



Life and Literature and Beyond

A Report to the Richard Lounsbery Foundation, Inc.

Smithsonian Libraries

Martin Kalfatovic | Biodiversity Heritage Library, Program Director

Grace Costantino | Biodiversity Heritage Library, Program Manager

Table of Contents

Executive Summary	3
Smithsonian Libraries Grant Budget.....	4
The Life and Literature Conference.....	4
Executive Summary	4
The Conference	5
Summary of Conference Findings and Recommendations	7
Life and Literature Conference Summary	12
No Cost Extension.....	12
The 2012 BHL Staff and Technical Meeting.....	13
BHL Technical Developments	14
Global Meeting: Berlin, Germany	15
Technical Meetings in Missouri	16
The New User Interface: Background and Details.....	17
Outreach.....	18
Conclusion	18

Executive Summary

In February, 2011, the Biodiversity Heritage Library (BHL) received an award of \$89,171.00 from the Richard Lounsbery Foundation to fund a conference to discuss the future of biodiversity literature digitization. This conference, entitled the Life and Literature Conference, was held at the Field Museum, November 14-15, 2011. 124 participants from approximately 30 countries attended, including librarians, biologists, computer scientists, publishers, students, educators, and others. Comprised of two plenary sessions, four panel discussions, and four break-out sessions, the conference gave a very effective structure for many BHL users to reflect on the uses and problems of biodiversity literature and to offer concrete suggestions to the BHL on how to improve its services and content.

As a result of a no-cost extension received for the Lounsbery grant, BHL staff were able to use remaining award funds to implement suggestions from Life and Literature.

The Biodiversity Heritage Library elected to use the remaining funds to support a Staff and Technical Meeting that would allow staff to revise BHL's strategic plans, as informed by feedback from the Life and Literature Conference, and prioritize projects and initiatives identified at the Life and Literature Conference in order to inform a work plan for BHL. In addition to this meeting, remaining funds were also used to directly support technical development for BHL. These technical developments involved travel expenses for BHL developers to attend meetings to address technical issues and supporting an upgrade to the BHL user interface. The user interface upgrade includes design updates, article-level access to BHL content, improved book navigation and viewing, improved PDF creation, and implementing improved taxon name finding algorithms.

Finally, staff utilized funds to support outreach activities, ensuring that the BHL project is widely known within the communities best positioned to take advantage of the resources it offers.

Thus, thanks to support from the Lounsbery Foundation, BHL was able to identify the needs of the communities it serves, prioritize development activities, begin developing a workplan to address these activities, support technical development, and further engage with its audiences.

Smithsonian Libraries Grant Budget

Presented below is the budget for the funds awarded to the Smithsonian Libraries from the Lounsbery Foundation. Please refer to the detailed financial documents accompanying this report for specific spending information.

Funding Categories	Expenditures
Life and Literature Conference Expenses	\$50,139.25
Staff and Technical Meeting	\$15,276.80
Global Meeting	\$4,894.06
Missouri Technical Meeting	\$1,269.29
Hardware and Software	\$414.97
Outreach	\$5,545.63
Indirect Costs	\$11,631.00
Total	\$89,171.00

The Life and Literature Conference

Executive Summary

To address the current and future issues of the digitization and accessibility of biodiversity literature the Biodiversity Heritage Library proposed a large, public conference titled “Life and Literature” and received funding from the Richard Lounsbery Foundation. The conference was held at the Field Museum, November 14-15, 2011. 124 participants from approximately 30 countries attended including librarians, biologists, computer scientists, publishers, students, educators, and others. The Life and Literature Conference was an enormous success and appreciated by the attendees. The Conference gave a very effective structure for many BHL users to reflect on the uses and problems of biodiversity literature and to offer concrete suggestions to the BHL on how to improve its services and content. BHL Member Institutions will refine and rank suggestions arising from the conference as a key step in preparing a revised project work plan.

The Conference

The purposes of the conference were:

- Engagement of current and future constituencies concerned with biodiversity literature.
- Integration and interoperability of the BHL with major biodiversity and digital library programs.
- Further innovative integration of digitized literature with biodiversity databases.
- Publishing models and legacy scientific literature.
- Setting the agenda for the BHL for the next 4 – 5 years.

There were two plenary speakers:

Dr. Richard Pyle:

Richard Pyle's main field of expertise involves the taxonomy and biogeography of coral-reef fishes. His other areas of interest include the use of advanced diving technology to document biodiversity inhabiting deeper regions of tropical coral reefs, and also the development of computer database systems (and associated data standards) for managing biodiversity information. He earned his PhD under the guidance of John E. Randall, and has worked in the Department of Natural Sciences at Bishop Museum since 1986. He has authored over a hundred scientific, technical, and popular articles, has been featured in dozens of documentary films, and has received a number of significant awards for his achievements. Pyle is also a commissioner of the International Commission on Zoological Nomenclature.

