ANCHORING BIODIVERSITY
INFORMATION
FROM SHERBORN TO THE 21st
CENTURY AND BEYOND

28th October 2011
Flett Theatre, Natural History Museum, London

Programme

10:00 Welcome and Logistics / Introduction to the Programme
Ellinor Michel (ICZN) & Graham Higley (BHL & NHM Libraries)

10:15 Opening Keynote: SHNH Annual Ramsbottom Lecture
Neal Evenhuis (Bishop Museum) Sherborn: Work history and impact of bibliography, dating and zoological informatics

10:55 Gordon McOuat (Univ of King’s College, Halifax) Naming and Necessity: Sherborn’s context in the late 19th Century

11:15 - 11:35 Coffee Break - posters & stands

11:35 - 12:40 Session 1: History of Taxonomic Literature, Indexing and Traditional Taxonomic Nomenclature
11:35 Edward Dickinson (Aves Press) Reinforcing the foundations: Filling in the bibliographic gaps in the historical legacy
11:55 F. Christian Thompson (Smithsonian) and Thomas Pape (Copenhagen) Systema Dipterorum: Sherborn’s critical influence in getting information control over a megadiverse group
12:15 Smithsonian Institution Libraries (Suzanne Pilsk, Martin Kalfatovic, Joel Richard) Unlocking the Index Animalium: From paper slips to bytes and bits
12:35 Nigel Robinson (Zoological Record) Sherborn’s Index Animalium integration into ION: access to all

12:55-13:30 Lunch - Pre-paid sandwich lunch in Flett foyer, posters & stands

13:30 Session 2: Current Taxonomic Practices
13:30 Chris Lyal (NHM) Digitising legacy taxonomic literature: processes, products and using the output
13:50 Henning Scholz (Museum für Naturkunde Berlin) BHL-Europe: Tools and Services for Legacy Taxonomic Literature
14:10 David Remsen (GBIF) Biodiversity Informatics: GBIF’s role in linking information through scientific names
**Poster presentations**

**Towards a complete list of family-group names in Coleoptera (Insecta) with comments on dates of publication**  

**Sherborniana – Artifacts of historical and heritage value from the Natural History Museum relating to CDS’s professional and personal collections**  
Paul Martyn Cooper

**Online synergy: Sherborn’s *Index animalium* and the Biodiversity Heritage Library**  
G. Costantino & L.K. Overstreet

**BLE – BHL - Europe’s virtual exhibition**  
Jiri Frank & Jiri Kvacek,

**The Global References Index to Biodiversity**  
Boris Jacob, Andreas Krausz & Melita Birthälmer

**Biodiversity literature in Europeana**  
Lizzy Komen & Jonathan Purday

**Calbug: Digitization of California’s Terrestrial Arthropod**  
Peter Oboyski, Joan Ball, Traci Grzymala & Kipling Will

**Charles Davies Sherborn: A magpie with a card index mind**  
Karolyn Schindler & Ellinor Michel

**Darwin: A systematic naturalist, a virtuoso or a miser?**  
Brian Rosen & Jill Darrell

**Sherborn’s *Index Animalium* - Systematic errors, mistakes and outdated judgements in the light of modern zoological nomenclature: An analysis based on the examination of 40,000 taxonomic names.**  
F. W. Welter-Schultes, A. Görlich & A. Lutze
ABSTRACTS

Charles Davies Sherborn and The Indexer’s Club
Neal L. Evenhuis
Bishop Museum, Honolulu, Hawaii 96817, USA
Email: NealE@bishopmuseum

Charles Davies Sherborn was an indexer. And he followed a long line of indexers. And a longer line of indexers followed him. They/we are all members of “The Indexer’s Club”. A club of obsessed individuals who, for some weird reason, find it necessary to not only facilitate a semblance of order, but to make sometimes incredibly huge amounts of information available to others [sacrificing their social lives and labouring on what spouses and colleagues may consider esoteric projects in order to save others from the same work]. And in doing so, encumbering most of the day and the wee hours of the night with a passion and fervour few other human beings can even begin to understand. This presentation will explore the bits of Sherborn’s life that led to that passion for indexing; and touch upon the impact he has had on bibliographies and researching the dates of publication; upon nomenclature; and upon the indexing of names — and it will attempt to explain why he did this and where we all can go as a result.

