and data storage as well as data processing services are offered. This will ensure an integrated and sustainable data flow to GBIF (see data flow for Diversity Workbench environments below).

Fig. 1. Resources with records hosted by Germany — Occurrence overview: The GBIF-D Fungi & Lichens publishers contribute with more than 200,000 georeferenced records.

Fig. 2. GBIF-D Fungi & Lichens work stations for generating high-resolution images of mycological and lichenological objects.

Fig. 3. GBIF-D Fungi & Lichens mobile work stations for gaining data in the field and the herbarium.

About GBIF-D:

GBIF was established in 2001 and Germany is one of the founding nations. Currently GBIF Germany (GBIF-D, www.gbif.de) is funded as a 3-year project by the Federal Ministry of Education and Research (BMBF, 01LI 1001 A-F).

Eight institutional nodes form the backbone of GBIF-D, each responsible for a range of taxonomic groups:

BGBM Berlin-Dahlem responsible for Coordination
DSMZ Braunschweig responsible for Bacteria & Archaea
BGBM Berlin-Dahlem responsible for Plants & Protists
BSM Munich responsible for Fungi & Lichens
MfN Berlin responsible for Insects
ZSM Munich responsible for Invertebrates II
Senckenberg Frankfurt responsible for Invertebrates III
ZFMK Bonn responsible for Vertebrates
MfN Berlin responsible for Fossils

For any questions concerning GBIF, GBIF-D, access to and provision of data please feel free to contact gbif-d@bgbm.org

About GBIF:

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access.
www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use.
http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together to mobilise scientific biodiversity data.

GBIF invites you to download species occurrence data freely and openly from http://data.gbif.org

GBIF & GBIF-D invite you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.

References

Jablonski, S., Kehl, A., Neubacher, D., Poschlod, P., Rambold, G., Schneider, T., Triebel, D., Volz, B. & Weiss, M. 2009: DiversityMobile — Mobile data retrieval platform for biodiversity research projects. — In Fischer, S., Maehle, E. & Reischuk, R. (eds.), Informatik 2009 — Im Focus das Leben. Beiträge der 39. Jahrestagung der Gesellschaft für Informatik e. V. (GI). GI-Edition: Lecture Notes in Informatics (LNI) — Proceedings 154: 27, 610—624.
Triebel, D., Neubacher, D., Weiss, M., Heindl-Tenhunen, B., Nash. T. H. III & Rambold, G. 2010: Integrated biodiversity data networks for lichenology — data flows and challenges. — In Nash, T.H. III, Geiser, L., McCune, B., Triebel, D., Tomescu, A.M.F. & Sanders, W.B. (eds.), Biology of lichens — symbiosis, ecology, environmental monitoring, systematics and cyber applications. — Biblioth. Lichenol. 105: 47—56
Weiss, M., Hagedorn, G. & Triebel, D. 2010: DiversityCollection information model (version 2.05.10, 16 December 2010).
http://www.diversityworkbench.net/Portal/CollectionModel_v2.05.10

DWB Environments

e.g.

Data flow between components of participating institutions

Regional computation centers

Data flow between components of national and international data provision projects

Participating institutions

Botanische Staatssammlung München, Universität Bayreuth, Julius Kühn-Institut Berlin, Freie Universität Berlin, Unabhängiges Institut für Umweltfragen Berlin, Senckenberg Museum für Naturkunde Görlitz, Universität Greifswald, Ludwig-Maximilians-Universität München, Eberhard Karls Universität Tübingen, Martin-Luther-Universität Halle-Wittenberg, Philipps Universität Marburg, Universität Regensburg

BGBM   DSMZ   ZSM   museum KOENIG   SENCKENBERG world of biodiversity

Out the drawer and on the web
Large scale metadata capture and publication of insect collections

A.S. Kroupa, F. Glöckler, A. Hoffmann, C.L. Häuser

Museum für Naturkunde,
Leibniz-Institute for Research on Evolution and Biodiversity
at the Humboldt University Berlin
Invalidenstr. 43,
10115 Berlin

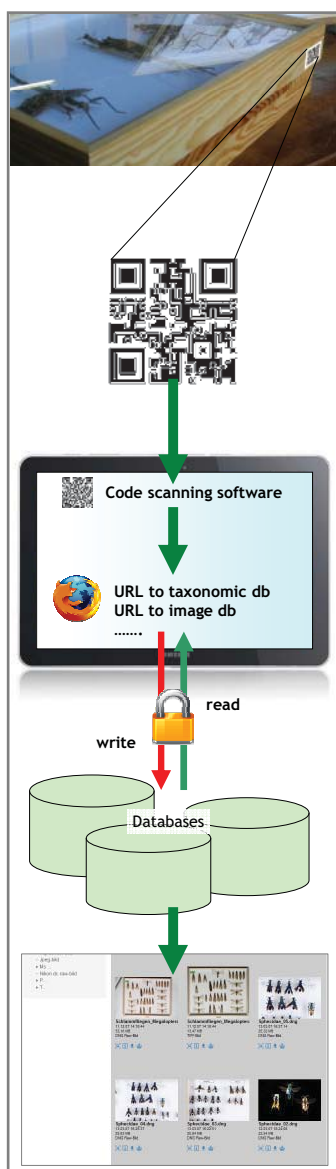alexander.kroupa@mfn-berlin.de
www.gbif.de

Natural history collections

Natural history collections represent a vast repository of biodiversity data of international significance. Although it would be for the scientific community of high value to digitize complete insect collections, the digitization of all specimens is not an achievable aim at present. In order to expose the holdings to new and established users of biodiversity data, the approach taken is a metadata catalogue on the basis of discrete collection units (drawers).

QR-Code (Quick Response)

Each collection drawer will receive a unique QR-Code, encoding an URL with the inventory identifier of the drawer. Scanning the QR-Code with a barcode-reader will redirect the user to a webpage with links to databases with information about the objects. This could be a taxonomic database, supplemented by digital images of the drawer.

Metadata

The essential metadata elements describing each drawer could be the following:
- List of taxa: preferably to the level of genus or family
- List of regions or countries: e.g. biogeographical regions and countries
- Time span of records: e.g. from 1880-1935
- Numbers of specimens
- Type status: numbers of type specimens

Publication on the web

All images of drawers will be published on the web, e.g. on a portal served to GBIF and/or on the website of the respective museum. Consequently, each specialist will have the opportunity to view the images and the accompanying metadata of the drawers before he/she might decide to examine the insect material directly, e.g. for a revision.

Conclusion

With such a digitization approach at a metadata level, an entire large insect collection could be digitized in a reasonable time, even if the collection contains hundreds or thousands of drawers. Using a QR-Code helps to simplify the recording of metadata and accelerate the process of metadata capture.

In some cases images can save the costs of the loans of material, as particular questions could already be answered by examining the images, especially if they are of high resolution/quality. One additional advantage of publishing images of entire collection units (drawers) is to the possibility for remote identification/verification of assorted specimens by a larger user community of specialists.

Fig. 1. Workflow from QR-Code on a drawer to the metadata recording in different databases up to publication on a webportal.

Literature
Berendsohn, W.G., V. Chavan & J.A. Macklin (2010) Recommendations of the GBIF task group on the global strategy and action plan for the mobilization of natural history collections. Biodiversity Informatics, 7, pp. 67-71.
Berendsohn, W.G. & P. Seltmann (2010) Using geographical and taxonomic metadata to set priorities in specimen digitization. Biodiversity Informatics, 7, pp. 120-129.
Berents et al. (2010) Towards demand-driven publishing: Approaches to the priorization of digitization fo natural history collection data. Biodiversity Informatics, 7, pp. 113-119.

Image credits: A.S. Kroupa

About GBIF-D:

GBIF was established in 2001 and Germany is one of the founding nations. Currently GBIF Germany (GBIF-D, www.gbif.de) is funded as a 3-year project by the Federal Ministry of Education and Research (BMBF, 01LI 1001 A-F).

Eight institutional nodes form the backbone of GBIF-D, each responsible for a range of taxonomic groups:

- BGBM Berlin-Dahlem responsible for Coordination
- DSMZ Braunschweig responsible for Bacteria & Archaea
- BGBM Berlin-Dahlem responsible for Plants & Protists
- BSM Munich responsible for Fungi & Lichens
- MfN Berlin responsible for Insects
- ZSM Munich responsible for Invertebrates II
- Senckenberg Frankfurt responsible for Invertebrates III
- ZFMK Bonn responsible for Vertebrates
- MfN Berlin responsible for Fossils

For any questions concerning GBIF, GBIF-D, access to and provision of data please feel free to contact gbif-d@bgbm.org

About GBIF:

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access. www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use. http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together to mobilise scientific biodiversity data.

GBIF invites you to download species occurrence data freely and openly from http://data.gbif.org

GBIF & GBIF-D invite you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.

Towards the second decade

New prospects for the botanical community within GBIF Germany

Sabine von Mering, Wolf-Henning Kusber, Jörg Holetschek, Regine Jahn & Walter G. Berendsohn

Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Germany
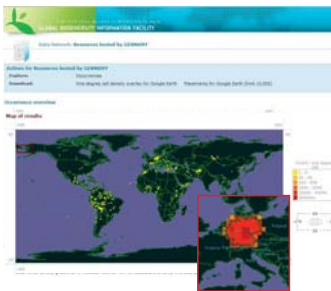
s.vonmering@bgbm.org
www.gbif.de

Fig. 1a. Georeferenced specimens from selected German herbaria; 1b (small inlet): Floristic data from Germany.

Fig. 2. *AlgaTerra*: Example for a Berlin Model database providing data to GBIF, which is being migrated to the CDM platform (developed by the European EDIT project).
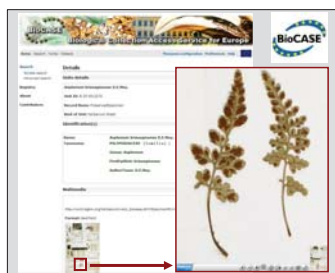
Fig. 3. Herbarium specimen of *Asplenium brissaginense*. Provided by BGBM using the BioCASe protocol and ABCD standard, and published in the European BioCASe portal (http://search.biocase.org). The detail shows the type specimen as displayed by the BGBM Image Server (FSI viewer).

