



# Life & Literature Conference

Chicago, IL, USA, Nov. 14-15, 2011

## Poster Session Guide

---

### 1) No Specimen Left Behind: Industrial Scale Digitisation of Natural History Collections

Authors: Vince Smith, Cybertaxonomist, Natural History Museum  
Vladimir Blagoderov, Manager of the Sackler Biodiversity Imaging Lab, Natural History Museum

Abstract: Digitisation of biodiversity literature and of specimens is crucial to mobilising our accumulated biodiversity knowledge and a cornerstone of Biodiversity Informatics research. Collection digitisation projects usually deal with isolated parts of collections based on taxonomy or geography or other criteria. Specimen digitisation still lags far behind literature digitisation in terms of process rates and workflows. Detailed imaging and capturing of associated metadata of individual specimens is enormously time-consuming. Industrial approach is needed to advance digitisation of three-dimensional natural history collections which includes:

- Determining significant parts of a collection with uniformly mounted specimens (e.g. drawers of insects, slides);
- Imaging and annotating en masse;
- Software and hardware automation (conveyors, OCR, automatic ROI recognition);
- Assigning uIDs/DOIs;
- Web publication of collection images to allow crowdsourcing metadata enhancement.

### 2) Romantic Natural History: Literature and Science in the Century Before Darwin's *Origin* (1859)

Author: Ashton Nichols, Walter E. Beach '56 Distinguished Chair in Sustainability Studies, Dickinson College

Abstract: This poster describes a website entitled Romantic Natural History. We often assume that Charles Darwin announced a new era in the scientific understanding of the natural world with the publication of his *Origin of Species* (1859). In fact, Darwin's theory was the culmination of decades of speculation about connections between human beings and nonhuman "nature." These ideas reflect not only in the work of natural scientists, philosophers, and theologians, but also the ideas of poets, novelists, and visual artists. Romantic Natural History surveys and organizes texts, images, and scholarship linking Romanticism and natural history in the period 1750-1859.

### 3) Online synergy: Sherborn's *Index Animalium* and the Biodiversity Heritage Library.

Authors: Leslie Overstreet, Smithsonian Institution Libraries  
Grace Costantino, Smithsonian Institution Libraries

Abstract: The collection-based science of taxonomy provides internationally recognized names for biological taxa (primarily genera and species) and creates the necessary foundation for many applied sciences. These names, whether currently accepted or in synonymy, have been published in the scientific literature since the mid-18th century, and finding their original appearance – to verify the taxon described, for example – can be almost as hard as finding a needle in a haystack. In zoology, C.D. Sherborn's *Index Animalium* (1902-1933) solves this problem; it provides the original source for every genus name and species epithet in the zoological literature from the 10th edition of Linnaeus's *Systema Naturae* in 1758 (the official start-date of binomial nomenclature in zoology) to 1850. The Smithsonian Institution Libraries' online version of the *Index Animalium* allows researchers to search the entire multi-volume work by name, epithet, or other keyword. With the citation thus provided, researchers can then access the cited text itself, scanned in full, on the website of the Biodiversity Heritage Library (BHL). BHL is an international consortium of natural-history institutions, supported by Internet Archive, dedicated to making the historical literature in the natural sciences freely available on the Internet. To date, tens of thousands of titles have been mounted on the site, and the work continues.

### 4) The Biodiversity Heritage Library for Europe

Authors: Henning Scholz, Museum für Naturkunde, Leibniz Institute for Research on Evolution and Biodiversity at the Humboldt University, Berlin

Graham Higley, The Natural History Museum

Jana Hoffmann, Museum für Naturkunde, Leibniz Institute for Research on Evolution and Biodiversity at the Humboldt University, Berlin

Abstract: BHL-Europe is mobilising and preserving digital European biodiversity literature and facilitating the open access to this literature through a multilingual BHL-Europe portal, the Global Reference Index to Biodiversity (GRIB), and Europeana. The BHL-Europe portal will not only be multilingual but incorporate functionalities not currently available in BHL-Classical. It will, for example, facilitate the search for common and scientific names of biological organisms as well as person names through the implementation of webservices (e.g. Catalogue of Life, VIAF). In order to serve a broader audience, the literature available in BHL-Europe is also accessible by Europeana, Europe's digital library, archive and museum.