Naming and Necessity: Sherborn’s Context in the late 19th Century
Gordon McOuat
History of Science and Technology Programme, University of King’s College, Halifax, NS, CANADA
Email: gmcouat@dal.ca

By the late 19th Century, storms plaguing early Victorian systematics and nomenclature seemed to have abated. Vociferous disputes over radical renaming, the world shaking clash of all-encompassing procrustean systems, struggles over centres of authority, and the issues of language and meaning had now been settled by the institution of a stable imperial museum and its catalogues, a set of rules for the naming of zoological objects, and a new professional class of zoologists. Yet, for all that tranquillity, the disputes simmered below the surface, re-emerging as bitter struggles over synonyms, trinomials, the subspecies category, the looming issues of the philosophy of scientific language, and the aggressive new American style of field biology — all pressed in upon the received practice of naming and classifying organisms and the threat of anarchy. In the midst rose an index. This paper will explore the context of CD Sherborn’s Index Animalium and those looming problems and issues which a laborious and comprehensive “index of nature” was meant to solve.

Reinforcing the foundations: Filling in the bibliographic gaps in the historical legacy
Edward C. Dickinson
Editor, Zoological Bibliography, Aves Press, BN20 7JG Eastbourne, UK
Email: info@avespress.com; edward@asiaorn.org

Ornithological nomenclature is based on the bibliographic legacy from Charles Davies Sherborn, working in the Natural History Museum, London, and from Charles Wallace Richmond, working at the Smithsonian Institution, Washington, D.C. Despite their significant foundations, a complete data series has not yet been achieved. Gaps in their original coverage, though few, have not been resolved. The post-1850, the end date of coverage of the Index Animalium the level of completeness declines. I will discuss the coverage of the gaps in ornithology and address the primary issues of completeness and accuracy.

Avian names in the Index Animalium have issues of accuracy in spellings, authorship and citation details. Most of the problems that can be pinpointed in ornithology will be paralleled in other zoological disciplines. Post-1850, ornithology is fortunate in the correspondence between Sherborn and Richmond. The Richmond Index to the Genera and Species of Birds, published on microfiche in 1992 and now available online, is founded on their collaboration. After Richmond, successive members of the Department of Birds at the United States National Museum were inspired and encouraged to update the resource regularly. Over the years since 1932 when Richmond died there were periods when this card index was well
maintained and others when less time was devoted to it. In addition, the information available to ensure it was comprehensive is likely to have been only marginally better in respect of the Americas than was available to the Zoological Record. There has been more deliberate work done to maximise the collection of avian generic names. The initial sustaining role played by the Zoological Society of London must be recognised as regards both the Zoological Record and the Nomenclator Zoologicus of Neave. Unfortunately, ornithologists have undervalued the importance of the bedrock of information that these initiatives provide and hence they have done little or nothing collectively to maintain and complete these resources.

The rare Book Room at the NHM holds what may be all Sherborn’s Index Animalium slips. They are appropriately separated, but old explanatory separators written by Sherborn are fading and the original sequences within the segments look disturbed. These need study and potentially reorganisation. For their long term preservation and wider availability scanning is recommended (after any agreed reorganisation), It is hoped that the museum, whose Trustees were publishers of the 33 volumes that cover 1801-1850, will assess the situation and if necessary seek to raise funding for these measures. Other Sherborn material should perhaps be brought together with the slip cabinet so that all material relating to the Index Animalium is together or fully cross-referenced. At the Smithsonian, the Department of Birds holds two card indexes which Richmond created to support his primary card index. These are being preserved and are accessible on site. The Richmond card index is now online under development at http://www.zoonomen.net/cit/RI/Genera/RIGen.html.

Finally, there are registers of timed periodical receipts and of exchanges involving the Smithsonian International Exchange program. Mining such information could well resolve long-standing disputes over precedence. Finally, zoological bibliographers should work with the global antiquarian book trade to develop a digital collection illustrating the wrappers of the many early part-works where these contain dates, original spellings or other vital information.

Sherborn’s critical influence in getting information control over a megadiverse group

F. Christian Thompson
Smithsonian Institution, Washington DC

Thomas Pape
Natural History Museum of Denmark, Copenhagen Systema Dipterorum

Email: tpape@snm.ku.dk

The order Diptera (Insecta), flies, is a megadiverse group, representing some 15% or more of the known species of organisms. Scientific names are tags to concepts (hypotheses), called species, by which we organize our knowledge of biodiversity. Our Systema Dipterorum provides an index to all scientific names related to flies, so access to our knowledge about them is readily available. Sherborn more than a century ago attempted to provide such an index to all animal names. He did provide an index to all names published up until and including 1850. We compare our indexes, revealing how standards have changed and the number of names increased. Today, more and better resources are being made available to us, such as the Biodiversity Heritage Library, and our standards are higher (new International Code of Zoological Nomenclature), but regardless of all the change, Sherborn for his time provided an almost perfect (99.9%) index.