Initially, German GBIF institutions focused on data from natural history collections. Mobilisation and digitisation of herbarium data was one of the major objectives of the botanical projects during the first decade of GBIF and continues to be a key activity in the future. GBIF Germany *Plants, Algae & Protists* currently delivers 5.8 million records to the global GBIF network. In 2011, the number of records increased by 17 %.

Botanical activities in GBIF Germany

The range of datasets is now going to be extended by the inclusion of additional data from herbaria (Figs. 1a), from living collections (botanical gardens and algal cultures), and from field observations (e.g., floristic surveys and vegetation databases; Fig. 1b). Some organism groups with low representation in GBIF - such as algae and protists - are now receiving increased attention. *AlgaTerra* (Jahn & Kusber 2005+, Fig. 2) comprises type information, specimen data as well as images, videos, and molecular data for terrestrial and limnic micro algae. A web-based Specimen Tool was developed in the German project *DNA Bank Network* for GBIF Germany, Plants, Algae and Protists to gather and publish further biodiversity data, e.g., to involve Citizen Science interest groups. With respect to multimedia objects, GBIF Germany will provide the image server technology for a joint German Digital Herbarium and other services for the joint publication of multimedia data (Fig. 3) as well as metadata.

BioCASe as a standard Provider Software in Germany

Most German data provider use the BioCASe protocol (Biological Collection Access Service) and ABCD standard. BioCASe enables unified access to distributed and heterogeneous collection and observational databases using open-source, system-independent software and open data standards and protocols (currently ABCD 2.06). Compared to the DarwinCore protocol, the ABCD standard allows delivering much more detailed data to the GBIF network. Support for ABCD/BioCASe users is provided by the BGBM Berlin.

Relevant related projects

For new collection types, such as DNA and tissue samples, the GBIF network and BioCASe technology are used within the *DNA Bank Network* initiated by the DFG (Gemeinholzer et al. 2009). Taxonomic information systems are increasingly based on the *Internet Platform for Cybertaxonomy* software developed by the European Distributed Institute of Taxonomy (EDIT). Close cooperation exists with the EU-funded project *OpenUp!* which will make culturally-significant multimedia content of natural history collections accessible via the web portal *Europeana*, thus furthering public appreciation and use of biodiversity information.

GBIF Germany *Plants, Algae & Protists* shares the objective of the German node to at least double the amount of data contributed by German GBIF providers until the end of 2013.

Literature
Gemeinholzer B., Droege G., Zetzsche H., Knebelsberger T., Raupach M., Borsch T., Klenk H.-P., Haszprunar G. & Waegele J.-W (eds.) 2009+. DNA Bank Network (continuously updated). DNA Bank Network Webportal. Published on the Internet at http://www.dnabank-network.org; accessed [08.09.2011].

Güntsch A., Kusber W.-H., Döring M., Ciardelli P. & Berendsohn W. G. 2007. Common access to distributed biodiversity information. – Pp. 45-48 in: Kusber W.-H. & Jahn R. (eds.): Proceedings of the 1st Central European Diatom Meeting 2007. – Berlin. doi:10.3372/cediatom.109

Jahn R. & Kusber W.-H. (eds.) 2005+. AlgaTerra Information System [online]. Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin. Published on the Internet at http://www.algaterra.org; accessed [26.09.2011].

About GBIF-D:

GBIF was established in 2001 and Germany is one of the founding nations. Currently GBIF Germany (GBIF-D, www.gbif.de) is funded as a 3-year project by the Federal Ministry of Education and Research (BMBF, 01LI 1001 A-F).

Eight institutional nodes form the backbone of GBIF-D, each responsible for a range of taxonomic groups:

- BGBM Berlin-Dahlem responsible for Coordination
- DSMZ Braunschweig responsible for Bacteria & Archaea
- BGBM Berlin-Dahlem responsible for Plants & Protists
- BSM Munich responsible for Fungi & Lichens
- MfN Berlin responsible for Insects
- ZSM Munich responsible for Invertebrates II
- Senckenberg Frankfurt responsible for Invertebrates III
- ZFMK Bonn responsible for Vertebrates
- MfN Berlin responsible for Fossils

For any questions concerning GBIF, GBIF-D, access to and provision of data please feel free to contact *gbif-d@bgbm.org*

About GBIF:

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access.
www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use.
http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together to mobilise scientific biodiversity data.
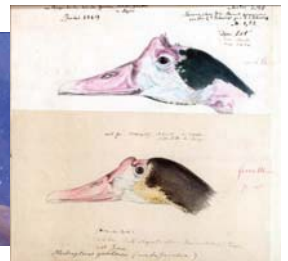
GBIF invites you to download species occurrence data freely and openly from *http://data.gbif.org*

GBIF & GBIF-D invite you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.

Enlarging information on vertebrate collection data

Innovative methods for rapid assessment

Renate van den Elzen, Peter Grobe, Zdravko Colic

Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, D-53113 Bonn
Corresponding author: r.elzen.zfmk@uni-bonn.de

r.elzen.zfmk@uni-bonn.de
www.gbif.de

The Zoological Research Museum Alexander Koenig, Bonn (ZFMK), Germany provides data on vouchers and observations of vertebrates (mammals, birds, fishes, amphibians and reptiles) for use in research, teaching and education as well as for the interested public. Three main goals will be accomplished:

- Acquisition of already digitized occurrence data
- A rapid acquisition of non-digitized vouchers and observations
- Addition of multimedia data to enrich specimen based information

Acquisition of already digitized data
Occurrence data, (specimen based collection and observation data), taxon names, and literature data are dispersed over several institutes and are available in different formats. The ZFMK will transfer all these data into one central repository, the Diversity Workbench Framework *www.diversityworkbench.net*. Tools will be developed to exchange the data with other applications and databases, like Laboratory Management Software, Web Portals, etc. This will ease access to collection data and allow data flow within the ZFMK and to other institutions.
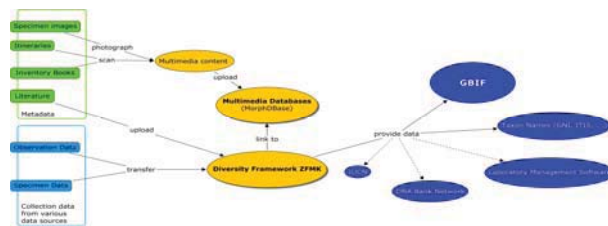
Rapid acquisition of new data
A rapid acquisition of specimen based collection data will be accomplished by scans and photographs: Multiple specimen based items like clutches of eggs in bird collections will be scanned. Amphibians or fishes in wet collections are photographed and presented as one single entity and documented by images. The information content of specimen based collection data will be enlarged by appending multimedia data. These are photographs, sound files and scans of literature, but also of old inventory books and itineraries that hold valuable and detailed information and illustrations of finding localities, landscapes and habitat (Figures 1 – 3).

Outlook
All data will be offered free to all via an institutional web portal as well as via GBIF. Ideally GBIF should provide the possibility to link vouchers, images or observations with complex literature references, old inventory books and/or itineraries.

Fig. 1. Scans of inventories, indexes or itineraries are also instruments for a synopsis of collection contents.

Fig. 2. Scans from the sketchbook of Georg Schweinfurth provide detailed information on the depicted species of Egypt.
A. Common dolphinfish B. Greater Hoopoe-lark with bill anomaly C. Rock Hyrax, *Procavia capensis abyssinicus* D. Striped Sand-snake (top) and Speckeld Sand-snake (bottom).

Fig. 3. The sketchbook of Georg Schweinfurth also gives insight into landscapes and habitats at a time, when photography was not yet a common tool.

Photocredits
Inventories, itineraries etc. of vertebrates are scanned at ZFMK by Thorsten Klug (Figures 1-3). The hardware is open to all GBIF-D partners.

About GBIF-D:

GBIF was established in 2001 and Germany is one of the founding nations. Currently GBIF Germany (GBIF-D, *www.gbif.de*) is funded as a 3-year project by the Federal Ministry of Education and Research (BMBF, 01LI 1001 A-F).

Eight institutional nodes form the backbone of GBIF-D, each responsible for a range of taxonomic groups:

- BGBM Berlin-Dahlem responsible for Coordination
- DSMZ Braunschweig responsible for Bacteria & Archaea
- BGBM Berlin-Dahlem responsible for Plants & Protists
- BSM Munich responsible for Fungi & Lichens
- MfN Berlin responsible for Insects
- ZSM Munich responsible for Invertebrates II
- Senckenberg Frankfurt responsible for Invertebrates III
- ZFMK Bonn responsible for Vertebrates
- MfN Berlin responsible for Fossils

For any questions concerning GBIF, GBIF-D, access to and provision of data please feel free to contact *gbif-d@bgbm.org*

About GBIF:

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access.
www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use.
http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together to mobilise scientific biodiversity data.

GBIF invites you to download species occurrence data freely and openly from http://data.gbif.org

GBIF & GBIF-D invite you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.

