

User: [REDACTED]

Please review the [Application Guidelines](#) for important details regarding the information requested in this application. Fields marked with an asterisk are required.

---

## Project Summary

Cover Sheet	CoverSheet_CLIRProposal.pdf
Applicant Institution	Smithsonian Institution
Collection/Project Title	Biodiversity Heritage Library Field Notes Project
Project Summary	<p>The Smithsonian, Internet Archive, Missouri Botanical Garden Raven Library, American Museum of Natural History, Yale Peabody Museum, Harvard University Herbaria Botany Libraries, Harvard Museum of Comparative Zoology Mayr Library, UC Berkeley Museum of Vertebrate Zoology, New York Botanical Garden Mertz Library, and Field Museum, seek to increase accessibility to original scientific documentation through the digitization of entire collections of archival field notes.</p> <p>Currently, field notes from related persons and expeditions are scattered across institutions, inaccessible to any but the determined researcher. By enabling the complete, online collocation of collections, the Project will significantly improve research opportunities for scholars with interests as diverse as climate change, evolution, history of science, and women and minorities in science.</p> <p>The Project will coordinate work to digitize field notes, assign metadata, and publish the field notes online through the Biodiversity Heritage Library and Internet Archive, with an emphasis on quality, quantity, and closely related content.</p>
Amount Requested	491713
Project Start Date	01/01/2016
Project End Date	12/31/2017

---

### Is this a collaborative project?

If so, include the names of the collaborating institutions below. Use the green **add** button to list additional partners as needed.

Yes/No	Yes
--------	-----

## Collaborating Institutions (if applicable)

Institution

Missouri Botanical Garden Peter H. Raven Library

Institution

American Museum of Natural History

Institution

Yale Peabody Museum

Institution

Harvard University Herbaria Botany Libraries

Institution

Harvard Museum of Comparative Zoology Ernst Mayr Library

Institution

UC Berkeley Museum of Vertebrate Zoology

Institution

New York Botanical Garden LuEsther T. Mertz Library

Institution

Field Museum

Institution

Internet Archive

---

**Materials to be digitized**

Enter below the estimated quantities and select the units of measurement and material types that most specifically describe the extent of original materials that will be digitized during the project. You may add as many different measurement/material types as you like by clicking the green **add** button found below this section, but each individual item should be accounted for in only one category.

If the quantities provided are rough estimates rather than precise descriptions, explain the method used for estimating those quantities in the space provided for additional information.

## Materials

Amount of Materials	460000
Unit of Measurement	pages
Type of Materials	mixed archival collections
Type of Materials: Other	

## Additional information (optional)

Based on the experience of the partners, one "field book" averages approximately 110 pages. The number of field books could be determined from catalog records and across all partner institutions came to approximately 4135 items. Some partners knew exact page numbers for each item, in which case those were used. For partners who had to estimate page numbers, the estimate used the average baseline of 110 pages per item.

**Quantities, Formats, and Specifications of Digital Files to be Created during the Project**

Enter estimated quantities of uniquely described digital files to be created through digitization, as well as the relevant digital format(s) created and technical specifications for those formats (dpi, minimum pixel dimensions, bit-depth, optical density, etc.).

If additional files are to be derived from those created in the digitization process for the purposes of backup, preservation and/or access, do not count these derivative files or formats in the totals entered; you may describe any derivative formats to be created and the purposes these will serve in the space provided for additional information.

## Digital Files to be Created

Quantity	499150.00
Format	TIFF, JPEG2000
Technical specifications	300 ppi, RGB

## Additional information (optional)

The project will create approximately 225,450 image files in TIFF and 273,699 image files JPEG2000 format for a total of 499,150 images files created (from which 499,150 image files in JPEG will be derived for access). The reason two types of master files are created is because of the different types of scanning equipment used by project partners - some create TIFF, some create JPEG2000, both of which are considered preservation-quality formats.

**Catalogs/repositories/services**

Provide names and complete URL(s) for the portals through which digitized content will be available to researchers and the general public.

## Portals

Portal

Biodiversity Heritage Library

URL

<http://biodiversitylibrary.org>

Portal

Internet Archive

URL

<https://archive.org/details/biodiversity>

Portal

Smithsonian Collections Search Center

URL

<http://collections.si.edu>

Portal

Botany Libraries, Harvard University Herbaria

URL

<http://huh.harvard.edu/pages/digital-collections-0>

Portal

Ernst Mayr Library of the MCZ at Harvard

URL

[http://library.mcz.harvard.edu/ernst\\_mayr/digital](http://library.mcz.harvard.edu/ernst_mayr/digital)

Portal

LuEsther T. Mertz Library, NYBG

URL

<http://mertzdigital.nybg.org/>

Portal

Yale Peabody Museum

URL

<http://drs.library.yale.edu/fedoragsearch/rest>

Portal

AMNH

URL

<http://digitallibrary.amnh.org/dspace/>

Portal

Field Museum

URL

<http://fieldmuseumlibrary.worldcat.org/>

Portal

Missouri Botanical Garden

URL

<http://www.botanicus.org>

## Description of Content: Public

### Description of materials

Provide a brief narrative description of the materials nominated for digitization, including their provenance, relevant associated people, organizations, events, and/or subjects.

