

https://dspace.org/

The biodiversity heritage library and African digital libraries in the global context.

Kalfatovic, Martin R.; Fourie, Anne-Lise

2015-05-29
Smithsonian Institution Washington, DC, United States
http://hdl.handle.net/10539/18400

Downloaded from a DSpace repository. https://dspace.org/

THE BIODIVERSITY HERITAGE LIBRARY AND AFRICAN DIGITAL LIBRARIES IN THE GLOBAL CONTEXT

Martin R. Kalfatovic
Smithsonian Libraries / Biodiversity Heritage Library
Smithsonian Institution
Washington, DC, United States

Anne-Lise Fourie
South African National Biodiversity Institute / Biodiversity Heritage Library Africa
Pretoria, Gauteng, South Africa

Abstract

The Biodiversity Heritage Library (BHL) was created in 2006 as a direct response to the needs of the taxonomic community for access to early literature. Designed with meeting these needs, the BHL was grant-funded and quickly proved its value to its target users. The original BHL organizational model, based on US and UK partners, provided a template for, first, BHL Europe, and then a series of global nodes, most recently, BHL Africa and BHL Singapore. As the BHL moved from project to a cornerstone of biodiversity infrastructure, sustainability, appropriate expansion, collaboration with national and pan-national digital libraries (Europeana, the Digital Public Library of America) became more important. Working within the unique and often challenging environments of Sub-Saharan Africa, BHL assisted in the creation of BHL Africa is an inclusive network of African libraries and institutions in Western, Eastern, Central and Southern Africa. The short term goals of BHL/BHL Africa are surveying the biodiversity content of Sub-Saharan partners, defining digitization and aggregation capacity, and delivering African content to the BHL portal. This presentation will cover principles of pan-institutional digital library development, working across multiple African institutions, address areas of growth, and formulate lessons learned through global BHL growth.

Overview

> BHL is one of the jewels in the crown of biodiversity informatics. It has delivered a resource that is already of high value to taxonomists, collection managers and naturalists around the world. This new partnership between GBIF and BHL will enable us to explore closer linkages between literature, collections and species concepts - helping to add further value to the data mobilized by our respective networks. Donald Hobern, GBIF Executive Secretary (GBIF 2014)

Operating as a consortium of natural history, botanical institutions, research institutions, and related organizations, the Biodiversity Heritage Library (BHL) is administered via a secretariat located at the Smithsonian Libraries in Washington, DC. BHL collections contain over 45 million pages from over 159,000 volumes. Services such as taxonomic name finding tools, custom PDF downloads, and open APIs allow users to easily locate and reuse these resources.

The BHL is revolutionizing the way scientific research and conservation is conducted, helping scientists to identify, describe, and conserve the world's species and the habitats that support them. Since its creation in 2006 and portal launch in BHL has served millions of users in nearly in over 240

countries and territories around the world. On average, BHL receives more than a million visits per year.

The BHL's open access collections and services enable scientists to find the information they need to identify, describe, and conserve the world's species and habitats. The services in the BHL were developed in response to the research needs of the scientific community. BHL offers taxonomic name finding tools, custom PDF downloads and a user feedback mechanism that enables users to contribute to collection development and enhancement, ensuring that BHL remains responsive to the needs of a growing global community.

Staffing

BHL central staff are located at the Smithsonian Libraries (Secretariat) and the Missouri Botanical Garden (Technical).

Secretariat staffing is as follows:

- **Program Director** (Martin R. Kalfatovic). Provides day-to-day leadership of the BHL and general outreach to user community and content providers.
- **Program Manager** (Carolyn Sheffield). Provides day-to-day management of the BHL; oversees all internal BHL communications, manages finances, and coordinates membership activities.
- **Collections Coordinator** (Bianca Crowley). Oversees and coordinates all collection development activities of the BHL. Serves as Chair of the BHL Collections Committee; serves as a liaison between the Collections Committee and Technical Team. Works with content providers on issues related to metadata, copyright, and rights clearance.
- **Outreach and Communications Manager** (Grace Costantino). Manages all external communications on both traditional (print and press release) and social media for BHL.
- **Digital Collections Librarian** (Jacqueline E. Chapman). Serves as the feedback coordinator and performs primary triage on incoming scanning requests, metadata edits, and technical issues, acting as a virtual reference librarian to BHL's user community and as a point of contact for BHL member librarians.