DNA Bank Network

How GBIF technologies enable sample referencing and data integration for biological DNA repositories

Gabriele Droege[1], Holger Zetzsche[1], Jonas Astrin[2], Walter Berendsohn[1], Johann-Wolfgang Wägele[2], Gerhard Haszprunar[3], Hans-Peter Klenk[4], Lisa Campbell[5] & Birgit Gemeinholzer[6]

[1] Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Koenigin-Luise Str. 6-8, 14195 Berlin, Germany
[2] Zoologisches Forschungsmuseum Alexander Koenig, Molecular Taxonomy & Biobank, Adenauerallee 160, 53113 Bonn, Germany
[3] Bavarian State Collection of Zoology, Muenchhausenstr. 21, 81247 Munich, Germany
[4] German Collection of Microorganisms and Cell Cultures, Inhoffenstr. 7B, 38124 Braunschweig, Germany
[5] The New York Botanical Garden, 2900 Southern Boulevard, Bronx, NY 10458, U.S.A.
[6] Justus-Liebig-Universität, AG Spezielle Botanik, Heinrich-Buff-Ring 38, 35392 Giessen, Germany

contact@dnabank-network.org
www.dnabank-network.org
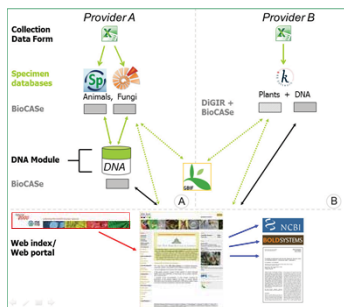
Fig. 1. Principal data architecture of the DNA Bank Network. Specimen and observation data are recorded platform-independent using the Collection Data Form in a standardized manner (right in the field). Data will be imported into specimen databases of individual Network partners. The DNA Module (A) or an alternate system (B) can be implemented for local DNA data management. Specimen data content is structured by wrappers (BioCASe, DiGIR, grey boxes) and transferred to the Network's shared web portal as well as to GBIF (green dotted arrows). DNA data are transferred via wrappers using ABCDDNA. Publications and online accessible DNA sequence data (blue arrows) can be linked to the related DNA sample. The Catalogue of Life checklist is used as a search backbone in the web portal (red arrow).
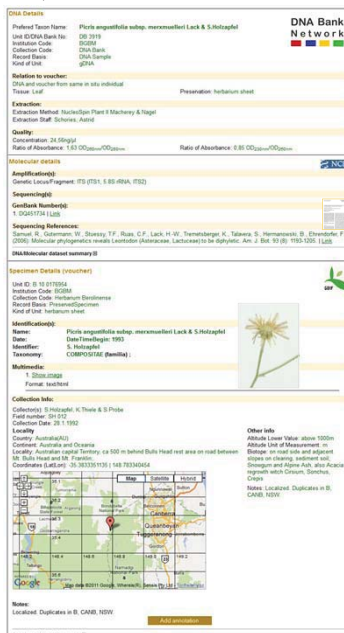
Fig. 2. Synoptic datasheet of a DNA sample. The datasheet is dynamically generated during a web query aggregating data from multiple sources. DNA sample data (on top) and specimen data (at bottom) is loaded from databases hosted and curated by the Network partners or respective data providers. Links to molecular data (center) go to e.g. GenBank, BOLD.

References:
TDWG, DNA Bank Network (2010): ABCDDNA – DNA extension for Access to Biological Collection Data. http://www.tdwg.org/standards/640/

© Photographs A-C: G.Droege
© Photograph D: DSMZ

What about voucher specimens of DNA sequences?

Accessible voucher specimens are the only reliable basis to verify the species identity of molecular sequences published in databases such as GenBank, EMBL, DDBJ, or BOLD. Although the deposition of voucher specimens in public research collections for taxonomic descriptions is routine, an equal diligence in deposition of voucher specimens for molecular based research is the exception rather than the rule.

Scope of the DNA Bank Network

The goal of the DNA Bank Network is to facilitate access to such genetic resources as well as to all related data, including digital images of vouchers. DNA banks focusing on DNA from organismic samples make their accessions available via the Network's shared portal and scientists can search for and requests voucher referenced DNA and tissues.

The Network's data architecture is built on GBIF

All partners of the DNA Bank Network are GBIF data providers. The realized Network's data model is based on GBIF infrastructure and IT tools and works on a three-level system: specimen databases, DNA databases and the index database of the web portal (Fig. 1). Both specimen data and DNA data are administered locally by the project partners in separate databases. Single datasets, namely specimen datasets and their related DNA datasets, refer to each other. Data exchange among databases at all three levels is facilitated by the same wrapper technology used by GBIF.

Open source software for data management

The DNA Module is a powerful tool for curation of DNA data and researcher requests. It can interact with every GBIF compliant database worldwide and enables dynamical references from a DNA data set to its respective specimen data sets.
The Specimen Tool (part of DNA Module) can be used for voucher data that are currently not availabe via GBIF. From 2012 this tool will be additionally hosted by GBIF.DE as a stand-alone portal and input tool for providing specimen and observation data to GBIF.

A standard for exchange of DNA data

Because ABCD and DarwinCore lack DNA specific elements a new DNA extension (ABCDDNA) was developed, which is proposed as a new standard to TDWG (TDWG, DNA BANK NETWORK 2010).

Synoptic data sets for voucher referenced DNA samples

The shared web portal provides information of all available DNA samples provided by the Network partners. Equal to the GBIF portal DNA data and specimen data are indexed in the Network's database. While searching for a DNA sample the Network's web portal sends one query to the relevant locally curated partner DNA database and another to the respective specimen database. Then, all information as well as digital images for the requested DNA sample is dynamically compiled into a virtual synoptic dataset (Fig. 2). Therefore redundant copies can be avoided and verifications can be accomplished by annotations. Thus, the accessible data to a certain DNA sample always represent the current state of knowledge.

About the DNA Bank Network

The main focus of the DNA Bank Network is to enhance taxonomic, systematic, genetic, conservation and evolutionary studies by providing:

• at-cost availability of organismic DNA samples covering the entire range of non-human biodiversity

• high quality, long-term storage of DNA material on which molecular studies have been performed, so that results can be verified, extended, and complemented,

• protocols for specimen and tissue preservation in the field, DNA extractions and data workflows

• Collection Data Form, DNA Module and Specimen Tool as open source software for standardized data recording for GBIF and the DNA Bank Network

• complete on-line documentation of each sample, including the provenance of the original material, the place of voucher deposit, information about DNA quality and extraction methodology, digital images of vouchers and links to published molecular data if available.

Partners of the DNA Bank Network

GBIF Germany (GBIF.DE)

• Botanic Garden and Botanical Museum Berlin-Dahlem

• German Collection of Microorganisms and Cell Cultures, Braunschweig

• Bavarian State Collection of Zoology, Munich

• Zoologisches Forschungsmuseum Alexander Koenig, Bonn

National Biological Information Infrastructure (NBII)

• The New York Botanical Garden*

* Supported in part by the National Science Foundation (DBI-0846412)

Sharing data and samples

At present more than 42,000 DNA and tissue samples are available via the central web portal.

All samples are provided under terms of agreement of the individual bank (MTA).

The Network's Vision

The DNA Bank Network holds the potential to become a shared global platform for voucher referenced DNA samples of non-human biological DNA repositories.

Further DNA banks worldwide are invited to join the DNA Bank Network.

About GBIF

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access.
www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use.
http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together to mobilise scientific biodiversity data.

GBIF invites you to download species occurrence data freely and openly from http://data.gbif.org

GBIF & GBIF-D invite you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.

OPEN UP! Opening Up the Natural History Heritage for Europeana

GBIF Meets the Arts
Natural History in a Cultural Context

W.G. Berendsohn, P. Böttinger, A. Güntsch, A.K. Michel, E. Zippel (2011)
Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Königin-Luise-Str. 6-8, 14195 Berlin, Germany.
Email: w.berendsohn@bgbm.org; OpenUpCoord@bgbm.org

GLOBAL BIODIVERSITY INFORMATION FACILITY

BioCASE

europeana think culture

Figure 1: Tapestry called "Krokus". Britta Rendahl , 1976. © Upplandsmuseet, Uppsala, Sweden. *www.europeana.eu*

Figure 3: Herbarium specimen *Crocus vernus* L. © BGBM, Berlin. *www.europeana.eu*

Figure 4: An east view on Yardley Oak. Robert Pollard, 1805. © The British Library Board – Online Gallery. *www.europeana.eu*

Figure 5: Type specimen of the goliath beetle Goliathus atlas Nickerl, 1887, and its original type labels. © National Museum Prague. *www.europeana.eu*

Project Details

OpenUp! is a 3-year-project (March 2011-February 2014) co-funded by the European Commision and by several European natural history collections (Fig. 7).
OpenUp! will connect the European virtual library Europeana to the providers in the network of the Biological Collection Access Service (BioCASE) in Europe and thus to GBIF.

Figure 2: Bronze Cat Coffin. © The Oriental Museum, University of Durham, Durham, UK. www.europeana.eu

Project Objective

Connecting the cultural and the natural history domains is the central idea behind OpenUp!. Europe's collections cover most of the world's described organisms, ranging from common and famous species to those that have already gone extinct. Some were collected during historical expeditions by well-known epochal explorers and scientists like Darwin or Humboldt. OpenUp! will make multimedia representations of these treasures available to the general public, often for the first time. The complementary taxonomic literature is provided by the Biodiversity Heritage Library (BHL) Europe project.

Project Activities

OpenUp! addresses technical developments (**mapping between different data standards, enrichment of metadata, caching strategy for a single access point to distributed data, etc.**). The project will also **incorporate multilingual metadata, in particular vernacular names of organisms,** for a multicultural audience. OpenUp! will contribute to the next development phase of Europeana by exemplifying the incorporation of an existing provision framework for distributed content (**BioCASE**).

The basic OpenUp! dataflows are established using software components originally developed in the EC-funded BioCASE project and its predecessors, which are now widely used in GBIF and other networks. The **BioCASE provider software** helps to hide internal data structures of OpenUp! participants and forms a consistent service-based information space for the multimedia content. Use of the **TDWG standard ABCD** (Access to Biological Collection Data) ensures provision of a broad spectrum of specimen data, which in turn provide the metadata linked up to Europeana using the **OAI-PMH interface definition**.

OpenUp! implements data quality measures to achieve a high level of completeness and correctness of metadata associated with multimedia objects. A dedicated **"data quality toolkit"** is being developed, which integrates a number of quality services including scientific name catalogues. The toolkit consults existing OpenUp! provider installations and compiles an XML-based report with quality problems found. The same quality services will also be used to enrich meta information with information derived from existing collection data. Figure 6 outlines the projected procedures.

Background

Europeana is building a cross-domain portal to Europe's cultural and scientific heritage. Up to now it provides access to **15 million digital images, texts, sound files, and videos** with a focus on cultural history (Fig. 1-5). Multimedia objects belonging to the natural history domain are still dramatically underrepresented, although being clearly within the scope of Europeana as a part of Europe's scientific heritage.