## 5) The GRIB – A Union Catalogue of and Scanning Management Tool for Natural History Libraries

Authors: Boris Jacob, Scientific Coordinator, Royal Museum for Central Africa / Tervuren  
Melita BIRTHÄLMER, BHL-Europe WP2 Leader, Museum für Naturkunde / Berlin, Germany

Abstract: The Global References Index to Biodiversity (GRIB) is a union catalogue of European natural history libraries. It contains deduplicated records from BHL-Europe and BHL partner libraries and also serves as a management tool to support the digitisation workflow in these libraries. For each of the bibliographic items the GRIB holds information on its digitisation status. This can either be 1) not digitized yet, 2) nominated to be digitized by a Scientist, or 3) intended to be digitized by a librarian. If it is 4) already digitized and accessible in electronic form, then the GRIB links to the full text.

## 6) Zea Books: Digital Imprint of the University of Nebraska-Lincoln Libraries

Authors: Sue Ann Gardner, Scholarly Communications Librarian, University of Nebraska-Lincoln  
Paul Royster, Scholarly Communications Coordinator, University of Nebraska-Lincoln

Zea Books is the open-access digital works imprint founded by the University of Nebraska-Lincoln Libraries in 2010. Intended to complement, not compete with, the University of Nebraska Press, it gives a voice to scholars whose works would not meet the financial publication demands of a traditional press. Not limited to Nebraska authors, titles to date include *De bestiis marinis, or, The Beasts of the Sea* (Steller), the *Dictionary of Invertebrate Zoology* (Maggenti, Maggenti and Gardner), and *A Nebraska Bird-Finding Guide* (Johnsgard). Operations are overseen by the publisher, Paul Royster, and executive editor, Sue Ann Gardner. An adjunct board of advisers includes the Director of the University of Nebraska Press and UNL faculty.

## 7) EJT - A European Journal of Taxonomy

Authors: Laurence Bénichou, Editorial Manager Scientific Publishing, Muséum National d'Histoire Naturelle  
Steven Dessein, Editor *Plant Ecology and Evolution*, National Botanic Garden of Belgium  
Isabelle Gerard, Head of Publications Service, Royal Museum for Central Africa  
Graham Higley, Head of NHM Library & Information Services, Natural History Museum, London  
Koen Martens, Editor-in-Chief HYDROBIOLOGIA and EJT Royal Belgian Institute for Natural Sciences

Abstract: Thousands of scholarly papers in natural history are published each year, but many taxonomic papers are published in small journals and do not get a high visibility. An investigation among the partners of the European consortium EDIT taught us that within this digital era, many taxonomic journals are faced with complex strategic and technical questions: visibility, access, format, and funding, issues difficult to tackle by individual institutions.

A group of natural history institutions wanted to break from this trend and launched a collectively owned, online journal in taxonomy under the name EJT ([www.europeanjournaloftaxonomy.eu](http://www.europeanjournaloftaxonomy.eu)). It takes there is a clear need for natural history institutions to act as public publishers/producers of taxonomic information.

What taxonomy needs is a journal that offers a communication channel for descriptive taxonomic work in botany, zoology and palaeontology, using the latest online scholarly standards and services.

Our poster describes this journal offering a long term, public business model that is favourable to the unique scientific environment of natural history science, taking into account the long shelf life of taxonomic papers and the correct use of nomenclature rules. This model guarantees open access without cost to authors.

## 8) Integration of African Solitary Bee Biodiversity Information

Authors: Willem Coetzer, South African Institute for Aquatic Biodiversity  
Connal Eardley, Plant Protection Research Institute  
Janine Kelly, Plant Protection Research Institute

Abstract: This project will build on another project, funded by SABIF in 2010, in which about 500,000 specimen records from three South African museums were cleaned and migrated to Specify6. We are trying to develop capacity for (Specify-based) biodiversity information management in South Africa and Africa.