George Dyson:

George Dyson is a historian of technology whose interests have included the development (and redevelopment) of the Aleut kayak (*Baidarka*, 1986), the evolution of digital computing and telecommunications (*Darwin Among the Machines*, 1997), and the exploration of space (*Project Orion*, 2002).

Turing's Cathedral, "a creation myth for the digital universe," will be published by Pantheon (USA) and Penguin (UK) in January 2012.

There were four panels:

Research, Informatics, and the Published Record.

The panelists were Dr. Sandra Knapp (Natural History Museum), Dr. Stanley Blum (California Academy of Science), Dr. Donat Agosti (American Museum of Natural History), and Dr. Elycia Wallis (Museum Victoria).

Publishers, Aggregators, and Authors - New Models and Access.

The panelists were Susan Skomal (BioOne), Jan Reichelt (Mendeley), Lyubo Penev (Pensoft), Catriona MacCallum (PLoS), Abel Packer (SciELO)

Learning and Education.

The panelists were Dr. Marie Studer (Encyclopedia of Life), Dr. Douglas Wilkin (CK-12 Foundation), Dr. Devin Reese (NSF Resource Center), Dr. Ken Walker (Museum Victoria), and Natalia Zamora (INBio).

Building Collaborative Networks for Science and the Humanities through Scientific Literature.

The panelists were Dr. Janet Browne (Harvard), Dr. Alain Touwaide (Smithsonian), Dr. Peggy Macnamara (Field Museum), and Dr. Chris Wildrick (Syracuse University).

These were followed by breakout sessions that focused on four key questions for the future of the BHL:

1. What new types of content and connections among content should BHL focus on bringing into its collection in the next 5 years?
2. What new technologies should BHL adopt and integrate, or plan for, in the next 5 years?
3. What other institutions, domains, or user groups should BHL engage within the next 5 years?
4. What does education mean for BHL? What is the audience? Targeted? Broad?

At the close of the conference, the four breakout sessions reconvened and presented their main findings for what BHL should maintain, expand, innovate, and avoid in future development.

Summary of Conference Findings and Recommendations

What follows is a synthesized summary of the notes, findings, and recommendations from the four breakout groups from the Conference.

The Life and Literature conference clearly established that the BHL is no longer a pilot project. It is part of the workflow for biological research and relied on by many.

BHL is used by people in diverse fields and in many ways:

1. Practicing taxonomists doing primary zoological, botanical, and, to a lesser extent, mycological research.
2. Historians in such fields as history of science, agricultural history, period re-enactment, and cultural history.
3. Artist, illustrators, conceptual artists.
4. Hobbyists.
5. Teachers
6. Librarians (public, special, academic)
7. Bioinformatics Professionals
8. Students in the fields of zoology, botany, history of science, mycology, and more

BHL is used by other research information systems and their users:

1. Biodiversity informatics systems such as the Encyclopedia of Life, BioStore, the European Distributed Institute of Taxonomy (EDIT), AlgaeBase, AnimalBase, WoRMS
2. Bibliographic aggregator/facilitator systems such as CrossRef, OCLC, Europeana
3. Museum Collection Information systems: California Native Plant Society's Inventory of Rare and Endangered Plants (www.rareplants.cnps.org), Royal Museum of Central Africa.
4. Online library catalog systems such as the library systems at the Smithsonian Institution, Harvard Museum of Comparative Zoology, Missouri Botanical Garden, New York Botanical Garden, California Academy of Science, and the Field Museum Library.
5. Global Names (<http://globalnames.org/>).

BHL Users and Stakeholders Want Many Things

Participants expect a significant expansion in scope and scale of the Biodiversity Heritage Library (BHL). In fact, some participants assigned almost any conceivable biodiversity-information need as a problem the BHL should attempt to solve. This tendency should not be viewed as a problem but rather as an opportunity to more clearly define the boundaries of the BHL.

