Library & Laboratory:
the marriage of research, data and taxonomic literature
Report of London Meeting,

Eighty participants from 22 countries gathered in London on Saturday and Sunday, 5-6 February 2005 to discuss the status and future of access to the taxonomic literature[1]. The goal of the meeting was to review the degree to which taxonomic information (especially research publications in taxonomy) have been made available in digital form to the research community, and to propose an agenda for actions that would improve the research environment for taxonomy. The participants were taxonomists; librarians; publishers; representatives of learned and professional societies, private foundations and government agencies; and specialists in information and communications technology. The meeting was planned by a committee of five representatives[2] of major natural history museums and their libraries, and support for the meeting was provided by the Alfred P. Sloan Foundation of New York and the National Biological Information Infrastructure of the US Geological Survey, Reston, Virginia.

The meeting was organized and held in conjunction with the first International Barcode Conference[3]. DNA barcoding is a technique for identifying species using short gene sequences from standardized positions in the genome. Barcode sequence data are being deposited in GenBank, the US nucleotide database at NIH. GenBank records are normally linked electronically to the journals in which the gene sequences are published, but relatively few taxonomic journals are available or indexed online. The Barcoding Initiative was eager to expand electronic access and indexing of taxonomic literature and helped to organize this meeting toward that goal.

The presentations[4] were organized into four sections:

- Needs: Who are the stakeholders concerned with the taxonomic literature and what are their needs and concerns?
- Culture and History: How have the institutions and cultural norms in taxonomy created today's research literature?
- Economics: How have market forces shaped today's research literature, and how will changing market forces affect the future of information access in taxonomy?
- Technology: What new models of information access have been introduced from which the taxonomic community could learn?

Participants spent considerable time in open discussion during and after these presentations. The following action items emerged from those discussions.

---

[1] See meeting agenda and participant list, appended to this report. Presentations made during the meeting are available online at http://biodivlibrary.si.edu/

[2] Graham Higley and Chris Lyal, Natural History Museum, London; Tom Moritz, American Museum of Natural History, London; Anna Weitzman, National Museum of Natural History, Smithsonian Institution, Washington, DC; and Tom Garnett, Smithsonian Institution Library, Washington, DC

[3] Held Monday-Wednesday, 7-9 February 2005 at the Natural History Museum, London.

[4] The presentations made during the meeting are available at http://www.barcoding.si.edu/ConferencePresentations.htm

Near-Term Action Plan:

1. The technology for massive digitization of books is now available and cost-effective. Taxonomic publications that are out of copyright could be digitized and placed in the open access domain. The Smithsonian will be convening a meeting of major natural history libraries in May 2005 to develop a roadmap for massive digitization of legacy literature. This roadmap should include a business plan for digitizing and making available all legacy taxonomic literature on an industrial scale.
2. Longer-term planning will require a database of the taxonomic literature that indicates (a) the portion that is already available in digital form, (b) the portion that is in the process of being digitized, and (c) the portion for which plans have been created for digitization. Short-term priorities can then be established for future digitization. Many elements of the taxonomic literature are not tracked by library utilities such as OCLC because much of the digitizing in this area is not done by libraries. At present, no organization has taken on responsibility for this task.
3. Many important taxonomic publications are not indexed by online services. This makes it impossible to link to the bibliographic citations for these publications. Journals and societies that were represented at the meeting were willing to have their indices and abstracts in PubMed, but PubMed has a process of vetting and approving their inclusion. PubMed should be willing to reach out to other journals and societies that are important to taxonomists and to invite these journals to apply for inclusion in PubMed. Greater clarity on the process and criteria used for admission of journals in PubMed is needed. The Consortium for the Barcode of Life (CBOL) and the Smithsonian's National Museum of Natural History will convene a meeting of the major online indexes (e.g., PubMed, Biosis, Agricola) with the goal of creating a unified list of journals for which online tables of contents, abstracts, and indices of content are not available. This would set the stage for negotiations between journals and indexing services and, eventually, a controlled vocabulary of journal names in taxonomy.
4. A global list of species' names is needed. Participants challenged CBOL to work toward establishing a complete online list of accepted/valid names and linking to them from the Barcode Section of GenBank. CBOL should convene a meeting in the near future that brings together representatives of GenBank and the major compilations of species' names (GBIF, Species2000, ITIS, uBio, etc.)
5. The presentations from this meeting should be placed on a website and all other related information (e.g., tools, open source software, links to known digitized literature, standards and projects) should be added as it becomes available. The Smithsonian will perform this task.