Technical Team staffing is as follows:

- **Technical Director** (William Ulate). Provides day-to-day leadership of technical aspects of the BHL. Manages the BHL Technical Advisory Committee (composed of staff from across all BHL institutions) for guidance on technical issues. Manages technical infrastructure of the BHL and programming staff.
- **Web Developer** (Mike Lichtenberg). Provides all web application programing and site maintenance and development.
- **Data Projects Coordinator** (Trish Rose-Sandler). Provides day-to-day management of data ingest and special project for BHL.

The BHL Executive Committee (EC) provides overall guidance for BHL on issues of policy. The Executive Committee is elected from the BHL Members. Current EC members are Dr. Nancy E. Gwinn (Chair, Director, Smithsonian Libraries); Constance Rinaldo (Vice-Chair, Director of the Ernst

Mayr Library of the Museum of Comparative Zoology, Harvard University); Jane Smith (Secretary, Director, Natural History Museum, London).

Additionally, BHL Members and Affiliates provide the equivalent of 14.5 FTE in staff time to support the activities of the BHL.

Governance

The BHL is a consortium of libraries and related institutions governed by a Memorandum of Understanding and BHL Bylaws. Participation in BHL is at two levels. The first is Member. BHL Members pay an annual dues of $USD10,000 and participate in the governance of BHL and are eligible to elect the Executive Committee. Membership is recommended by a Membership Committee and approved by a vote of the members. The second is Affiliate. BHL Affiliates participate in the operations of BHL (through contributions of content, staff time, and related activities) and are recommended by the Membership Committee and approved by a vote of members.

There are also content contributors to BHL who are neither Members or Affiliates. These include scientific societies, data service providers and others.

Funding for BHL comes from a variety of sources. Member dues pay for certain central activities; members and affiliates also provide subvention of activities (BHL Secretariat and Technical Team) through line item funding or internal and external grants.

In kind donation of activities and FTEs provide nearly $USD1.6m in value to BHL (calendar year 2014).

Partnership

The BHL actively partners with content providers (libraries, publishers), large biodiversity organizations (the Global Biodiversity Information Facility, Encyclopedia of Life), as well as metadata aggregators (DPLA, Europeana).

BHL is currently actively engaging in the larger natural history community to build partnerships. Institutions that participate in the Council of Botanical and Horticultural Libraries (CBHL), SLA Natural History Caucus, International Association of Aquatic and Marine Science Libraries (IAMSLIC), and the American Alliance of Museums (AAM).

BHL staff are involved in these and similar organizations are have proven effective evangelists for participation in BHL. BHL has over eight years' experience in working with members of this community on issues related to metadata, rights (successfully negotiating the CC0 license for all metadata early on).

BHL Secretariat staff have a detailed "new member onboarding" checklist of activities for use in orienting new participants in the BHL. The BHL has grown from the original 10 US and United Kingdom Members to 17 Members and 6 Affiliates. These are supplemented by global nodes (with their own governance structures) in Europe, China, Australia, Sub-Saharan Africa, Egypt, Mexico, and Brazil.

Global BHL

The spread of the BHL meme was a direct outcome of the early success of BHL. BHL, formally launched in May 2007, had, by November 2008 attracted enough attention to generate interest in European institutions seeking EU funding for a congruent project. BHL invited interested European partners to various planning and technical meetings which culminated in the successful EU proposal that led to BHL Europe in January 2009.

The other global nodes took various paths, as outlined below. One prevalent thread in the formation of many of the global nodes was the relationship between BHL and the Encyclopedia of Life (EOL) and ties between EOL and other biodiversity organizations in those countries/regions.