OpenUp! closes this gap by serving multimedia content to the BioCASE network and making this compliant and accessible to Europeana. Content served includes specimen images, movies, and animal sound files with a reference to an observation or collection event. Participants committed to serve at least 1.1 million objects by the end of the project. **All providers to BioCASE are automatically also serving GBIF.**
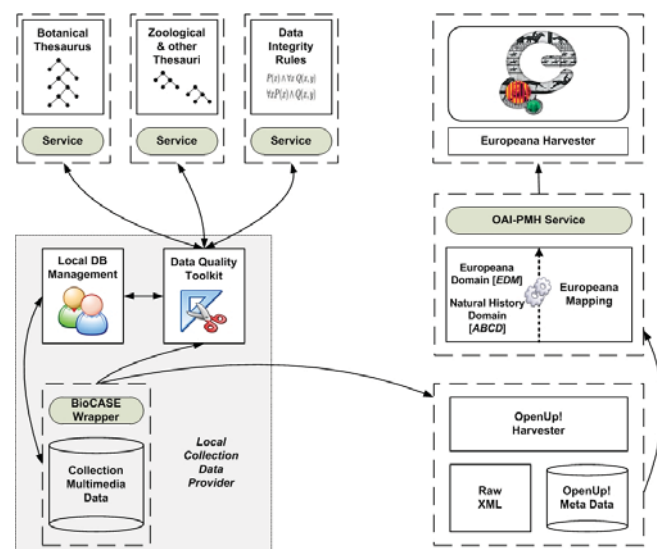

Figure 6: Data processing steps in the mobilisation of multimedia content from natural history institutions for Europeana

Project Participants

OpenUp! is an initiative of CETAF, the Consortium of European Taxonomic Facilities, together with several European GBIF Nodes.

URLs:
BioCASE – Biological Collection Access Service – *www.biocase.org*
BHL-Europe – Biodiversity Heritage Library Europe – *www.bhl-europe.eu*
CETAF – Consortium of European Taxonomic Facilities – *www.cetaf.org*
Europeana – *www.europeana.eu/portal/aboutus.html*
GBIF – Global Biodiversity Information Facility – *www.gbif.org*
OpenUp! – Opening up Europe's natural history heritage for Europeana *www.open-up.eu*

**Find more information at *www.open-up.eu*
or contact the OpenUp! project coordination: *openupcoord@bgbm.org***

Figure 7: The OpenUp! consortium consists of 23 partners from 12 European countries

GBIF-D enables free access to Biodiversity Data: How to contribute?

M. Gleisberg[1]*, Ch. Häuser[2], R. Jahn[1], W. Kiessling[2], R. Melzer[3], J. Overmann[4],
D. Triebel[5], M. Türkay[6], R. van den Elzen[7] & W.G. Berendsohn[1]

[1]Botanic Garden and Botanical Museum Berlin-Dahlem, Berlin; [2]Museum fuer Naturkunde, Berlin;
[3]Zoologische Staatssammlung, Munich; [4]German Collection of Microorganisms and Cell Cultures
Braunschweig; [5]Botanische Staatssammlung Munich; [6]Senckenberg, Frankfurt;
[7]Zoologisches Forschungsmuseum Alexander Koenig, Bonn
*corresponding author: m.gleisberg@bgbm.org

gbif-d@bgbm.org
www.gbif.de

SPONSORED BY THE
Federal Ministry of Education and Research

The GBIF network includes primary biodiversity data (specimens as well as observations) from a large variety of data sets.

Fig. 1. The Occurence Overview-Map generated by the GBIF Data Portal is showing the geographical distribution of primary biodiversity data hosted by German institutions as of September 2011 http://data.gbif.org

Fig. 2. Oecophora bractella: Observation record in the GBIF portal http://data.gbif.org

Based on the occurrence data available in the GBIF Network the niche modelling algorithm implemented in the GBIF portal shows the predicted distribution for Mycena pterigena (Fig. 3).
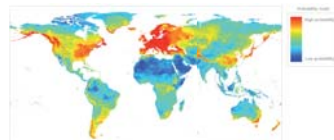
Fig. 3. Niche modelling for the fungus Mycena pterigena in the GBIF portal http://data.gbif.org

Photo credits:
Vertebrate fauna, sound files prepared to be available via BioCASe provider software at www.europeana.eu (K.-H. Frommolt, MfN Berlin).

Fieldwork during the excursion, here observation of migrating birds, Ussuria/Russian Far East (M.Gleisberg, BGBM Berlin-Dahlem).

Oecophora bractella, see observation record in Fig. 2 (D. Hobern, www.gbif.org).

Herbarium digitization at B. New GBIF-D hardware installed at the BGBM Berlin-Dahlem for all GBIF-D partners (W.-H. Kusber, BGBM Berlin-Dahlem).

Mictyris longicarpus, specimen from the Rueppell-Collection (M. Türkay, Senckenberg Frankfurt).

Mycena pterigena (Ch. Lange, www.gbif.org).

What is GBIF?

The Global Biodiversity Information Facility (GBIF) is an international initiative of 57 countries and 47 organizations. GBIF has the overall mission of mobilizing the world's primary biodiversity data via the Internet. The GBIF network currently provides access to more than 302 million records. This includes specimens as well as observation data from a large variety of data sets.

Which role does GBIF-D play within the global network?

Main objective of GBIF-D is the mobilization of suitable data from Germany's research community, focussing on "species occurrence data", i. e. specimens and observations of organisms. In addition, GBIF-D offers expertise on technical aspects of data capture and database networking. Stable workflows from data entry to open and free online access are established.

One of the technologies used in the global network was developed in Germany within the EU-funded BioCASE (Biological Collection Access Service for Europe) and SYNTHESYS projects, and so was the international data exchange standard ABCD (Access to Biological Collection Data).

How can you contribute to GBIF?

Organizations, institutions and individuals holding collection and observation data are invited to share their primary biodiversity data within the GBIF network. Databases can be connected by using provider software like BioCASE or can join one of the established databases, networks and repositories such as SESAM, OBIS, PANGAEA, DiversityWorkbench, Specify, etc. Be a part of the biggest biodiversity information community worldwide!

What is your benefit of GBIF & GBIF-D?

GBIF-D offers expertise on technical aspects of data capture and database networking. Stable workflows to support the provider from data entry to open and free online access are established. GBIF is far more than the mobilization of natural history collections data. We encourage all scientists present to integrate their data into the GBIF network and to continue using GBIF data for their own work, in order to benefit from the global effort GBIF represents.

GBIF is the only functional initiative directly targeting biodiversity research and researchers that operates on a global scale. It offers the chance to connect national and international efforts in the field in a synergistic manner. With tools like the GBIF-Community-Site it is easy to keep up-to-date on what is happening with regard to the availability of biodiversity data and the state of biodiversity informatics tools.

Outlook:

Managing the sharing of global biodiversity data is a complex task: collaboration is absolutely essential to tackle the growing environmental crises. So do not hesitate to share your data and join the network.

About GBIF-D:

GBIF was established in 2001 and Germany is one of the founding nations. Currently GBIF Germany (GBIF-D, www.gbif.de) is funded as a 3-year project by the Federal Ministry of Education and Research (BMBF, 01LI 1001 A-F).

Eight institutional nodes form the backbone of GBIF-D, each responsible for a range of taxonomic groups:

- BGBM Berlin-Dahlem responsible for Coordination
- DSMZ Braunschweig responsible for Bacteria & Archaea
- BGBM Berlin-Dahlem responsible for Plants & Protists
- BSM Munich responsible for Fungi & Lichens
- MfN Berlin responsible for Insects
- ZSM Munich responsible for Invertebrates II
- Senckenberg Frankfurt responsible for Invertebrates III
- ZFMK Bonn responsible for Vertebrates
- MfN Berlin responsible for Fossils

For any questions concerning GBIF, GBIF-D, access to and provision of data please feel free to contact gbif-d@bgbm.org

About GBIF

GBIF makes digital biodiversity data openly and freely available on the Internet for everyone, and endorses both open source software and open data access. www.gbif.org

GBIF provides scientific biodiversity data for decision-making, research endeavours and public use. http://data.gbif.org

GBIF is a network of data publishers who retain ownership and control of the data they share. Linked datasets provide a more robust representation of biodiversity than any single dataset.

GBIF provides access to primary biodiversity data held in institutions in developed and developing countries. Data shared through GBIF are repatriated data.

GBIF is a dynamic, growing partnership of countries, organisations, institutions and individuals working together to mobilise scientific biodiversity data.

GBIF invites you to download species occurrence data freely and openly from http://data.gbif.org

GBIF & GBIF-D invite you to join the GBIF network and share your biodiversity data, as well as participate in developing new tools and services.

GBIF.DE

SENCKENBERG world of biodiversity

Mobilising Migratory Species Data –
integrating Global Register of Migratory Species and GBIF data

Klaus Riede (ZFMK) & Francisco Rilla (CMS)

Zoological Research Museum Alexander Koenig (ZFMK), Adenauerallee 160, 53113 Bonn
UNEP/CMS Secretariat, United Nations Premises,Hermann-Ehlers-Str. 10, 53113 Bonn, Germany
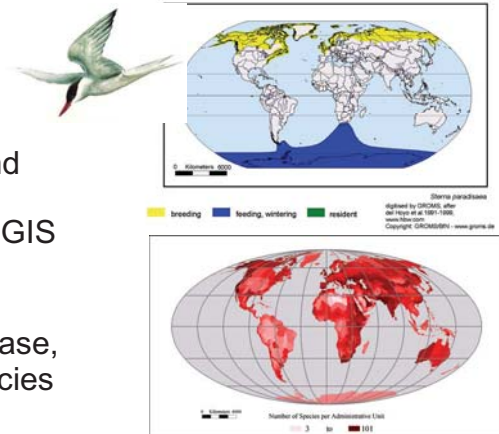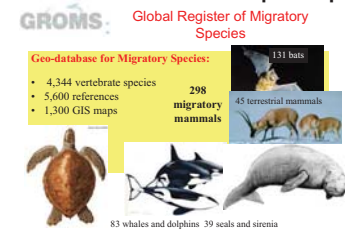
The number of migratory animals can only be estimated at present (between 5,000 and 10,000 species).
Therefore, in 1998 the UNEP/CMS Secretariat initiated the development of the Global Register of Migratory Species (GROMS - www.groms.de) (funded by The German Federal Environment Ministry, in cooperation with the Zoological Research Museum Alexander Koenig in Bonn, Germany).