Unlocking the Index Animalium: From paper slips to bytes and bits

Suzanne Pilsk
Smithsonian Institution Libraries, Smithsonian Institution, Washington D.C., USA

Co-authors: Joel Richard, Martin Kalfatovic
Email: PilskS@si.edu

Smithsonian Institution Libraries received funding in 2004 to digitize Sherborn’s Index Animalium. The initial project was to digitize the pages images and re-key the data into a simple data structure. As the project evolved, a more complex database was developed to
enable quality searching to retrieve species names and to search the bibliography. The OCRed, scanned *Index Animalium* was re-keyed to the specifications of 99.995% accuracy rate. Working off the lessons learned by MBL WHOI Library’s project for Neave’s *Nomenclator Zoologicus*, simple expressions were used to break apart the re-keyed text. Coinciding with the development of the Biodiversity Heritage Library (2005), it became obvious there was a need to integrate the scanned *Index Animalium*, BHL’s scanned taxonomic literature, and taxonomic intelligence. The challenges of working with legacy taxonomic citation, computer matching algorithms, and making connections have brought us to today’s goal of making Sherborn available as open linked data. The goal is to allow repurposing of data, partnering with others to allow machine-to-machine communications and sharing information for broad discovery and access.

**Sherborn’s *Index Animalium* integration into ION: access to all**

**Nigel Robinson**  
Zoological Record Director, York Operations Thomson Reuters  
Email: nigel.robinson@thomsonreuters.com

As a list of zoological taxonomic names used in the scientific literature, the *Index to Organism Names* (ION) has provided a reference service to researchers for a number of years, based on the content of international publications analysed for inclusion in *Zoological Record*. However, this content is limited to names included in *Zoological Record* (1864-present). This project extends the names database and brings new opportunities through collaboration between the Smithsonian Institute, Thomson Reuters, and the Biodiversity Heritage Library to provide a more complete resource to complement other taxonomic names services.

**Digitising legacy taxonomic literature: processes, products and using the output**

**Chris Lyal**  
Department of Entomology, Natural History Museum. London  
Email: c.lyal@nhm.ac.uk

To date, most digitisation of taxonomic literature has led to a more or less simple digital copy of a paper original – the output has effectively been an electronic copy of a traditional library. While this has increased accessibility of publications through internet access, for many scientific papers the means of indexing and locating them is much the same as with traditional libraries. OCR and born-digital papers allow use of web search engines to locate instances of taxon names and other terms, but OCR efficiency in recognising names is still relatively poor, people’s ability to use search engines effectively is mixed, and many papers cannot be directly searched. Instead of building digital analogues of traditional publications, we should consider what properties we require of future taxonomic information access. Ideally the content of each new digital publication should be accessible in the context of all previous published data, and the user able to retrieve nomenclatural, taxonomic and other data / information in the form required without having to scan all of the original paper and extract target content manually. This opens the door to dynamic linking of new content with extant systems – automatic population and updating of taxonomic catalogues, ZooBank and faunal lists, all descriptions of a taxon and its children instantly accessible with a single search, comparison of classifications used in different publications, and so on. The means to do this is currently marking up content into XML, the more atomised the mark-up the greater the possibilities for data retrieval and integration. Mark-up requires XML that accommodates the required content elements and is interoperable with other XML schemas, and there are now several written to do this, particularly TaxPub, taxonX and taXMLit, the last of these being the most atomised. Building on earlier systems for mark-up of legacy literature ViBRANT is developing a new workflow and seeking to increase the automated component of the process. Manual and automatic data and information retrieval is demonstrated by projects such as INOTAXA and Plazi. As we move to creating and using taxonomic products through the power of the internet, we need to ensure the output, while satisfying the requirements of the Code, is fit for purpose in the future.
BHL-Europe: Tools and Services for Legacy Taxonomic Literature

Henning Scholz
Museum für Naturkunde, Berlin
Email: henning.scholz@mfn-berlin.de

Literature research is the base for the scientific work of taxonomists. Therefore, large and well-curated natural history libraries are a very important prerequisite to carry out scientific projects efficiently. The library work, however, has several serious limitations that slow down the work significantly. The natural history library corpus is highly fragmented and scattered. In particular much of the early published literature is rare or is only available in a very few libraries. A lot of time and effort is involved to find and collect all scientific works that are necessary for a specific project.

Today, quick and easy access to digital literature is more and more important to facilitate scientific work. Over the last few years a large number of library resources for taxonomists have been made available online. Since 2007, the Biodiversity Heritage Library (BHL) project is digitising the biodiversity literature holdings of numerous libraries in the UK and USA and making them available on the internet.