Discussions in the breakout sessions identified a number of areas where the BHL requires significant enhancement:

- ✦ Contextualization information for educators
- ✦ Improved accuracy of Optical Character Recognition
- ✦ Clearer citation resolution

Breakout Sessions Revealed that BHL should avoid:

- ✦ Imposing layers of interpretation

- ✦ Charging
- ✦ Closed access
- ✦ Being wedded to old technology
- ✦ Being bought out
- ✦ Limiting access
- ✦ Growing too big too fast
- ✦ Loss of focus (“spread too thin”)
- ✦ Bad metadata

Breakout Sessions Revealed that BHL should innovate:

- ✦ Functional workflow to correct poor OCR
- ✦ Identification and annotation of illustrations
- ✦ Better ways of sorting search results, e.g. by taxon name occurrences, by date, by original description, by co-occurrence of other terms.
- ✦ Full text searching of the OCR by journal or book volume.
- ✦ User created collections at every level of access
- ✦ Click & go: Connect grey literature to published literature
- ✦ Click & go: Connecting published sources to published literature
- ✦ Explore ways to serve in-copyright content
- ✦ Compare literature to current environmental/conservation biology literature
- ✦ Discovery layer services
- ✦ Common names – allowing people to add and correct

- ✦ Article-handling, article metadata.
- ✦ Multi-lingual access
- ✦ Annotation

Breakout Sessions Revealed that BHL should expand:

- ✦ Content based on taxonomic communities (ViBrant, Scratchpads)
- ✦ Accept boutique scanned collections
- ✦ Connect w/ Mendeley
- ✦ Connect w/ ZooRecord
- ✦ Define CiteBank & make a trusted repository
- ✦ Article level access
- ✦ BHL Resources & funding!
- ✦ BHL citations in Wikipedia
- ✦ Connect better to Google
- ✦ Make scan-on-demand easier

What is feasible in the next 5 years?

What can be done in the next year?

- ✦ Make BHL Expeditions tool available by the end of 2012!
- ✦ Collect and coordinate feedback and focus groups on interfaces and use cases
 - ✦ BHL US performed user testing on UI in October 2011
 - ✦ Major outcomes: make it easier to find illustrations and improve pagination

- ⤴ BHL US has a feedback link on their website, through which they have and will continue to receive feedback on UI improvements
- ⤴ Introduce news feed on home page
 - ⤴ Utilize the 2 million photos that BHL China will be putting up.
 - ⤴ Advertise that BHLEU has 14 languages.
 - ⤴ Link to all other BHL sites and harmonize portals
- ⤴ Include maps with illustrations
- ⤴ Have autofill on forms such as create your own PDF.

What can we do in the next 1- 2 years?

- ⤴ Coordinate across countries to work on educational packages (e.g., tools, features and resources) for youth
 - ⤴ Make site more aesthetically pleasing to draw in users
- ⤴ ID intermediaries to talk to and get message out
 - ⤴ Teachers, librarians, game developers, museum educators
 - ⤴ Volunteer advocates
 - ⤴ DVDs/CDs and preset materials for teachers to present to students, especially in developing countries where access to the internet is challenging.
- ⤴ Larger presence on social networks
 - ⤴ Right now on US and UK site, very active twitter feed/Facebook/blog/Flickr
- ⤴ Coordinate Focus Groups!
 - ⤴ ideas for use, who will use site?

- ✦ not just for science, but for art and other audiences

What can we do in the next 3-5 years?

- ✦ Integrate GIS component into BHL searching.
- ✦ Develop Tools and API so game designers and media producers can use BHL info in what they create
- ✦ Create way for users to pull and integrate BHL computable data (wing length, life span, etc)
- ✦ Partnerships with publishers to get more modern materials

Life and Literature Conference Summary

BHL Member Institutions will refine and rank suggestions arising from the conference as a key step in preparing a revised work plan for the next three to five years. Subsequent to the Life and Literature conference, a matrix of technical development was drafted for review by BHL stakeholders as a first step in the process.

The Life and Literature Conference was an enormous success and appreciated by the attendees. The results of a post-conference survey are included with this report (*Life and Literature Conference: Post-conference Survey Results*). The Conference provided an effective structure for attendees to reflect on the uses and problems of biodiversity literature and to offer concrete suggestions to the BHL on how to improve the current environment.

No Cost Extension

Thanks to cost efficiencies at the Field Museum, which hosted the Life and Literature Conference, a balance of \$27,400.75 remained after all conferences expenses were addressed. Thus, in March, 2012, the Biodiversity Heritage Library received a no-cost extension from the Lounsbery Foundation , extending the award from its original end date of 12/31/2011 to 11/30/2012, with final reports due to the foundation on 2/28/2013. In addition to the extension, BHL received permission to use the remaining funds to support some of the recommendations resulting from the Life and Literature Conference.