Key lessons learned from my close involvement in the formation of the global nodes are:

- Successful nodes have passionate and committed individuals leading the formation of the node;
- Successful nodes grow organically with the participation of individuals and institutions; nodes that are driven by institutional mandate are less productive;
- Seed funding can create sustainable nodes under the right circumstances even when substantial further funding is unavailable (this is the case in both BHL Europe and BHL Australia which, though primary funding is over or substantially decreased, the elements of collaboration and passion for the BHL has kept the node active). This was also the case for BHL, where the six year MacArthur funding allowed time for the BHL and its partners to mature, take on appropriate leadership roles and grow towards sustainability;
- Being a successful BHL global node can take on many forms. A major lesson learned in the formation of global nodes is that replicating the BHL does not work across different regions and cultures, creates costly and inefficient redundancies, and leads to confusion in the biodiversity community.

Summary of Global BHL nodes

There are currently BHL global nodes located in Europe, China, Australia, Brazil, Egypt, Sub-Saharan Africa, Singapore, and Mexico. Areas of future growth include India, Russia, Canada, Latin America, Japan, and others.

The Biodiversity Heritage Library (2006)

When the BHL was formed in 2006, it arose through a direct need expressed by biodiversity researchers and informaticians, specifically, access to the historic literature of taxonomy. Preceded by a series of workshops in Washington and culminating in the "Libraries and Laboratories" conference in London (2005), the original ten members of BHL officially formed at a meeting held at the Smithsonian in 2006 (Marcum)

Building on a long history of cooperation, the ten members quickly identified the goals of BHL and sought funding to bring them about. Six institutions took the lead, Smithsonian Libraries, the Missouri Botanical Garden, the American Museum of Natural History, the Museum of Comparative Zoology, Marine Biological Laboratory/Woods Hole Oceanographic Institution, and the Natural History Museum (London).

BHL Europe (January 2009)

The BHL Europe project was created under the leadership of Henning Scholtz (Museum für Naturkunde, Berlin) and Graham Higley (Natural History Museum, London). Working under the constraints of the EU funding model, BHL Europe was created to aggregate existing European scanned content with existing BHL content in a European repository and multilingual portal. BHL Europe was successful in aggregating existing BHL content in Europeana. With the funding for the project officially ending in June 2012, ongoing BHL Europe work is done primarily with support for staff time from selected institutions (Natural History Museum, London; National Museum Prague, et al.) and volunteer work from a number of project participants.

BHL Egypt (December 2009)

BHL Egypt is another direct outgrowth of the international collaborations of the Encyclopedia of Life (EOL). The EOL worked closely with Dr. Ismail Serageldin of the Bibliotheca Alexandrina (BA) to establish an Arabic language portal. The BA was also a historic partner of the Internet Archive and maintained a mirror of IA content. The congruent creation of the BHL Egypt node was a natural outgrowth of the EOL/BA partnership. BHL Egypt has not yet contributed content to BHL, it has, however, created a complete mirror of BHL content in Egypt.

BHL China (January 2010)

Formed under the auspices of the Chinese Academy of Sciences (CAS), the BHL China node is associated directly with the CAS Institute of Botany. The BHL China node was a direct outgrowth of the EOL's relationship with the CAS and the interest in Dr. Keping Ma of the CAS in forming a BHL node. With support from the CAS/IB, BHL China has supported a single Internet Archive Scribe machine to digitize 19th and 20th century literature. To date, BHL China has not expanded to other institutes of the CAS or in aggregating other existing content.