Since 2009, the eContentplus project Biodiversity Heritage Library for Europe (BHL-Europe) is developing four different access routes to the biodiversity literature digitised by many European and global partners over the last years. With the Global References Index to Biodiversity (GRIB, http://grib.gbv.de/), BHL-Europe provides in collaboration with the EDIT project a union catalogue of library holdings of many European and US libraries. This will facilitate the search for literature, either digitised or not. This tool will also facilitate the management of digitisation projects all over the world and collect scan request from the scientific community. For an effective access to already digitised literature, BHL-Europe is building a multilingual portal for the scientific community. This portal will also have functionalities currently not available in the BHL portal. The BHL-Europe Portal will, for example, facilitate the search for common and scientific names of biological organisms as well as person names through the implementation of various webservices (e.g. Catalogue of Life, VIAF). The backbone of the portal is a preservation and archive system built on a customised storage infrastructure housed by the Natural History Museum in London. We are currently collecting digitised literature from 27 different content providers on our servers, including all the content that is currently available through the BHL portal (http://www.biodiversitylibrary.org). In order to serve also a broader audience, the digitised literature available by BHL-Europe is also accessible by Europeana, Europe's digital library, archive and museum (http://www.europeana.eu/).

Biodiversity Informatics: GBIF’s role in linking information through scientific names

David Remsen
Senior Programme Officer. Electronic Catalog of Names of Known Organisms
Global Biodiversity Information Facility Secretariat, Copenhagen, Denmark
Email: dremsen@gbif.org

The Global Biodiversity Information Facility provides access to primary biodiversity data. It currently provides access to over 300 million data records from over 8000 different databases. Data customers expect to be able to retrieve data records organised around taxa. The challenges in integrating these data, originating in different sources, is considerable and requires access to both taxonomic and nomenclatural authority files. These challenges and the subsequent capabilities these resources enables are presented.

LANs: Lists of Available Names – a new generation for stable taxonomic names in zoology

Daphne Gail Fautin
Professor, Ecology and Evolutionary Biology Curator, Natural History Museum (Biodiversity Institute) University of Kansas, USA & ICZN
Email: fautin@ku.edu
Miguel Alonso-Zarazaga
Article 79 of the ICZN Code, which appeared first in the Fourth Edition, outlines a procedure for adding large numbers of names to the List of Available Names simultaneously, as a Part of the List. This feature has gained importance with the development of Zoobank, because the LAN can be an important adjunct or component of Zoobank. Article 79 describes a deliberative process, detailing steps for submission and for consideration by the public and Commission, and their chronology: submission must be by “an international body of zoologists,” and the proposed Part must be available for “comments by zoologists” for 12 months, followed by another 12-month period for comments on the proposed Part as revised in light of comments received. However, Article 79 it is mute about the contents of the submission. It is clear that adding a Part to the List will prevent long-forgotten names from displacing accepted ones – thus, for taxa on the List under the provisions of Article 79, nomenclatural archeology will not be worthwhile. Beyond that, Commissioners who participated in writing the Fourth Edition are divided about the intent of Article 79: some aver it is intended to document every available name within the scope of the Part, others it is to pare the inventory of names within the scope of the Part. The comprehensiveness of the names in the Part is critical because, according to Article 79.4.3, “No unlisted name within the scope (taxonomic field, ranks, and time period covered) of an adopted Part of the List of Available Names in Zoology has any status in zoological nomenclature despite any previous availability” (names may subsequently be added only “in exceptional circumstances,” according to Article 79.6). Under the first interpretation, the Part functions as a strictly nomenclatural archive. Under the second interpretation, the Part pares away nomina dubia, so Parts of the List resulting from actions under Article 79 are like the Approved Lists of Bacterial Names that took effect on 1 January 1980 – taxonomically recognizable as well as nomenclaturally available. It is critical that a consistent basis for implementing Article 79 be adopted; it is unrealistic to expect unanimity, given the diversity of opinion among those who helped craft Article 79.

Approaches to preserving digitized taxonomic data: prints, manuscripts, specimens

Chris Freeland
Director, Center for Biodiversity Informatics, Missouri Botanical Garden, St Louis, MO, USA
Email: chris.freeland@mobot.org

The availability of digitized taxonomic data has increased dramatically over the past twenty years as national funding agencies have strengthened their support of digitization activities and as scanning devices have become less expensive and easier to operate. As such, natural history museums and libraries have taken on new responsibilities for managing electronic information as ways of providing enhanced opportunities for educational outreach and scholarly dissemination. Museums and libraries have to consider how best to create and care for electronic resources given a volatile technology landscape with rapidly changing file formats and display devices. This session will address methodologies for responsible curation of digitized prints, manuscripts, and specimens, and will outline best practices for safeguarding digitized taxonomic data to ensure longevity of resources.