The Biodiversity Heritage Library elected to use the remaining funds to support a Staff and Technical Meeting that would allow staff to revise BHL's strategic plans, as informed by feedback from the Life and Literature Conference, and prioritize projects and initiatives identified at the Life and Literature Conference. In addition to this meeting, remaining funds were also used to directly support technical development for BHL. These technical developments involved travel expenses for BHL developers to attend meetings to address technical issues and supporting an upgrade to the BHL user interface. The user interface upgrade includes design updates, article-level access to BHL content, improved book navigation and viewing, improved PDF creation, and implementing improved taxon name finding algorithms.

The 2012 BHL Staff and Technical Meeting

On September 27-28, 2012, thirty-one staff members representing all 14 BHL member institutions convened at the Ernst Mayr Library at the Museum of Comparative Zoology at Harvard University for the 2012 BHL Staff and Technical Meeting. As a combined meeting, it brought together not only those that manage the digitization workflow at each member institution, but also those that work to keep BHL's technical infrastructure running smoothly and constantly improving.

Lounsbery funds were used to cover travel costs for one representative from each BHL member institution, as well as meeting room fees and catering costs.

One key purpose of the meeting was to review BHL's strategic plan, including mission statement, vision statements, and project goals, based on the feedback received from the Life and Literature Conference. Secondly, staff prioritized the suggested projects and initiatives for BHL identified at the Life and Literature Conference (combined with those already under development or identified by the BHL Technical Team). As a result of the meeting, elements for a 2 year BHL workplan, comprised of the highest priority projects and initiatives, were outlined. Additional projects and initiatives were designated as either lower priority, and thus to be addressed 2+ years from now, or priority funding opportunities. Priority Funding opportunities are high priority projects for which BHL does not have sufficient staff or resources to address, therefore requiring the acquisition of additional funding.

As a result of the meeting, a new BHL vision statement and revised BHL project goals were drafted. Elements for a revised BHL mission statement were also outlined. The BHL Secretariat is currently revising those

suggestions from the meeting and will present final recommendations to the BHL Steering Committee in time for the Spring BHL Institutional Council Meeting.

The following Projects and Initiatives were identified as Priorities for a 2 Year Workplan for BHL:

- Develop a Permissions Plan in order to increase permissions scanning
- Develop a BHL Fundraising Plan
- Develop a BHL Marketing and Outreach Plan
- Pursue partnerships with collaborators that generate funding opportunities
- Implement full-text searching
- Include more government and university collections
- Collaborate with partners that already have existing digitized literature

The Following Projects and Initiatives were identified as Priority Funding Opportunities:

- K-12 science and art programming using BHL content
- Increase manual pagination for BHL content
- Allow Annotation and Mark-up
- Allow common names searching
- Citizen Science Projects
- Provide multi-lingual access
- BHL in a Box (i.e. Allow local serving of BHL data on a general user's computer)
- Add GIS component
- Mobile BHL
- Pursue and Generate BHL games (i.e. games that improve BHL content)
- Support user-created collections and community building

BHL Technical Developments

Lounsbery funds were also used to support key BHL technical developments by supplementing travel costs for BHL developers to attend technical meetings and supporting an upgrade to the BHL user interface that will include article-level searching, improved book navigation and viewing, improved PDF creation, and the implementation of new taxon name finding algorithms. These improvements were all identified as priorities at the Life and Literature Conference.

As outlined, Lounsbery funds were used to support travel costs for two developers from BHL member institutions: Joel Richard of the Smithsonian Libraries and John Mignault of the New York Botanical Garden. Mr. Richard and Mr. Mignault traveled to the Museum für Naturkunde in Berlin, Germany, to attend the BHL-Global Meeting to discuss global technical issues, and the Missouri Botanical Garden (home to BHL's technical team) in St. Louis, MO, to work with BHL's lead developer, Mike Lichtenberg, to implement new user interface developments and discuss other key technical issues. Lounsbery funds were also used to purchase new computer equipment to replace outdated hardware within the BHL Secretariat.

Global Meeting: Berlin, Germany

On June 7-8, 2012, the Museum für Naturkunde (a BHL-Europe member) in Berlin, Germany, hosted the third BHL-Global Meeting. Representatives from nearly all BHL-Global (gBHL) nodes participated, including Europe, the US, Australia, Brazil, and Egypt. Partners provided project updates and engaged in crucial discussions regarding project sustainability, content synchronization and replication, communication and outreach collaboration, and workflow, scanning, and licensing cooperation.

The key technical issue discussed was the challenge of replicating content across all BHL nodes. There are currently six global nodes (BHL-US/UK, BHL-Europe, BHL-China, BHL-Egypt, BHL-Brazil, and BHL-Australia), all of which are producing their own content. There is currently no method to allow for all content to be available across all portals. Achieving global replication and synchronization was one of the tasks identified at the Life and Literature Conference for BHL to address within the next year. The global meeting brought each program node's technical staff together to discuss solutions to the problem, including Joel Richard and John Mignault, whose attendance was supported by Lounsbery funds.