BHL Australia (June 2010)

Formed as an outgrowth of the Atlas of Living Australia (ALA), and, again, under the auspices of the BHL/EOL partnership, BHL Australia was created with the primary goals of creating a portal that would feed literature to the ALA and also to digitize Australian content. With leadership provided by the Museum Victoria, a network of partners (ALA, etc.), BHL Australia created a low cost scanning kit (based on the ATIZ scanner) and facilitated upload of content to the Internet Archive (for ingest into BHL). The most important outcome of BHL Australia was the portal. Built on the model of the BHL portal, the Australian portal added new functionality, an enhanced book viewer, and a high level of design. BHL Australia also served as a metadata mirror for the BHL with a real time data sync established. With the end of the initial ALA funding, BHL Australia was supported by the Museum Victoria. With remaining ALA funds and BHL support, it was decided to merge the best of both the BHL Australia portal and the BHL portal. The new hybrid portal was launched in 2013. In April 2014, the metadata sync was officially ended as it was no longer necessary for ALA work.

BHL SciELO (December 2010)

After initial planning discussions led by Abel Packer (then head of BIREME) in 2009, the Brazilian government funded a meeting of a number of Brazilian biodiversity institutions and organizations in

February 2010 in São Paulo. Attended by librarians, researchers, and informaticians from around the country, BHL staff (Tom Garnett, Chris Freeland, and Martin Kalfatovic) conducted workshops on organization, collections development, and technology with their Brazilian counterparts. After the meeting, technical difficulties in importing equipment delayed the start of scanning. BHL SciELO remains a vigorous partner and has recently (2013) begun a flow of content into BHL.

BHL Africa (April 2013)

BHL Africa was the first global BHL node that was formed without a previous EOL relationship. BHL Africa was also the first formed directly under aegis of librarians. With the encouragement of Bryan Heidorn (a longtime friend of the BHL and an attendee at the 2005 London meeting), BHL approached the JRS Foundation to provide funding for African librarians and researchers to attend the 2011 BHL conference, Life and Literature held in Chicago.

Previously, the Smithsonian had a long history of work in Kenya and solid partnerships with the National Museums of Kenya (NMK) and the Kenyan Wildlife Service (KWS). Librarians at the NMK were in frequent contact with their Smithsonian counterparts and assisted in identifying Kenyan participants for the Chicago meeting. Likewise, SANBI's library and information center network provided good contacts within South Africa. At the conclusion of the Chicago meeting, however, the hoped for cohesion did not occur. One goal of the meeting, a focused, joint proposal to the JRS Foundation was not realized. With remaining JRS funds from the Chicago meeting, BHL (via the Field Museum) and SANBI proposed supplemental JRS funding to support a meeting in Cape Town in conjunction with a planned SANBI biannual meeting.

Unlike the Chicago meeting, this meeting would be run in the form of a workshop with specific outcomes identified, the primary outcome being the formation of an organizational structure that would create a BHL Africa. The Cape Town meeting was successful in that it created the rough framework for BHL Africa, identified key leaders (Anne-Lise Fourie of SANBI and Asha Awano of NMK), and technical staff (Lawrence Monda, NMK). Diligent work by Fourie and Awano led to the planning for the April 2013 launch of BHL Africa. The large meeting was funded with remaining JRS funds from the previous meeting and significant support by SANBI (Smithsonian Libraries, 2013).

BHL Singapore (February 2014)

Contacts initiated by Judith Warnement (Harvard Botany Libraries) with the Singapore Botanical Garden (SBG) regarding permissions to digitize SBG publications. Nancy E. Gwinn, attending the 2013 International Federation of Library Associations (IFLA) meeting in Singapore me with Elaine Ng (Chief Executive Officer, National Library Board) to coordinate contacts in Singapore. Ng saw the BHL model as something Singapore could contribute to and organized a larger meeting involving other Singapore institutions including the SBG, the National Archives, the National University of Singapore, and Lee Kong Chian Natural History Museum. Following on these discussions, BHL Program Director traveled to Singapore in February 2014 to finalize the memorandum of understanding with Singapore which was signed by the BHL Singapore representative, Eric Chin (Director, National Archives, Singapore).