New workflows for “gold-standard” taxonomy

Sandra Knapp
International Association for Plant Taxonomy (IAPT); Int’l Trust for Zoological Nomenclature (ITZN); Dept. Of Botany, The Natural History Museum;
Email: s.knapp@nhm.ac.uk

Monographs have long been considered the “gold-standard” of taxonomic treatments. These comprehensive compilations of all knowledge about particular taxa or groups of taxa traditionally culminated a taxonomist’s career and often took a lifetime to produce. Descriptive taxonomy has been likened to a cottage industry, the antithesis of the big science approach to tackling global problems. Taxonomy itself is currently in a state of flux; opinions differ as to whether revolution or evolution is necessary, and even
then, what should actually happen when we do decide. Do we still need “gold-standard”, life-consuming single-author taxonomic works? Are traditional monographs doomed to go the way to dinosaurs? I will explore the elements of monographic, “gold-standard” taxonomic work and examine the workflow we traditionally use to produce them. I will suggest that taxonomists are important for only some parts of this workflow, and with that explore what modern taxonomists are really for and by extension what a modern monographic/taxonomic workflow might look like.

ZooKeys: Streamlining the registration-to-publication pipeline

Lyubomir Penev, Teodor Georgiev & Pavel Stoev
Pensoft Publishers, Sofia, Bulgaria
Email: info@pensoft.net

Electronic registration of nomenclatural acts is a dynamic process which still needs to be elaborated in accordance with the Biological Codes. The main questions to resolve are: (1) When a registration should take place, before or after publication? (2) Who is doing the registration? (3) Who approves the registration record? (4) Who confirms the validity of the registration record at the moment of publication? (5) At which point a registration record should be validated (e.g., a new name available), on the day of electronic or printed publication? At the International Botanical Congress in Melbourne in July 2011 it became clear that the registration process will be organized in different ways for the different organisms' kingdoms (Fungi, Plants, Protista and Animals).

ZooKeys was the first journal to provide a mandatory registration in ZooBank for all newly described taxa and to include the ZooBank's LSIDs in the original publication. In our view, the registration of nomenclatural acts and the quality control of the bibliographic metadata in ZooBank should be a responsibility of taxonomy publishers. Besides, we are convinced that registration of nomenclatural acts should be mandatory, independently of that will these be published on paper or online-only.

Within the framework of the EU FP7 project ViBRANT, and in a close collaboration with Zoological Record, ZooBank, the International Plant Name Index (IPNI), MycoBank and Index Fungorum, we are elaborating a workflow and associated software tools to streamline the registration within the editorial, publication and dissemination process. The talk will present Pensoft's vision on how to make the registration process secure and cost-efficient, through the currently developed Pensoft Writing Tool (PWT).

Towards an open taxonomy

Roderic D. M. Page
Professor of Taxonomy, Institute of Biodiversity, Animal Health and Comparative Medicine
College of Medical, Veterinary and Life Sciences, University of Glasgow
Email: r.page@bio.gla.ac.uk

Taxonomy is in many ways still pre-digital. Most taxonomic databases are little more than digitised index cards linking names to often-cryptic bibliographic citations, oblivious to the growing volume of scientific literature that is now online. A growing fraction of taxonomic literature is becoming freely available, either through adoption of Open Access publishing models, or through digitising efforts such as the Biodiversity Heritage Library. Yet much of the most basic information about biodiversity, namely taxonomic description, remains either behind a pay wall, or only available in paper form. This talk sketches the goal of an "Open Taxonomy." The first step towards this goal is digitally linking scientific names to the primary literature using standard identifiers such as DOIs. I argue that until we make serious inroads into this task, taxonomic knowledge will remain in a ghetto largely ignored by the wider scientific community.