As a main result, the GROMS (www.groms.de)

- identified and categorised 4,430 species of migratory vertebrates, based on
- 5,600 bibliographic references.
- 1,300 GIS shape maps on global scale, coding winter- and breeding areas. This dataset facilitated a
- GIS analysis of global migrant diversity (Fig. 1) and easy GIS „calculation" of range states by intersection with national boundaries.
- Results were compiled within an MS-Acces/mySQL database, including IUCN Red List 2000 threat assessment): species not listed by CMS, but threatened!

Travellers without passport

Global Register of Migratory Species

Geo-database for Migratory Species:
- 4,344 vertebrate species
- 5,600 references
- 1,300 GIS maps

298 migratory mammals

131 bats
45 terrestrial mammals
83 whales and dolphins 39 seals and sirenia

breeding   feeding, wintering   resident

- Elevated diversity of migrants in temperate zones
- Migrants will be lost if conservation is limited to Hot Spots

As a follow-up, the Convention on Migratory Species (CMS) signed a Memorandum of Cooperation (MoC) with the Global Biodiversity Information Facility (GBIF) at Bonn, 30 October 2008, to work together to develop and share biodiversity data on migratory species, to facilitate the access to the GROMS database (http://www.groms.de/) in combination with GBIF primary data.

A flash application available at http://groms.gbif.org integrates GROMS data and polygons with points from the the GBIF data portal (programmed by Javier de la Torre, http://www.vizzuality.com/. The above examples illustrate a higher density of GBIF observation and specimen points in Europe, but lack of data in the rest of the Asian/African ranges of species (e.g. *Sterna albifrons*), according to the GROMS global shape files (mainly based on del Hoyo et al., 1992-2001).
However, the present application does not allow direct access/query of GBIF points or other GIS operations, such as intersection or overlay with environmental GIS layers.
Together with on-doing data repositories on animal tracking and tagging data, such as movebank (http://www.movebank.org/), such integrated GIS data and functionalities could improve the availabilty and usability of data on migratory species considerably.

References: de la Torre J: http://www.vizzuality.com/projects/gromsizzuality
del Hoyo J Elliott A & Sargatal J (eds) (1992-2001), Handbook of the birds of the world
Kays R, Wikelski M, Fountain T & S Tilak (200/) Movebank project http://www.movebank.org/
Riede K (2004) Global Register of Migratory Species - from Global to Regional Scales. Final Report of the R&D-Projekt 808 05 081.330 pages + CD-ROM, ISBN 3-7843-3845-3

BioVeL
Developing a Biodiversity Virtual e-Laboratory
An innovative way to accelerate biodiversity science

MAIN OBJECTIVES

Research on biodiversity mobilises a range of disciplines (taxonomy, ecology, genetics, soil science, geography, climate science, geochemistry, eco-informatics, etc.). **Each scientific community has its own data, tools, vocabulary, and concepts. To understand the dynamics of biodiversity and ecosystem services in a rapidly changing environment, it is urgent to improve data sharing and networking of experts.** Up-to-date methodologies using e-platforms and agile work approaches contribute to this goal.

Biodiversity Virtual e-Laboratory (BioVel) is an FP7 European project that addresses these issues with:

- Provision of key services for sharing of workflows and better data analysis, with unified access: taxonomy, phylogenetics, metagenomics, ecological niche and population modelling, and ecosystem functioning and valuation (Fig. 1)
- User-friendly interfaces with a catalogue of services provided
- An e-Science support centre & training activities

DATA
- Global Species Databases
- Distribution
- Field Survey
- Climate
- Autecology
- Substrate
- Landuse
- Dispersal / Migration
- DNA

SERVICE SETS
- Taxonomy
- Phylogenetics
- Metagenomics
- Ecological Niche Modelling
- Ecological Population Modelling
- Ecosystem Functioning and valuaton

SUPPORTING SERVICES SET
- Geospatial mapping
- Shim services

APPLICATION AREAS
- Carbon Sequestration
- Ecosystem
- Invasive Species

Figure 1 : Relationship between data, workflows and application

ACTIVITIES

Networking – To strengthen the research community across disciplines

- Elaborate the content of workflow packs and service sets for chosen topical science areas
- Create consensus in and across the communities of interest

Services – To build and deploy e-Science environment for the biodiversity research community

- Provision and operation, within the myGRID environment of biodiversity related services
- Tailoring the myGRID environment for biodiversity research

Joint research activities – To improve ease of use of workflows with software adapted to the specific needs of users

- For biodiversity workflow developers
- For biodiversity workflow consumers

EXPECTED IMPACTS

The greatest benefit for the European research community

By developing e-Science technology for biodiversity, European countries will directly benefit from knowledge and expertise assembled in this e-Infrastructure

An easy way to combine complex data sets from different disciplines

BioVel will provide a generic concept for successfully linking climatic, remote sensing, geographic, or ecological information with the spatio-temporal occurrence of taxa, genotypes or habitats. It will establish links with key bodies as GBIF and bring GBIF web services into work flows.

A major input to the LifeWatch ESFRI Research Infrastructure and GEO BON

BioVel will be a key contribution towards implementation of these initiatives by enabling better sharing of skills and data, and faster production of outputs in Biodiversity science

BioVel and policy makers

Decision-making depends on knowledge-based scientific assessments to respond to the threats of biodiversity loss and degradation of ecosystem services. BioVel will help towards building a dynamic science-policy interface.

Towards a Taxonomic Backbone for Europe

Yde de Jong*, Nihat Aktac, Ward Appeltans, Javier Atalah, Walter Berendsohn, Louis Boumans, Thierry Bourgoin, Mark Costello, Henrik Enghoff, Marc Geoffroy, Werner Greuter, Anton Guentsch, Michael Guiry, Francisco Hernandez, Charles Hussey, Roger Hyam, Juliana Kouwenberg, Alessandro Minelli, Roisin Nash, David Ouvrard, Bart Vanhoorne, Eckhard von Raab Straube.

Presenting author: yjong@uva.nl

What?

Integrating EU taxonomic e-infrastructures

Both regional and global, expert-edited species databases have emerged in Europe:

- Fauna Europaea
- Euro+Med PlantBase
- European Register of Marine Species
- World Register of Marine Species
- Index Fungorum
- AlgaeBase and other Global Species Databases that contribute to Species 2000 and its Catalogue of Life.

BUT, the content, supporting technical infrastructure, governance and management, and plans for their long-term sustainability vary significantly.
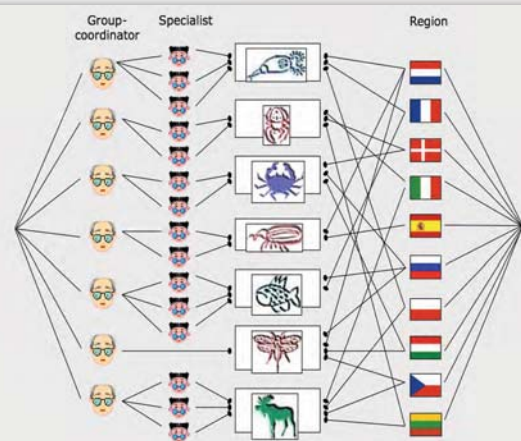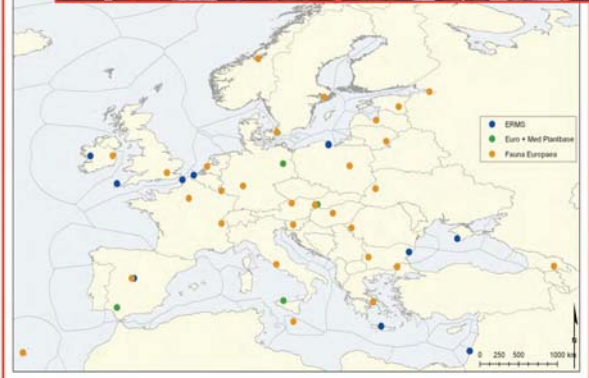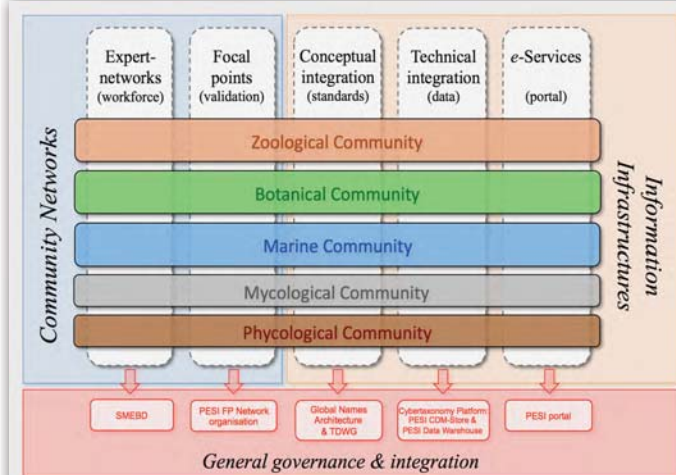
What next?

The *Pan-European Species-directories Infrastructure* (PESI) project (www.eu-nomen.eu/pesi) is integrating both European expertise and information networks and infrastructures …

✓ **upgrading databases content and infrastructures** to make them more compatible in terms of content, enable data integration using data standards, and create proper information infrastructures.

✓ **building the network of experts** in species taxonomy across Europe and Palaearctic, updating and controlling the quality and taxonomy of species inventories (**www.editexpertnet.org**).

✓ **establishing long-term governance** of the databases through the scientific *Society for the Management of Electronic Biodiversity Data* (SMEBD) (**www.smebd.eu**).

✓ **create a common data-portal** to service a wide range of users, providing the first synonymised inventory of all species in Europe (**www.eu-nomen.eu/portal**).

✓ **contribute to global advancements on taxonomic indexing** supporting the efforts on establishing a *Global Names Architecture* as a common taxonomic reference service coordinated by GBIF (**www.gbif.org/informatics/name-services**).

✓ **develop strategies** for securing the long-term sustainability of these networks and infrastructures as part of the LifeWatch construction plan (**www.lifewatch.eu**).