Mid-Term Action Plan:

6. Schemata are being developed for biodiversity data and a central clearinghouse for information is needed to avoid unnecessary proliferation of competing standards. GBIF plans to explore the development of a dictionary of schemata.
7. There is no universally recognized system for the identification of entities in taxonomy (specimens, publications, etc.) A system of globally unique identifiers (GUIDs) is a vitally important component of a distributed system, and must be

integrated into the developing system of interoperable databases. GBIF will be working on this in the next year, including a determination of the most appropriate form of GUID to use.

8. One of GBIF's four work programmes is an electronic catalog of taxonomic names (ECAT). The results of the meeting on species' names convened by CBOL should be forwarded to GBIF/ECAT for consideration for its programme of work.
9. TDWG and GBIF are leading efforts to develop data standards for biodiversity information. Standards need to be coordinated and/or developed which integrate biodiversity domains, including taxonomy, ecology, marine science, and molecular biology. At present, no organization has taken responsibility for this task, although it is considered to be in the long-term purview of GBIF.

Long-Term Action Plan:

10. The Digital Divide between taxonomists in industrial and developing countries is severe. Access to taxonomic literature and other information resources is critically needed in developing countries.
11. The use of standards for biodiversity data of all types will permit interoperability, thus reducing redundancy of data entry and increasing prospects of linkage between data. Standards and interoperability are developing through *ad hoc* partnerships but especially through TDWG and GBIF.
12. Authors of taxonomic papers are commonly asked or required to sign exclusive agreements with the publisher. This essentially signs the author's copyright over to the publisher. In placing their papers in traditional printed journals with these exclusive rights, authors are prohibited from offering access to their work digitally through other conduits. When the publisher's copyright expires, the contents of these journals are added to the backlog of undigitized literature. The participants in the meeting suggested that taxonomists should not continue to contribute to the backlog of 'legacy literature'. In negotiating with publishers, taxonomists should avoid signing exclusive agreements and should try to find journals that couple print publication and open digital access.
13. Professional recognition and promotion is commonly tied to print publications and citations of papers in traditional journals. New systems of recognition are needed that include online publication as part of an individual's research performance. In addition, systems of ranking and citations are needed for on-line publications. Once in place, these new metric systems of performance related to online publishing may be adopted by universities and research institutes as part of their performance appraisal systems.

Outreach Action Plan. Participants in the meeting agreed that outreach and education should be integrated into all the action plans described above. In so doing, the results of this meeting could be used as a reference point for advancing future work. Taxonomists, librarians, administrators, funders, government agencies, the Global Taxonomy Initiative national focal points and other national focal points should be participants in these outreach and education efforts.

The goals of the outreach and education efforts should be to:

- Increase the number and geographic coverage of taxonomists;
- Increase support for taxonomy;
- Improve the access of biodiversity rich but resource poor countries to taxonomic resources, most of which are located in other regions of the world;
- Break down the digital divide through capacity building (GBIF, Bio-NET International, and many others are undertaking some of this action);
- Implement the overall design for the information infrastructure of taxonomy in a goal-driven manner. Overall design includes the specific requirements for data contest as well as workflow processes;
- Create toolkits for information access that are easy to use, thereby lowering the barriers to participation in taxonomy.