Towards a Global Names Architecture: The future of indexing scientific names

Richard Pyle
Bishop Museum, Honolulu, HI, USA & ICZN
Email: deepreef@bishopmuseum.org
For more than 250 years, the taxonomic enterprise has remained almost unchanged. Certainly the tools of the trade have improved: months-long journeys aboard sailing ships have been reduced to hours aboard jet airplanes; advanced technology allows humans to access environments that were once utterly inaccessible; GPS has replaced crude maps; digital hi-resolution imagery provides far more accurate renderings of organisms that even the best commissioned artists of a century ago; and primitive candle-lit microscopes have been replaced by an array of technologies ranging from scanning electron microscopy to DNA sequencing. But the basic paradigm remains the same. Perhaps the most revolutionary change of all – which we are still in the midst of, and which has not yet been fully realized – is the means by which taxonomists manage and communicate the information of their trade. The rapid evolution in recent decades of computer database management software, and of information dissemination via the internet, have both dramatically improved the potential for streamlining the entire taxonomic process. Unfortunately, the “potential” still largely exceeds the reality. The vast majority of taxonomic information is either not-yet digitized, or digitized in a form that does not allow direct and easy access. Moreover, the information that is easily accessed in digital form is not yet seamlessly interconnected. In an effort to bring “reality” closer to “potential”, a loose affiliation of major taxonomic resources, including GBIF, the Encyclopedia of Life, NBII, Catalog of Life, ITIS, IPNI, ICZN, Index Fungorum, and many others have been crafting a “Global Names Architecture” (GNA). The intention of the GNA is not to replace any of the existing taxonomic data initiatives, but rather to serve as a dynamic index to interconnect them in a way that streamlines the entire taxonomic enterprise: from gathering specimens in the field, to publication of new taxa and related data.

**POSTER ABSTRACTS**

Towards a complete list of family-group names in Coleoptera (Insecta) with comments on dates of publication.

P. Bouchard, Y. Bousquet & A.E. Davies

Coleopterists recently synthesized data on all known extant and fossil Coleoptera family-group names for the first time (Fig. 1). A catalogue of 4887 family-group names (124 fossil, 4763 extant) based on 4707 distinct genera in Coleoptera was given. A total of 4492 names were determined to available. Names were listed in a classification framework. The authors recognized as valid 24 superfamilies, 211 families, 541 subfamilies, 1663 tribes and 740 subtribes.

For each name, the original spelling, author, year of publication, page number, correct stem and type genus were included. The original spelling and availability of each name were checked from primary literature.

Here we provide information about the resources that were used to infer the correct date of publication of works in which Coleoptera family-group names were proposed. We compare these resources with those that previous workers, such as C.D. Sherborn, would have used for similar projects before the advent of computers and the internet.

Online synergy: Sherborn’s *Index animalium* and the Biodiversity Heritage Library.

Grace Costantino & Leslie K. Overstreet
Smithsonian Institution Libraries, Washington D.C. USA
Email: overstreetL@si.edu
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.

**BLE – BHL - Europe's virtual exhibition**

**Jiri Frank & Jiri Kvacek**  
National museum in Prague, Czech Republic

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

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 of interesting to a wide audience.

**The Global References Index to Biodiversity**

**Boris Jacob**  
Royal Museum for Central Africa, Tervuren, Belgium

**Andreas Krausz**  
Head Office of the Common Library Network GBV (VZG), Göttingen, Germany

**Melita Birthälmer**  
Museum für Naturkunde, Leibniz Institute for Research on Evolution and Biodiversity at the Humboldt University Berlin, Germany

The Global References Index to Biodiversity (GRIB) is a union catalogue of European natural history libraries. It contains de-duplicated 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.

**Biodiversity literature in Europeana**

**Lizzy Komen & Jonathan Purday**  
Europeana, c/o the Koninklijke Bibliotheek, National Library of the Netherlands, The Hague, The Netherlands

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

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." Neelie Kroes, EC Vice President

Calbug: Digitization of California's Terrestrial Arthropods

Peter Oboyski, Joan Ball, Traci Grzymala & Kipling Will
Essig Museum of Entomology, University of California, Berkeley, CA, USA
Email: kipwill@berkeley.edu

Sherborn's legacy now transcends the capture of taxonomic literature to the capture of individual specimen data from museum collections on which literature records are based. While the basic concepts and standards in data management still apply new challenges need to be met, including new data types and formats, sharing data across platforms, and the sheer volume of information to be managed.

Although most biological data standards are now well-established, databasing of entomology collections has lagged behind other collections largely due to the quantity of specimens and the highly abbreviated and inconsistent data found on very small specimen labels. Calbug is an NSF funded collaborative of the eight major entomology collections in California that intends to capture 1.1 million specimen-level data records from our combined holdings in a Darwin Core-compliant MySQL relational database.

We will analyze these data using geospatial technology to understand the relationship between changes in distribution and the precise nature and extent of habitat modification. Given that successfully capturing 1.1 million records would only account for a small fraction of our combined holdings, development of time-saving methods and technology for getting data from labels into databases is paramount. In the initial stage of the project we have focused on developing and testing methods and workflows to radically increase the rate of data capture, while maximizing data quality appropriate for the biotic change analyses. Digital imaging of data labels provides a more easily viewed verbatim archive of specimen data and allows subsequent off-site data entry from image files using manual entry, crowd-sourcing, and automated OCR and data parsing.