At the close of the Berlin meeting, it was decided that all global partners would upload their content to Internet Archive, and that each global node could then pull that content into their own portals, thus allowing for global replication and synchronization. However, since that meeting, additional issues with using Internet Archive for this purpose were discovered, requiring further discussions among BHL's technical staff.

Technical Meetings in Missouri

On October 18-19, 2012, Joel Richard and John Mignault traveled to the Missouri Botanical Garden to meet with Mike Lichtenberg, BHL's lead developer, to further discuss various technical issues, many identified at the Life and Literature Conference, and provide development assistance for the new BHL user interface implementation. Technical issues discussed included further examination of a new method to allow for global replication and synchronization.

Technical Issues Discussed Included:

- The new BHL user interface
 - See details in section below
- Macaw
 - Macaw is software developed by Joel Richard at the Smithsonian Libraries that can be used to upload content to BHL that was scanned outside of the Internet Archive digitization workflow. Incorporating content digitized via “boutique” methods was a priority identified at Life and Literature.
- Article-level access to BHL content, another Life and Literature priority.
- The NEH Art of Life Project
 - BHL received a grant from the National Endowment for the Humanities to improve access to images in BHL. The grant involves creating and implementing an algorithm to automatically identify images in the BHL corpus.
- Copying BHL Content to SIL
 - Currently, Internet Archive is used as the repository for BHL content. BHL technical staff are working on migrating all BHL content to servers at the Smithsonian Libraries in order to maintain better control over BHL data.
- Syncing Content
 - Global replication and synchronization, as previously discussed.

The New User Interface: Background and Details

Informed by requirements identified at the Life and Literature Conference and October 2011 usability surveys, BHL partnered with BHL-Australia staff at the Museum Victoria to update the BHL user interface, merging the user-praised aesthetics and book viewer of the BHL-Australia portal with the functionality of the US/UK portal, while also implementing new features. The new UI will launch late winter, 2013.

Improvements Include:

- **Updated Design:** The website's design has been upgraded to reflect the celebrated aesthetics of the BHL-Australia portal.
- **Article and Chapter Access:** The ability to search BHL by article or chapter titles has been implemented. To date, over 82,000 articles and chapters have been indexed and are searchable within BHL. Additional articles and chapters will become available as the collections continue to be indexed.
- **Open Data Enhancements:** BHL's APIs, OpenURL interface, and Data Exports have been modified to include available article and chapter information.
- **Book Viewer Updates:** The BHL book viewer has been updated, allowing users to view up to 40 pages on screen at once and more easily navigate to a specific page within a book. Users can also view OCR text alongside page images, and, where the books have been indexed, users can navigate directly to the articles or chapters within.
- **PDF Creation Improvements:** The custom PDF creation process has been improved, allowing users to select pages for their PDF while in the book-viewer mode and more easily review the PDF before creation.
- **Improved Taxon Name Finding Algorithms:** BHL will implement a new algorithm capable of identifying previously undiscovered taxon names throughout the BHL corpus. Test applications of this algorithm on a portion of the BHL corpus have already resulted in an increase in nearly 50 million name instances in BHL, translating to over 20 million new unique names identified. These newly-identified names are currently available in BHL.

Outreach

In addition to the Life and Literature Conference and technical development, Lounsbery funds were also used to support BHL outreach activities, including fees for a booth at the Ecological Society of America 2012 Annual Conference and materials to promote BHL. Reaching out to the scientific and educational communities was one of the suggestions arising from the Life and Literature Conference, and these outreach activities allowed BHL staff to do just that.

Conclusion

The generous grant from the Lounsbery Foundation has not only allowed the Biodiversity Heritage Library to identify the needs of a wide and disparate community of users in regards to biodiversity literature digitization, but has also allowed the project to begin implementing many developments to support those needs. Staff now have a clear understanding of the tasks BHL must accomplish in order to remain relevant and indeed revolutionary within the biodiversity informatics community, and have begun drafting a work plan to tackle these tasks. Furthermore, BHL will soon launch a new user interface that addresses many of the suggestions proposed at the Life and Literature Conference, and technical staff have engaged in vital discussions regarding additional development challenges. Staff have also engaged in valuable conversations with scientific and educational audiences, ensuring that the BHL project is widely known within the communities best positioned to take advantage of the resources it offers. Such accomplishments would not have been possible without support from the Lounsbery Foundation.