Day 1: 5 February 2005			
Opening Address	1030 - 1100	Welcome The big idea – why is this important?	Scott Miller, Smithsonian Institution
	1100 – 1115	COFFEE BREAK	
Session 1 Research Needs and Necessary Content. *Defining stakeholders and their needs.*	1115 – 1215 1215 - 1315	1. Supporting the 'Barcoding of Life' 2. Supporting Ecologists 3. Supporting Systematists 4. Supporting applied research. 5. **Research Needs and Necessary Content: Discussion**	Session Chair: Anna Weitzman Rapporteur: Tom Moritz Robert Hanner, CBOL/Coriell Institute Charles Godfray, Imperial College Sandra Knapp, NHM Fred Grassle, OBIS/Rutgers University
	1315 – 1445	LUNCH BREAK	
Session 2 **Evolution of Cultural and Institutional Practices.** *What are the current impediments to providing open access to and use of all systematic literature; includes both taxonomic and publishing norms.*	1445 – 1545 1545 - 1645	6. BioMed Central - a model for publishing in an open access world. 7. Reducing legal blocks & barriers to Open Access 8. Nomenclatural blocks and barriers. 9. Taxonomic support and output: blocks & barriers. 10. **Changing Cultural and Institutional Practices: Discussion**	Session Chair: Chris Lyal Rapporteur: Graham Higley Marianne Josserand, BioMed Central Heather Joseph, BioOne Andy Polaszek, ICZN Chris Lyal, NHM & Anna Weitzman, Smithsonian Institution
	1645 – 1700	COFFEE BREAK	

Session 3 Funding/Resources/Economics. *The micro-economic and macro-economic issues that surround the enabling of open access models.*	1700 – 1745 1745 - 1830	11. Macro-economic case for open access. 12. Micro-economic cases for various classes of systematics publishers to move to open access. 13. A view from a funding organisation. **Funding/Resources/Economics: Discussion**	Session Chair: Tom Moritz Rapporteur: Anna Weitzman Tom Moritz, AMNH Raym Crowe, Chainbridge A speaker from the funding sector, TBD
	1830 - 2130	NETWORKING BUFFET & DRINKS	

Day 2: 6 February 2005			
Session 4 **Technological Developments and Opportunities.** *What are the components of a prospective institutional model for publishing and what are the already existing models that might be considered? What will we need to build...?*	1030 - 1130 1130 - 1215	14. Ontological analysis. 15. Taxonomic data standards. 16. Taxonomic literature content and interchange standards. 17. Putting it all together - technological integration at a global level. 18. **Technological Developments and Opportunities: Discussion**	Session Chair: Tom Garnett Rapporteur: Chris Lyal Bryan Heidorn, Univ. of Illinois Neil Thomson, NHM Anna Weitzman, Smithsonian Institution & Chris Lyal, NHM Donald Hobern, GBIF
	1215 – 1230	COFFEE BREAK	
Discussion and Report. *What do we need to do next? Timescales and priorities; including current activities and expected outcomes for the next 3 years.*	1230 - 1400	19. **Discussion and resolution.** 20. **Presentation and Discussion of Draft Report for CBOL**	Session Chair: Graham Higley Rapporteur: Tom Garnett Anna Weitzman, Smithsonian Institution
	1400	LUNCH BREAK/END OF MEETING	