BHL Mexico (December 2014)

Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) was a long time

participant in the Encyclopedia of Life (EOL). José Sarukhán Kermez (National Coordinator, CONABIO), while attending an EOL meeting in Washington in 2014, met with BHL Chair Nancy E. Gwinn and BHL Program Director Martin R. Kalfatovic to outline methods of participation of CONABIO and other Mexican institutions in the BHL. The outcome of these discussions led to a workshop held at the CONABIO offices in Mexico City during the first week of December 2014. This workshop led to a clarification of how CONABIO could work in the BHL context and mobilize Mexican biodiversity resources (Smithsonian Libraries, 2015).

Technology

The BHL portal is located at the Missouri Botanical Garden (MBG) and runs under a .Net environment. BHL offers various web APIs, downloads, and other services for export and download of BHL data and content. BHL partners with the Internet Archive for file staging and storage. Additionally, complete copies of the BHL scan data is held at the Smithsonian Institution and the Bibliotheca Alexandrina; a partial copy is stored via the BHL Europe project in London. Additionally, BHL has developed software service, MACAW[1], that simplifies ingest of partner content into BHL. MACAW is an open source tool with all code available on GitHub[2]. MACAW performs three major tasks in the process of digitizing a book-like item:

- Import and management of the images from the scanner, camera, or existing files
- Collection of the page-level metadata that describes the physical aspects of the page
- Post-processing and exporting/uploading the digital book to other systems.

Content

BHL focuses on the core literature of biodiversity that supports taxonomic research. Additionally, a wider group of materials from the areas of agronomy and other tangential areas (e.g. geology, anthropology) are included in the BHL corpus.

Biodiversity-relevant content

At its core, the BHL is focused on the subject matter relevant to the work of zoologists, botanists, evolutionary biologists, taxonomists, systematists, ecologists, natural history collections managers, scientific illustrators, biological science historiographers, and amateur scientists & hobbyists. This representative subset of BHL subject areas shows core and supporting subject matter. When deciding what to scan from our own collections, we try to maximize our scanning dollars by prioritizing the core literature, especially those titles that have a high concentration of taxonomic names. Content deemed to be irrelevant to biodiversity studies is subject to removal according to our Deaccession Policy.

Collection boundaries

To serve a wide, interdisciplinary audience, the boundaries of the BHL collection are deliberately inclusive. As E.O. Wilson once stated, "Biologists are inclined to agree that it [biodiversity] is in one sense everything," (Reaka-Kudla, et al., 1997) which is why content at the periphery of the core

1 https://macaw.library.si.edu/
2 https://github.com/cajunjoel/macaw-book-metadata-tool

taxonomic zoological and botanical literature is present in the collection. In the case of early taxonomic literature, core content was, at times, published in journals of a broader scope than is the practice today. Rather than review journal volumes piece by piece and scan items in isolation, the BHL seeks to provide the complete set of volumes for any given title. In addition, the BHL incorporates open access content scanned by other digital library projects to supplement its collection and enrich the range of content available for use through its taxonomic name finding and data export services.

Date range of content

The bulk of material in the BHL is in the public domain in the United States, meaning that the publication date is prior to 1923. Materials published after 1922 are available in the BHL for one or more of the following reasons:

- BHL has received explicit permission to provide the content online from the copyright holder;
- U.S. Federal government publications are in the public domain;
- Works for which the copyright was not renewed, according to the Stanford University Copyright Renewal Database;
- Works made available via open access repositories such as the Internet Archive.

Types of materials

The BHL provides access to a range of scholarly and general science materials. Content in the BHL collection consists primarily of books and journals, but a growing number of articles are appearing in the collection as a result of the work of BioStor. As of April 2015, there are nearly 150,000 articles indexed and searchable in the BHL collection. Article information harvested through BioStor.org is being indexed within the BHL collection on a regular basis. Try our Advanced Search options and select the "Articles/Chapters" tab to find articles. Additional article level access is possible by locating the journal title and navigating to the appropriate volume and page.

Materials from third-party websites may be indexed within the BHL so as to provide a single point of access to the world's biodiversity content. In these cases, BHL will not include the full text within its repository, but link out to selected content in external trusted repositories. It is preferred that full-text content be deposited in BHL via the Internet Archive and served through the BHL's website. The tools and services that make BHL unique, such as taxonomic name finding services, can only be used if the content is served through the BHL site.