National BioPortal:
The getaway for Tanzania Biodiversity Data and Information

TanBIF

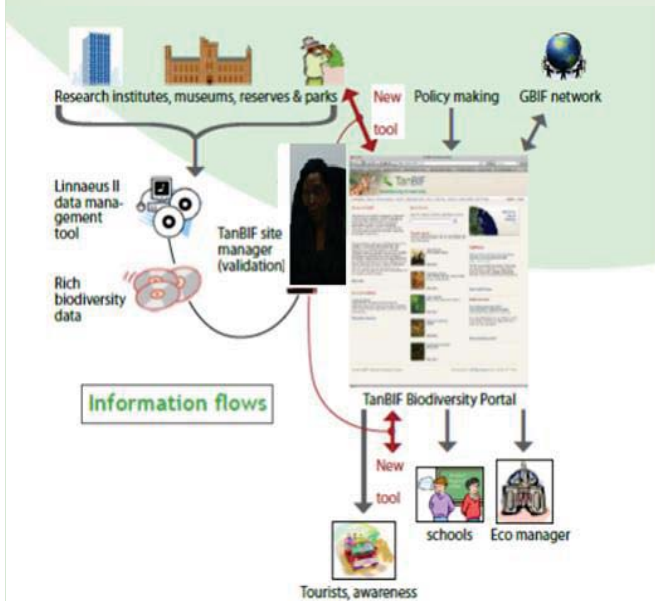*National Facility of Biodiversity Data and Information*

Figure 1: TanBIF BioPortal Functional Structure

- The National BioPortal is designed to meet Tanzanian biodiversity data and information needs
- It has a web-interface to increase visibility and accessibility
- It includes the Linnaeus II Species Bank compiler for publishing and accessing Rich Species Data
- Integrated Publishing Toolkit (IPT) has been incorporated for publishing and accessing Primary Biodiversity Data
- Greenstone Information Management System has been incorporated for sharing and accessing publications on Tanzania Biodiversity
- National BioPortal harvests data from local and international sources; and uses the Catalogue of Life as a core validated taxonomic index to match searches to connected data sources.

Currently more than 170 Rich Species Data, 309,900 Primary Biodiversity Data and 1,000 Tanzania notes and Records are accessible on the Portal
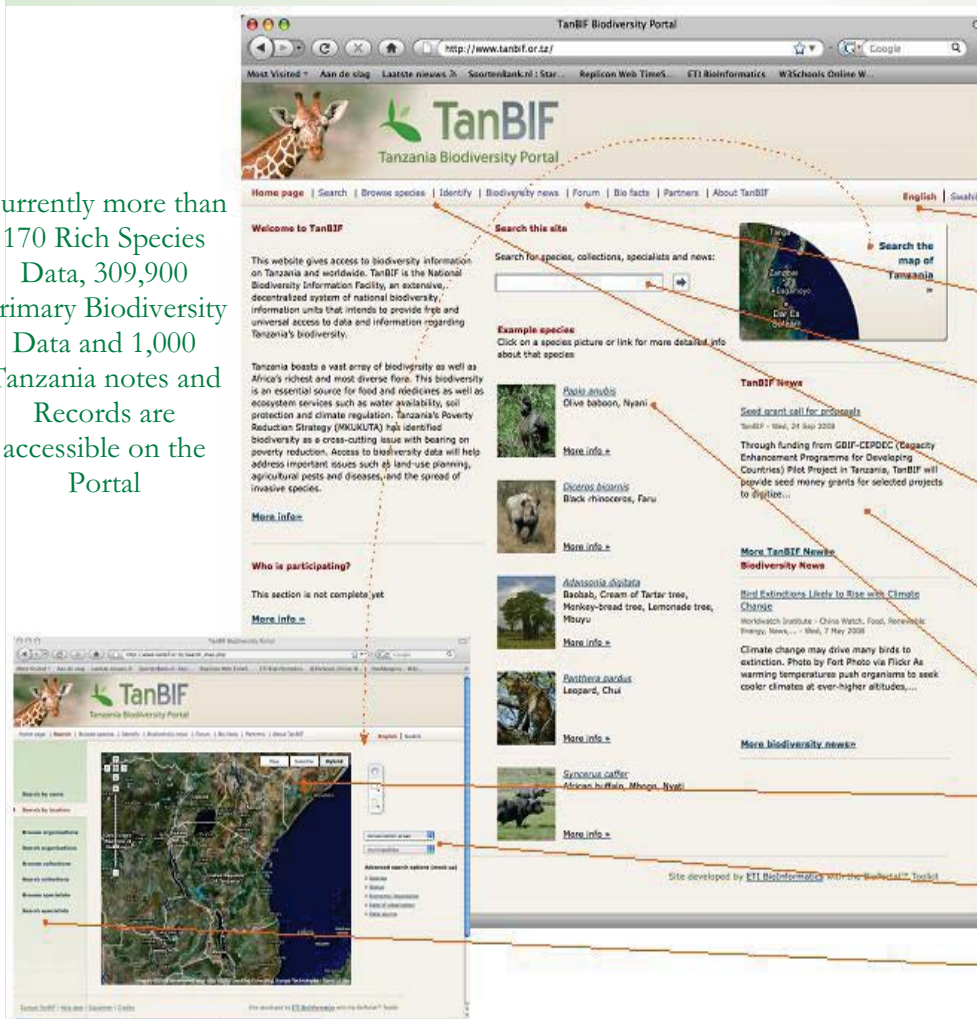
1. Language choice button (currently English and Swahili). The portal can accommodate other languages.

2. The forum provides a virtual meeting space for those working on and interested in Tanzanian biodiversity. It is managed by the site manager.

3. The search box gives 'Google-like' access to all connected information sources. Species names, names of experts or collections, institutes, botanic gardens and zoos, parks and nature reserves can be found. The portal provides an overview with web links, and displays location.

4. Browse species / identify. The TanBIF portal includes a national speciesbank holding detailed information on species (scientific name, synonyms, common names, taxonomic and ecological description, illustrations) as well as computer-aided identification keys.

5. The column 'biodiversity news' harvests continuously updated news from RSS sources in Tanzania and abroad. News items point to the source websites.

6. The 'species panel' displays thumbnails of random species derived from information sources that are connected to the NLBIF portal. The images are linked to the connected speciesbanks.

7. The interactive map that supports geographic searches uses GoogleEarth and GoogleMaps. After selecting an area with the mouse, information of that region is displayed.

8. With the 'select box' parks, nature reserves and municipalities can be selected and displayed on the map.

9. The 'select' menubar gives access to NCD metadata which holds information on research institutions, national parks, biological collections and specialists in Tanzania. There are links to the websites of the institutes and locations are displayed on the map.

TanBIF-GIS Tool
For Ecological Niche Modelling, Prediction and Assessment of Biodiversity

H. Gideon[1], P. Nyinondi[1], G. Oyema[1], M. Bakari[1], H. Fikiri [1] and T. Sutton[2]

1. Tanzania Biodiversity Information Facility Tanzania Commission for Science and Technology

2. Linfiniti Consulting CC, Cape Town, South Africa

Introduction

TanBIF-GIS Tool is software that enables analysis and modelling of primary biodiversity data to support (biodiversity related) decision-making activities such:

Land use planning, implementation of biodiversity conservation and sustainable use strategies, design of protected areas and biodiversity risk assessment among others.

Perform meaningful analysis by integrating: biodiversity data including occurrence data, species-level data/ecosystem data, and other types of data such as geospatial, climatic, demographic, economic datasets.

Basic online GIS Operations

TanBIF-GIS Tool has sets of devices that enable mining, visualization, and combine various geospatial layers (such as climate, topography, vegetation, land use etc.), economical, and primary biodiversity data from various online sources, such as GBIF Darwin Core, Spatial Data Library and others. User can perform simple operations, for instance estimating the extent of overlapping between two or more layers.
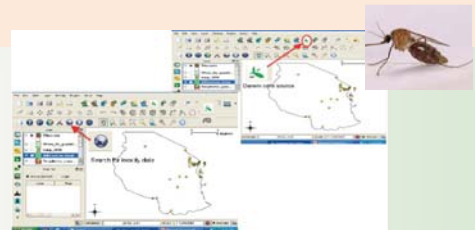

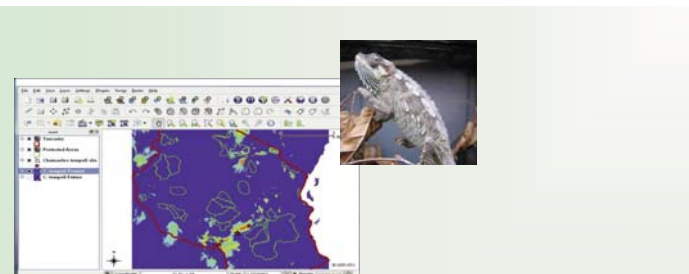*Figure 1: Mining data from GBIF Darwin Core sources and mapping Anopheles gambiae distribution in Tanzania*

Figure 2: Modeling the potential Tanzanian distribution for Chameleo tempeli created using the openModeller plugin. The model algorithm used was 'environmental distance' and the environmental variables used were average annual precipitation, annual minimum and maximum rainfall.

Ecological Niche Modelling

TanBIF-GIS Tool is an integration of openModeller Ecological Niche Modelling Software into Quantum GIS and thus offers access to the most common niche modelling algorithms plugins, including GARP, Climate Space Model, Bioclimatic Envelopes, Support Vector Machines, Environmental Distance and others. User can plug in species occurrence data and environmental layers on which the model should be based, creating a fundamental niche model and projecting the model into an environmental scenario.

Predicting Change

The dynamic change prediction relies on historical transitions and models specified for the future. TanBIF-GIS tool offers the options to modify biotic and/or abiotic parameters to identify actual or potential changes in biodiversity. It allows export of the project results in formats interoperable with the World Mapping Services and Open GIS Consortium Standards
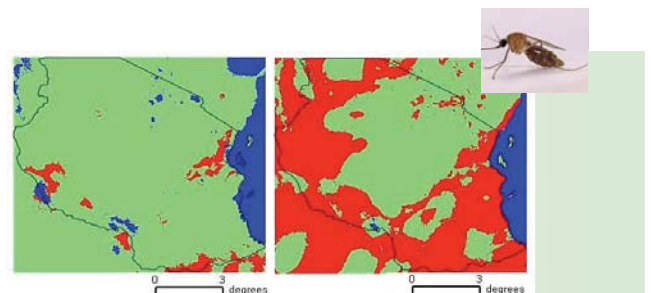

Figure 3: Modeling the potential Tanzanian distribution for Anopheles gambiae (left map is for year 2010 and right map is prediction for year 2050) . Environmental variables used were average annual precipitation, minimum and maximum Temperature.