Participant List

Name	Organization	Country
Donat Agosti	American Museum of Natural History	USA
Gabriel Ameka	Ghana Herbarium, Department of Botany, University of Ghana	Ghana
Tetiana Andrianova	M.G. Kholodny Institute of Botany, KIEV	Ukraine
Jesse Ausubel	Sloan Foundation	USA
Shelley L. Ball	National Centre for Advanced Bio-protection Technologies, Lincoln University, Canterbury	New Zealand
James H. Beach	Biodiversity Research Center, Univ. of Kansas	USA
Rafael Borroto	Instituto de Ecologia y Sistematica, Habana	Cuba
Thierry Bourgoin	Museum Nat. Histoire Naturelle	France
Vishwas Chavan	National Chemical Laboratory, Pune	India
Katherine Chiang	Mann Library, Cornell University	USA
Mike Claridge	Cardiff University, Wales	UK
José Clavijo	AndinoNET-BioNET-INTERNATIONAL, Museo del Instituto de Zoología Agrícola	Venezuela
Raym Crowe	Chainbridge	UK
Neil Davis	University of California, Berkeley	USA
Ebrahim Ebrahimi	Plant Pests & Diseases Research Institute, Tehran	Iran
Jayanthi P. Edirisinghe	University of Peradeniya, Peradeniya	Sri Lanka
Kate Edmondson	Natural History Museum, London	UK
Elizabeth Ferguson	Blackwell Publishing	UK
Gail Fordham	Natural History Museum, London	UK
Tom Garnett	Smithsonian Institution	USA
George Garrity	Michigan State University	USA
Birgit Gemeinholzer	Freie Universität Berlin	Germany
Charles Godfray	Imperial College	UK
Fred Grassle	OBIS/Rutgers University	USA
Robert Guralnick	Univ of Colorado	USA
Robert Hanner	Coriell Institute for Meidcal Research	USA
Jan Haseplagh	Flanders Marine Institute	Belgium
Bryan Heidorn	University of Illinois	USA
Donna Herendeen	Librarian, The National Agricultural Library, Beltsville, MD USA	USA
Graham Higley	Natural History Museum, London	UK
Donald Hobern	Global Biodiversity Information Facility	Denmark
Douglas Holland	Missouri Botanical Garden	USA
Heather Joseph	BioOne	USA
Ravindra P. Joshi	Philippine Rice Research Institute (PhilRice)	Philippines
Marianne Josserand	PubMed	UK
Martin R. Kalfatovic	Smithsonian Institution Libraries	USA
Mikhail V. Kalyakin	Zoological Museum of Moscow Lomonosov State University	Russia
Jeyaraney Kathirithamby	Univ. of Oxford, Department of Zoology, South Parks Road, Oxford	UK
Paul Kirk	CABI Bioscience	UK
Sandra Knapp	The Natural History Museum	UK
Drew Koning	American Museum of Natural History	USA
Frank Krell	The Natural History Museum, Dept of Entomology	UK
John Kress	Department of Botany, United States National Herbarium, Smithsonian Institution	USA
Thomas F. Lahr	National Biological Information Infrastructure	USA
Po-Feng Lee	Academia Sinica Nankang, 115 Taipei	Taiwan

Chris Lyal	The Natural History Museum	UK
Catherine Lyons	Names for Life	UK
Santiago Madriñán	Universidad de los Andes, Bogotá	Colombia
Julien Masanès	Internet Archive	Netherlands
John McNeil	International Code for Botanical Nomenclature, Royal Botanic Garden, Edinburgh	UK
Scott Miller	Smithsonian Institution	USA
Tom Moritz	American Museum of Natural History	USA
Bob Morris	Univ. Massachusetts Boston	USA
Bridget Neal	Smithsonian Institution	USA
Catherine Norton	uBio, Marine Biology Laboratory, Woods Hole	USA
Thomas Orrell	Integrated Taxonomic Information System	USA
Alan Paton	Royal Botanic Garden Kew	UK
Andrew Polaszek	International Commission on Zoological Nomenclature	UK
Martin Pullan	Royal Botanic Garden Edinburgh	UK
David Remsen	uBio, Marine Biology Laboratory, Woods Hole	USA
Connie Rinaldo	Museum of Comparative Zoology, Harvard University	USA
Gary Rosenberg	Academy of Natural Sciences Philadelphia	USA
David Schindel	Smithsonian Institution	USA
Richard Smith	BIONET International	UK
Sabine Stohr	Swedish Museum of Natural History	Sweden
Sri Sulandari	The Indonesian Institute of Sciences	Indonesia
Alistair Taylor	Natural History Museum, London	UK
Neil Thomson	Natural History Museum, London	UK
Shun-Chern Tsaur	Academia Sinica,Research Center for Biodiversity, Taipei	Taiwan
Erik J. van Nieukerken	National Natural History Museum Naturalis, Leiden	Netherlands
Ronald Vonk	Zoological Museum, Amsterdam	Netherlands
Richard Wakeford	The British Library	UK
Charles M. Warui	National Museums of Kenya	Kenya
Elizabeth Watson	The Swedish Museum of Natural History	Sweden
Anna Weitzman	National Museum of Natural History, Smithsonian	USA
Shen-Horn Yen	National Sun Yat-Sen University	Taiwan