Decisions to link to content on third-party websites are exceptions and will be made on a case-by-case basis when it is not possible to deposit content within the BHL repository. For example, see content from our Brazilian partners, BHL SciELO.

Coverage

The BHL seeks to provide the most comprehensive collection of legacy botanical and zoological taxonomic literature possible. The foundation of the BHL collection is based on the collections of its library consortium. See BHL Consortium Membership for more information about the BHL

participating institutions. A concerted effort has been made to provide all public domain content held within the general collections of BHL consortium libraries so long as the materials fall within the scope of biodiversity relevant subject matter. Every effort is made to avoid duplicating the digitization of like materials among BHL consortium library holdings. Duplication may be deliberate in cases where different copies of a book each have unique features, such as annotations. If duplication is accidental, items may remain in BHL until such time as proper review and de-accessioning takes place.

Content held within special collections or rare book collections is made available when possible. As rare materials cannot be shipped to a scanning facility for digitization, only those libraries that have in-house scanners are able to contribute such materials to the BHL collection. Often, the condition, size, or physical location of rare materials precludes the ability to make these materials available in digital form.

In addition to content contributed from BHL consortium libraries, the BHL collection is supplemented by:

- user-requested content accepted through the BHL feedback form
- open access biodiversity relevant materials, already in digital form, as made available by other digital library projects and scanning partners, such as the Internet Archive.
- in-copyright titles for which obtain permission to provide in our collection.

More information available on the BHL Collections Wiki Page (Biodiversity Heritage Library, 2015).

Service and outreach

The BHL has an active outreach program managed by Grace Costantino. BHL maintains standard social media outlets. The complete Outreach and Communications plan is attached. Key Goals of the Outreach plan are:

- Outreach Goal One: Increase awareness about BHL and grow our audience;
- Outreach Goal Two: Keep people informed about project developments, services, and how to use them;
- Outreach Goal Three: Raise awareness and foster dialogue within BHL audiences about biodiversity topics and corpus publications using BHL collections.

The audiences for BHL Outreach are:

- Scientists/Researchers
- Citizen Scientists/People Interested in Biodiversity
- Bibliophiles
- Artists/Art Historians
- Librarians
- Taxonomists
- Techies
- Educators

Conclusion

The Biodiversity Heritage Library has been a successful model for digital library development and collaboration for a variety of reasons. Chief among these has been the BHL focus on a specific use case, namely, providing access to a core constituency of taxonomists. Expanding beyond this core group in carefully planned stages (e.g. those interested in the visual images within BHL) and partner growth (e.g. the global growth via targeted institutions) has broadened both the participant and user base. Likewise, the integration of BHL into core functions of its participating institution has put BHL on the path to financial sustainability. BHL looks to the future with planned and sustainable growth, enriched content, and appropriate services.

References

Biodiversity Heritage Library (2015). Collection Development Policy (2015). Retrieved at: http://biodivlib.wikispaces.com/Collection+Development+Policy

GBIF (2014). Biodiversity Heritage Library joins GBIF as associate. http://www.gbif.org/page/7790

Marcum, Deanna. (2013). Biodiversity Heritage Library: Ithaka S+R Study. Retrieved at: http://www.sr.ithaka.org/sites/default/files/reports/SR_BHL_20140129.pdf

Reaka-Kudla, M.L., D.E. Wilson, and E.O. Wilson, eds. (1997). *Biodiversity II, Understanding and Protecting Our Biological Resources*. Washington, Joseph Henry Press.

Smithsonian Libraries (2013). Press Release. Biodiversity Heritage Library for Africa Launches. Retrieved at: http://library.si.edu/sites/default/files/pdf/general_pages/2013BHLAfrica.pdf

Smithsonian Libraries. Press Release. Biodiversity Heritage Library Adds CONABIO as New Member (2015). Retrieved at: http://library.si.edu/sites/default/files/pdf/general_pages/bhlconabiorelease.pdf