An Online Biodiversity Informatics Training Curriculum:
Marshalling the Collective Knowledge of the GBIF Community

A. Townsend Peterson and Jorge Soberón, Biodiversity Institute, University of Kansas

ONLINE NOW AT:
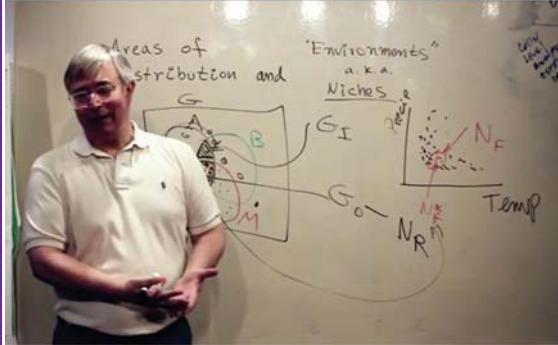http://ornithology.biodiversity.ku.edu/biodiversity-informatics-training

Biodiversity Informatics – The Field
Information on biodiversity has become increasingly digital and increasingly available in recent decades. These advances have taken place thanks to technological improvements, as well as to sociological shifts that have made it more acceptable and desirable that data be shared openly and efficiently. GBIF has played an important role over the past decade in enabling this efficient access, such that now over 300M primary biodiversity records are openly available to researchers around the world. With such rich stores of information available, however, a premium has been placed on insightful, effective, and appropriate methods for creating new information, managing existing information, and understanding the importance and implications of this information … thus was born the field of biodiversity informatics. The field, however, is young, and has seen neither broad syntheses of methodologies nor effective guides for how to carry out key processes.

The Online Biodiversity Informatics Training Curriculum
In view of the rapid development of a new field, but without a forum for training with current methods, we have initiated an online facility for sharing expertise. The site (URL given above) provide a a broad educational curriculum that will cover major topics in the field of biodiversity informatics (see Thematic Summary at right; additional themes can be added without trouble). The goal is to have a lesson on each topic presented by at least one expert in the field, providing users with access to lectures and presentations by those who have developed or used extensively the tools or challenges in question. This curriculum is still under development, so only a few themes have lessons built so far, but as it builds and improves, and as methodologies evolve, presentations can be revised, or other experts can present alternative methods.

What Makes Up A Training Module?
Each training module has three components: (1) a digital video of a lecture or demonstration by the expert that is posted on YouTube and thereby openly available to all, (2) links to key literature (note: whenever possible, we assure that this literature is openly accessible, though such is not always possible, given recent trends toward commercialization of the scientific literature), and (3) sample data sets and worked examples. These three elements can be mixed and matched quite flexibly to reflect the particular needs of a given thematic module.
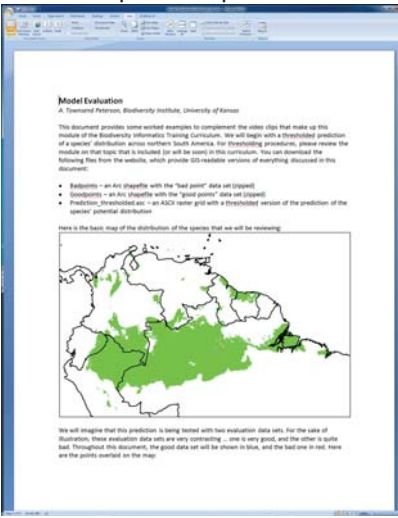
YouTube Page

Relevant Literature

Worked Examples and Sample Data

THEMATIC SUMMARY

Data Capture
• Choosing and designing a database management system
• Designing a data model
• Imaging specimens - 2-dimensions
• Imaging specimens - 3-dimensional
• Imaging specimens - microscopic
• OCR, natural language processing, other automated approaches
• Manual data capture (by taxon/preservation)
• Data capture for observation and survey data
• Citizen science automated data capture
• Capturing tagging/tracking data

Data Quality
• Georeferencing
• Taxonomic validation
• Error flagging (detection versus resolution)
• Data quality assessment
• Metadata (record-level versus dataset-level)
• Repatriating improvements
• Data annotation from users

Data Publishing
• IPT
• Darwin Core
• Data publishing networks and network architecture
• Persistent identifiers - GUIDs and DOIs

Analysis and Interpretation
• Ecological niche modeling and the BAM diagram ✓
• Niche modeling –data preparation
• Niche modeling - theory
• Niche modeling - Maxent
• Niche modeling - model validation ✓
• Niche modeling - applications to invasive species
• Niche modeling - applications to climate change
• Niche modeling - applications to biodiversity discovery
• Measuring biodiversity inventory completeness ✓
• Place prioritization and hotspots
• "Biodiversimetrics"
• Survey gap analysis
• Species inventories and endemicity

NOTE: Training modules that are complete are indicated by ✓

Assuring High-quality Contributions
As with any effort in academia, an important aspect is quality control. As a consequence, we have created a means of assuring effective peer review for each module in the training curriculum presented herein. As part of the online journal *Biodiversity Informatics* (http://journals.ku.edu/index.php/jbi/), we have created a special section of the journal for "Biodiversity Informatics Teaching Modules" … in effect, each module in the curriculum will be submitted to the journal with the elements described above, and will see peer review according to the normal academic standards. These submissions, once reviewed and edited, then have a greater associated level of confidence: rather than just being something posted on the Internet, they have seen review and approval by colleagues who can judge scientific merit.

How to Contribute
All members of the GBIF and broader biodiversity informatics communities are strongly encouraged to participate in this initiative, as the idea is to capture the expertise of the entire community. An "expert" is someone who has special experience with a given technique, and who can communicate that knowledge to others effectively through digital video media. Potential contributors should contact Town Peterson (town@ku.edu) and Jorge Soberón (jsoberon@ku.edu) for guidance, as the apparatus for building this curriculum are only now being set up. In general, though, contributors should plan to assemble the three elements described above for submission in digital formats to the journal for peer review.

GBIF-enabled Knowledge of Birds of the World:
A Challenge to Fill Gaps and Complete Knowledge

A. Townsend Peterson, Laura Russell, and Jorge Soberón, Biodiversity Institute, University of Kansas
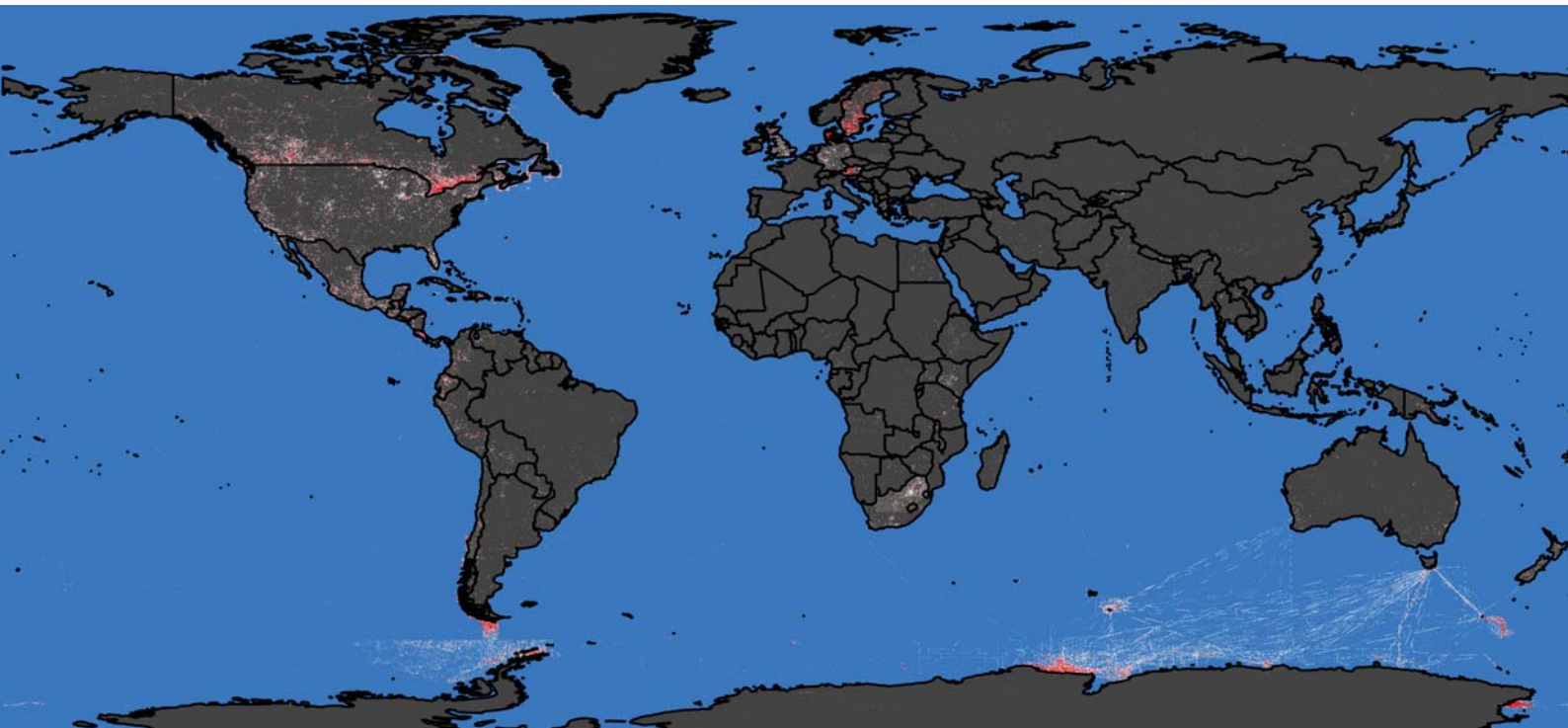
The World Birds Project
Birds represent the best-known major taxonomic group in the world, with >99% of essentially all local avifaunas on Earth known to science. Approximately 12,000 species of birds are currently recognized. Given this well-known taxonomy, and given that birds are the object of the attention of millions of bird-watchers, birds offer a very special opportunity to assess one taxon on global scales.