Specimen handling, both in terms of time and risk to specimens, remains a significant obstacle to retrospective data capture from entomological collections. Georeferencing is also a challenge due to the highly abbreviated and inconsistent nature of location data on specimen labels, but a number of strategies that combine computer and human data handling are being used.

A magpie with a card-index mind: Charles Davies Sherborn 1861-1942
Karolyn Shindler & Ellinor Michel,
ICZN & Natural History Museum, London, UK
Email: karolynshindler@aol.com

Charles Davies Sherborn was geologist, indexer and bibliographer extraordinaire. He was fascinated by science from an early age - although there are probably very few small boys who attempt to construct volcanoes in their gardens, the consequent explosion resulting in a visit from the police. Like so many Victorians, the young Sherborn was a passionate natural history collector and was obsessed with expanding his collection of land and freshwater shells. He later described himself as being a ‘thorough magpie’ and having ‘a card-index mind’, and these two traits coalesced in his monumental *Index Animalium*, an index of every known living and extinct animal from 1758 to 1850. The *Index* was a true labour of love - and shamefully little financial reward - that occupied 43 years of his life. One of the first visitors through the doors of the Natural History Museum in South Kensington when it opened in 1881, Sherborn began work there seven years later as one of the small band of unofficial scientific workers, paid by the number of fossils he prepared. By the time of his death in 1942, Sherborn’s corner in the Museum was the first port of call for generations of scientists seeking advice, information – or an invitation to one of his famous “smoke and chat” parties.

In addition to his work on the Index, Sherborn is also responsible for rescuing the correspondence, manuscripts and books of Sir Richard Owen, the great Victorian comparative anatomist and the prime mover behind the creation of the NHM. The papers were “in a
cow-shed, exposed to rats and rain”. The manuscripts were piled twelve feet high, while the correspondence filled countless packing cases. But for Sherborn, this was very heaven. In high excitement he wrote to a friend, “I must husband all my time and strength now, for it is a giant’s task set before me, and this must cap...my other works”.

Darwin: A systematic naturalist, a virtuoso or a miser?

Brian Rosen & Jill Darrell
Dept. of Zoology & Dept. of Palaeontology, Natural History Museum, London, UK
Email: B.Rosen@nhm.ac.uk

Although the publication details of many natural history specimens are generally well-documented in museum collections, it is also not uncommon for the published status of other specimens, such as types, to be incomplete, uncertain or unknown. This can happen when original documentation was insufficient, or when it has become separated from the specimen, mislaid, or even lost altogether. Here, we suggest a methodological framework (‘collection trajectory’) for reconstructing or recovering such information.

Charles Robert Darwin (1809-1882) was a prolific collector of geological and biological specimens. As he said of himself, “The passion for collecting, which leads a man to be a systematic naturalist, a virtuoso or a miser, was very strong in me, & was clearly innate, as none of my sisters or brother ever had this taste”. Provisional results of a recent survey carried out by us suggest that the Natural History Museum holds over 14,000 of his specimens, including significant types, and almost entirely biological. The bulk of them are beetles and barnacles, though the Museum also holds most of his birds, mammals (living and fossil), fish, reptiles and amphibians, as described in the ‘Zoology of the Beagle’ (Darwin 1838-43). Substantial Darwin collections also exist elsewhere, notably his geological material at the University of Cambridge.

Many labels of Darwin’s specimens give only very brief information, while other specimens which might have been collected and/or studied by Darwin, bear little or no evidence of that. Many people therefore think that Darwin’s specimen documentation was poor. In fact, he was extremely methodical in this, and worked hard to practise his own advice that the “collector’s motto” should be “‘Trust nothing to the memory’, for the memory becomes a fickle guardian when one interesting object is succeeded by another still more interesting.” As a result, he left us a complex legacy of lists, field notebooks and diaries. Ironically though, this complexity can make it difficult to find the necessary information about a given specimen. It is therefore essential to understand the relationship between these various sources, and also Darwin’s specimen numbering system.

How can this legacy be used to recover relevant information for any given specimen? Our ‘generalized collection trajectory’ provides a nine-point framework for working methodically through all the potential sources of information about a given collection of Darwin’s. We also suggest that a similar approach might be used for other people’s collections where similar problems exist. For Darwin in particular, this led us to revise and extend earlier reviews of Darwin’s specimen lists. In our poster, (and as an aid to identifying previously uncertain specimens), we include illustrations of the number-tags Darwin used for his dry specimens, and give examples showing how our trajectory approach sheds further light on some examples of NHM coral reef specimens.