One of the main objectives of the current JRS-funded project is to make available online the *Catalogue of Afrotropical Bees* (Eardley and Urban, 2010). This catalogue lists 2,755 valid bee names and 6,989 invalid bee names in 26,671 citations of 1,229 literature references. The catalogue has already been imported into Specify6. The catalogue is interesting because it represents a 30-year effort to tag legacy biodiversity literature semantically, on the theme of bees in Africa, even though the authors didn't necessarily foresee the recent developments in biodiversity informatics. The catalogue also includes 6,194 mentions of 59 countries where bees occur, 4,005 mentions of 1,219 visited plant species, 182 mentions of 115 plant species that bees nest in, 93 mentions of 66 parasite species hosted (some parasites are themselves bees) and 50 mentions of 37 hosts parasitised by parasitic bees.

How do we make the literature text itself available online?

Information on bees and pollination is very important in conservation and agriculture, particularly in the face of global change. There are excellent networks and collaborations on bee taxonomy and pollination ecology in Africa, which would benefit immensely from easier, integrated, structured and enriched access to bee biodiversity information and literature.

## 9) BLE - BHL-Europe's Virtual Exhibition

Authors: Jiri Frank, Manager, National Museum in Prague  
Jiri Kvacek, WP5 leader BHL-Europe, National museum in Prague  
Jana Hoffmann, Project Assistant BHL-Europe, Museum für Naturkunde

Abstract: The Biodiversity Library Exhibition (BLE) is a virtual exhibition of the digital content in the Biodiversity Heritage Library for Europe. It is a dissemination and e-learning tool which highlights specific biodiversity content and makes it accessible for a wider audience. The first two exhibitions will feature BHL-Europe's content on "spices" and "expeditions," presenting beautiful illustrations and informative text in old and rare books. It will also provide useful information for the visitor, e.g. recipes. The attractive design and easy to use interface of BLE has a great potential to show that historical literature on biodiversity can be interesting to a wide audience.

## 10) Biodiversity Literature in Europeana

Authors: Lizzy Komen, Business Project Coordinator, Europeana  
Jonathan Purday, Senior Communications Advisor, Europeana  
Jana Hoffmann, Project Assistant BHL-Europe, Museum für Naturkunde

Abstract: Europeana.eu provides online access to the digital resources from Europe's museums, libraries, archives and audiovisual collections. Europeana currently provides access to over 19 million items from 27 EU countries. BHL-Europe adds substantial value to Europeana by making available a great amount of biodiversity literature. Europeana is the EU's most visible expression of our digital heritage. [...] Europeana has established itself as a reference point for European culture on the Internet. It reflects the ambition of Europe's cultural institutions to make our common and diverse cultural and scientific heritage more widely accessible to all.

## 11) Primary Biodiversity Data Records in Legacy Literature Databases

Authors: Arturo H. Ariño, Museum of Zoology and Ecology of the University of Navarra  
Estrella Robles, Museum of Zoology and Ecology of the University of Navarra

Abstract: Automated extraction of primary biodiversity data records (PBDR) (i.e. the basic triad of taxon/location/time) from existing literature is desirable as a way to help fill gaps and increase fitness-for-use of global repositories of biodiversity data digitally stored from specimens and observations. Current efforts at extracting taxonomic data and their context from legacy literature through digitizing and OCR, such as Global Biodiversity Information Facility's (GBIF) Global Names Architecture (GNA), TaxonX, Innotaxa, Plazi, Fieldjournal, and other automated XML markup and tagging procedures applied to digitised literature increasingly available at BHL, are yet to produce unambiguous PBDR. Existing historical literature presents a high degree of formal variation which makes modelling in an XML schema quite difficult, so we still rely on manual parsing and digitization or markup for each complete PBDR.

This labor-intensive effort entails selective digitization because of its associated cost, and therefore may result in patterning of the acquired data, with high potential for gaps in knowledge. We explore some of these potential gaps by looking at patterns resulting from manual digitization of primary biodiversity data records into Zootron 4, a vintage taxonomic database including about 200,000 worldwide occurrence records of fauna manually captured from scientific literature over a period of more than two decades by biodiversity researchers according to their own selective interests.

Four broad classes of patterns were found: Taxonomic, geospatial, human-dependent, and chronological. However, these may reflect both intrinsic patterns existing in the examined literature and biases introduced by the researchers' selective processes. Incremental analysis involving other similarly recorded PBDRs, as well as comparisons with other patterns resulting from alternate sources of PBDRs such as collections and observations, may help recognize the main source of each pattern. For that, a standardization of datasets, for example through an extension of Darwin Terms, may be desirable.