This project aims to provide an overview of current knowledge of bird diversity and distribution globally—the ultimate objective is to stimulate digitization and integration of bird data from the gaps in current knowledge, so as to complete the picture. Such a data set would be enormously useful in planning for bird conservation, guiding future ornithological sampling efforts, and enriching biogeographic analyses.

The GBIF data portal provides access currently to data records documenting tens of millions of primary bird occurrences—3,396,913 specimens and 39,689,026 observational records. These data, however, has never been analyzed in terms of knowledge of world avifaunas. "Completeness" is based on species accumulation in the inventory as a function of numbers of records; see box below.

This poster provides a first glance at the project's results. Above is a map summarizing completeness values worldside, at a spatial resolution of 10' (cells of about 19 km on a side at the Equator)—in that map and others on this poster, gray cells have some information in the GBIF-enabled data realm; red cells are well-known in terms of their avifaunas (i.e., records for >50% of species likely to occur are available to the scientific community).

Northwestern South America: Knowledge of birds is shown in relation to topography (green = lowlands, brown = mountains); note areas of the Andes that have seen very little survey, and that the Amazon of Colombia remains poorly documented. Note the excellent information available for the birds of Costa Rica and for the Panama Canal Zone, but little information otherwise.
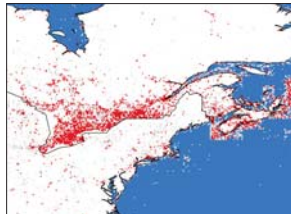

Northeastern North America: This view of eastern Canada and the United States shows the difference between the two countries in integration of data sets—the US has not yet integrated its Breeding Bird Survey and Christmas Bird Count data
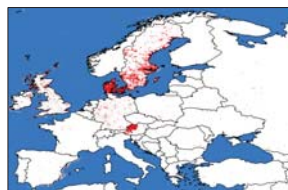

Europe: Thorough knowledge of Denmark, Sweden, Austria, England, Germany; spotty otherwise


South Africa: Thanks to extensive data sharing by the South African National Biodiversity Institute, very thorough knowledge is available for bird distributions across the country; neighboring countriesare not as well documented.
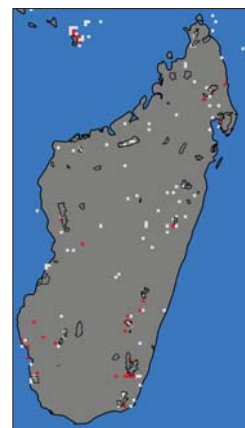
Madagascar: Overall sparsely known in terms of avifauna, Madagascar's avifauna moreover is known largely from records in or near protected areas (here shown as black outlines).
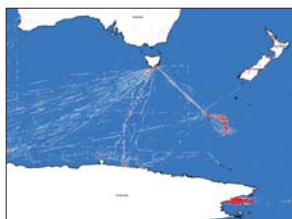
Australian South Seas: The Australian Antarctic Data Centre has made >150,000 records of seabirds available via GBIF, creating an excellent picture of bird distributions in the region. Note the lines radiating out from Tasmania, which reflect ship sampling routes.
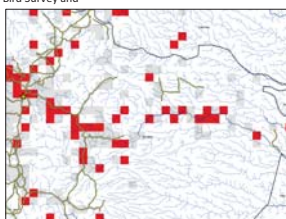
Northeastern Equador: In this view, the relationship between good characterization of avifaunas and access to sites can be appreciated. The yellow line segments indicate roads; blue lines indicate rivers. The well-known sites are aligned closely to these means of access.

New Guinea: Note overall sparse knowledge, but that Papua New Guinea (eastern half) has many more well-known cells than Indonesian New Guinea (western half); this difference reflects less reporting of data existing from Indonesia, as well as greater difficulty of access.

Calculating Completeness

For any region, in this case a 10' cell of latitude and longitude, a certain number of bird species will have been recorded, ranging from 0 up to hundreds. This quantity is termed S_{obs}, or the observed number of species known to occur there. The frequencies with which these species occur can be very informative about how well known the avifauna of that region is—a very well known avifauna will be documented by multiple records per species. Hence, we focus on species that have very few records--Anne Chao, a statistician at the National Tsing Hua University, Taiwan, developed important tools to this end. One of the simplest and most robust of her equations uses the frequency of species that have been detected exactly twice (here denoted F_1) and the frequency of species that have been detected exactly twice (here denoted F_2). From S_{obs}, F_1, and F_2, we can calculate an expected number of species S_{exp} for the region as $S_{exp} = S_{obs} + F_1^2 / 2F_2$. Finally, comparing known and expected numbers of species, we can calculate an estimate of inventory completeness $C = S_{obs}/S_{exp}$.

Key Role of Data Quality
The original (raw) data set held 10,914 unique names for bird species. However, only 8035 of these names were shared with any one taxonomic authority. That is, 26.4% of names in the GBIF-mediated data archive are not reliably associated with a particular reference taxonomy, creating considerable confusion. We undertook the lengthy labor of "fixing" these non-standard names … after 2 months of work, we had a complete lexicon permitting full translation of the GBIF data names set into two standard world bird taxonomies (Sibley-Monroe, Morony-Bock-Farrand), with records for 8228 species overall.

This cleanup of the bird data had important implications for the results of our analyses of inventory completeness. In the graph at right, the vertical gray bar indicates no change in C, which was the case for 9534 cells. However, in hundreds of cases, the clean taxonomic data yielded higher C values (bars to the right of the gray bar), and only in a few cases did C values decrease (bars to the left of the gray bar). Hence, data quality is an important concern in serving biodiversity data to the scientific community—without careful attention to the details of making data "fit for use," the products based on analysis of those data can be compromised seriously.
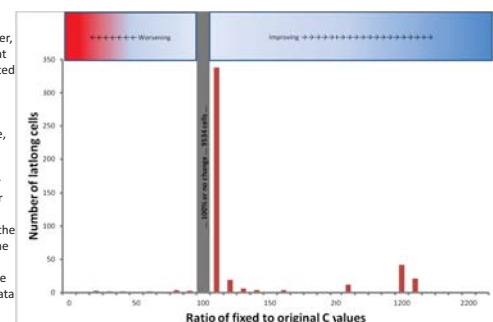
Where is this Project Headed? We anticipate further exploration of this data resource, with the eventual goal of developing metrics of regional inventory completeness, as well as of data "gappiness." We then plan to challenge the global ornithological/birding communities to digitize and share their data strategically so as to complete the global vision of bird distributions and diversity worldwide. Stay tuned!

InBIF

Ministry of Environment and Forests · **Indian Biodiversity Information Facility**
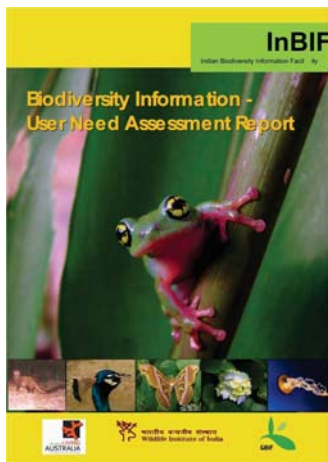
GBIF
www.gbif.org

Wildlife Institute of India – Atlas of Living Australia Mentoring Project

Karthikeyan Vasudevan[1], Omkar Dhavale [1], Gautam Talukdar [1], John Tann [2], Donald Hobern [2],, V. B. Mathur [1], Hem Pandey [3]
1 Wildlife Institute of India, Dehradun, India
2 Atlas of Living Australia, Canberra, Australia
3 Ministry of Environment and Forests, Government of India, New Delhi, India

InBIF — Indian Biodiversity Information Facility — Biodiversity Information – User Need Assessment Report

InBIF — Indian Biodiversity Information Facility — Biodiversity Information Action Plan

Proceedings at the Workshop on "Strategies for Digitisation and Mobilization of Natural History Collections Data"

• <u>User Needs Assessment</u> – The needs and expectations of users were identified and information on users and data providers were compiled and analyzed.

> **Outcome** - Priority areas for InBIF were identified based on the needs of the users; the ways of sharing in biodiversity information were understood

• <u>Interaction with Atlas of Living Australia</u> – Visits by professionals from both organisations took place.

> **Outcome** - Designed and executed User Need Survey; built confidence and professionalism in our approach to launching InBIF

• <u>InBIF Action Plan</u> – Preparing a road map for biodiversity informatics for 5 years for the country with a network of partners
> **Outcome** - A document that provides guidelines on dissemination of biodiversity through requisite infrastructure, processes and capacity building

• <u>Capacity Building</u> – Workshop on "Strategies for Digitisation and Mobilization of Natural History Collections Data" with inputs from ALA; Student dissertation project on digitization and archival of primary biodiversity data was initiated.

> **Outcome** - Increased level of competence in personnel; mobilization of primary Biodiversity data through student projects

• <u>Sensitization and Team Building</u> – Identifying key personnel and provided them training at the InBIF coordinating node; motivating students to take up projects on data discovery.

> **Outcome** - Personnel were trained in the latest software toolkits allowing the submission of data papers; primary Biodiversity Data mobilized through student dissertation

भारतीय वन्यजीव संस्थान
Wildlife Institute of India

Established in 1982, WII is an internationally acclaimed Institution, which offers training program, academic courses and advisory in wildlife research and management. The Institute is actively engaged in research across the breadth of the country on biodiversity related issues.

ATLAS OF LIVING AUSTRALIA
sharing biodiversity knowledge

WII-ALA Mentoring project is a joint initiative to build a national database of our flora and fauna. The project brings together a huge array of information on Australia's biodiversity, accessible through a single web portal.

Partners in this collaborative project include CSIRO, museums, herbaria, other biological collections, the Australian Government and the community.

MOBILIZING BIODIVERSITY DATA THROUGH DIGITIZATION OF SPECIMEN COLLECTIONS : A CASE STUDY ON THE HERPETOLOGICAL COLLECTION AT THE WILDLIFE INSTITUTE OF INDIA MUSEUM
(Dissertation Report)