AnimalBase and Index Animalium

Welter-Schultes, F.
Zoologisches Institut der Universität, Göttingen
Görlich, A. & Lutze, A.
E-mail: fwelter@gwdg.de

Appreciating Sherborn’s tremendous work implies understanding to which extent Sherborn’s index data can be used for nomenclatural purposes today. In the course of the AnimalBase project to digitise early zoological literature and provide a taxonomic names database we cross-checked our own manual examinations of 40,000 new names in the original sources with those of Sherborn’s Index Animalium. For each examined work we extracted all new names under the present-day nomenclatural rules (4th edition of the ICZN Code), and compared our results with Sherborn’s list extracted from the same work.

It was crucial to know how to read the Index, only 70% of the 420,000 names in Sherborn’s list were marked as new (300,000 new names). We found that Sherborn’s data were consistent with our own
finds at an average rate of 80-90%. The degree of reliability of Sherborn's data differed by work and by animal group, and depended on various factors. The rate of misspellings in Sherborn's manual work was low, lower than in the AnimalBase project, but naturally not zero. The proportion of overlooked names in each work depended on its style. Sherborn did not have all important works at his disposal.

Some categories of systematic errors and mistakes were under Sherborn's responsibility (obvious difficulties in understanding foreign languages except Latin, careless examination of difficult works to save time, neglecting subspecific names), others have to do with the nomenclatural rules having changed in the past 100 years (criteria for availability of names, corrections of incorrect Latin, authorships for names, unavailability of non-binominal works).

Sherborn was confronted with many problems we also had in our own work. This included the difficulty to maintain a common standard over time. We came to the conclusion that anyone who intends to repeat Sherborn's job will inevitably be fascinated by his low non-systematic error rates.

Charles Davies Sherborn provided the bibliographic foundation for current zoological nomenclature with his magnum opus *Index Animalium*. In the 43 years he spent working on this extraordinary resource, he anchored our understanding of animal diversity through the published scientific record. No work has equalled it since and it is still in current, and critical, use.

Until now, Sherborn's contribution has been recognised by professional taxonomists worldwide but he has escaped the celebration of his accomplishment that is his due. This symposium is in his honour in the 150th year of his birth at the Natural History Museum, London, with an international panel of experts on bibliography and biodiversity bioinformatics, linking a view of the past with an active debate on the future of the related fields.

Sponsored by
The International Commission on Zoological Nomenclature (ICZN), the Linnean Society of London, the Natural History Museum London, the Society for the History of Natural History (SHNH), Biodiversity Heritage Library – Europe, ZooKeys/Pensoft, ViBRANT

This meeting could not have happened without the support of the host Institution: The Natural History Museum, London.

Thanks to all those concerned who helped bring this together!

Charles Davies Sherborn

You served us well: and to the task you chose
Brought courage and a matchless care and zeal.
Perfection such as few have ever sought,
Or, seeking, have achieved in what they wrought,
Earned the just mark of Oxford's scarlet seal.

And we, true sons of Adam, who suppose
The myriad creatures of this world to tell,
Using the means you gave to guide our way
Through the maze of names bestowed since Linné's day,
In tribute murmur, "Friend, you served us well":

July 1942            H.DIGHTON THOMAS
The International Commission on Zoological Nomenclature (ICZN) acts as adviser and arbiter for the zoological community by generating and disseminating information on the correct use of the scientific names of animals.

The ICZN is responsible for producing the International Code of Zoological Nomenclature - a set of rules for the naming of animals and the resolution of nomenclatural problems.

www.iczn.org

The Linnean Society of London is driven by a single purpose as outlined in our first charter, 200 years ago: The cultivation of the Science of Natural History in all its branches

Founded in 1788, the Society is named after the great Swedish scientist Carl Linnaeus (1707-1778). His botanical, zoological and library collections have been in its keeping since 1829. These are now accessible on the web.

www.linnean.org

The Society for the History of Natural History is celebrating its 75th anniversary as a friendly international society for everyone who is interested in natural history in the broadest sense. This includes botany, zoology and geology as well as natural history collections, exploration, art and bibliography. Everyone with an interest in these subjects – professional or amateur – is welcome to join.

www.shnh.org.uk

This meeting has been made possible by support from the following sponsors and supportive organisations:

Sponsors:
- ICZN – Int’l Commission on Zoological Nomenclature
- SHNH - Society for the History of Natural History
- Linnean Society
- BHL-Europe – Biodiversity Heritage Library-Europe
- Pensoft Publishers (ZooKeys)
- NHM - Natural History Museum, Science Directorate
- ViBRANT – Virtual Biodiversity

Supportive organisations