#### Description

Materials to be digitized include bound notebooks, loose field notes, unpublished manuscripts, collections of correspondence, and photographs and photo albums related to the last 200 years of scientific expeditions carried out by staff employed by or associated with the partner institutions. All field notes collections were initially created by scientists employed by the holding institution or in partnership with the holding institution during the course of their careers. Upon retirement, scientist's field notes were retained by their employing institutions, under archival control for description and preservation purposes, as reference materials for the expeditions, specimens, organizations, persons, and scientific observations they recorded. Most are still actively used by current curators and researchers. Each is a unique and valuable document that may be inaccessible to the casual researcher. Digitization and online, digital publication will bring these valuable documents to the general public for the first time, enabling vastly expanded types and amounts of research. Furthermore, due to changing employers and affiliations during their lifetimes or to expeditions which employed many scientists who, upon their return, dispersed back to their own institutions, collections related by a common creator or expedition may be scattered amongst many institutions. Digitization allows these collections to be reunited online in a way impossible to replicate physically, critical for any researcher interested in the full context of the scientific expeditions, discoveries, or collections described therein. Project partners engaged in scanning will digitize larger or smaller amounts of material, depending on the size of their collections, with an emphasis on the digitization of related collections, and partners with smaller collections will provide additional support assigning metadata and ingesting digitized collections into the Internet Archive and Biodiversity Heritage Library.

**Geographic Scope**

Describe the range of geographic regions represented in the nominated collection(s).

Geographic scope

Materials to be digitized have an international reach, and are known to cover North, Central, and South America, as well as the Caribbean, China, Indonesia, and the Arctic and Antarctic. Materials may provide additional coverage of Europe, Africa, Asia, and Australia/New Zealand.

**Date range of materials to be digitized**

Enter the earliest and latest dates the original materials in the nominated collection(s) were created, in whole years. Dates should be formatted as YYYY BC/AD – YYYY BC/AD (e.g. 356 BC - 1542 AD).

Date range of materials

1815 AD - 2000 AD

**Collection level descriptions (if applicable)**

If applicable, identify and provide the URL(s) for any collection-level descriptions currently available online.

## Collection URL(s)

Collection name(s)

Smithsonian: National Museum of Natural History, Pacific Ocean Biological Survey Program, records, circa 1961-1973, with data from 1923

Collection URL

[http://collections.si.edu/search/results.htm?q=record\\_ID:fbr\\_coll\\_NCDC33](http://collections.si.edu/search/results.htm?q=record_ID:fbr_coll_NCDC33)

Collection name(s)

Botany Libraries, Harvard University Herbaria

Collection URL

<http://id.lib.harvard.edu/aleph/000603352/catalog>

Collection name(s)

Ernst Mayr Library of the MCZ at Harvard

Collection URL

<http://id.lib.harvard.edu/aleph/000605398/catalog>

Collection name(s)

LuEsther T. Mertz Library, NYBG

Collection URL

<http://willow.nybg.org/search/?Xfield+notebooks&SORT=/Xfield+notebooks>

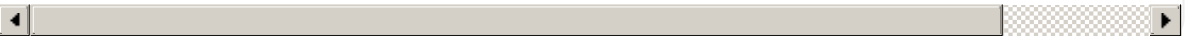


Collection name(s)

Yale Peabody Museum

Collection URL

<http://findingaids.library.yale.edu/HLTransformer/HLTransServlet?stylename=yul.ead2002.xhtml.xsl&pid=ygm:entar.000001&query=remi&stylesheet-cache=yes&hlon=yes&filter=&hitPageStart=1>



Collection name(s)

AMNH

Collection URL

<http://libcat1.amnh.org/search~S0?Xwhitney+rollo+beck&SORT=D/X>



Collection name(s)

Missouri Botanical Garden

Collection URL

<http://www.biodiversitylibrary.org/collection/engelmannpapers>

## Description of Content: Confidential

### Current arrangement and description(s) of materials to be digitized

Provide a brief narrative that summarizes the physical arrangement and the level(s) of processing, cataloging, or other descriptive work that has previously been done for the nominated collection(s). Include the date(s) this descriptive work took place and the standard(s) and/or current format(s) of the records that were created.

#### Current Arrangement

Smithsonian: Collections are arranged by series, box, and folder, and have EAD finding aids and descriptive item-level MODS and MARC XML metadata, created from 2010-2015. Metadata is available online via Collections Search Center.

Harvard Botany Libraries: Collection is arranged by format. A FileMaker database provides item level access. The collection is discoverable in Harvard's online catalog, HOLLIS, and on the library's webpage.

MCZ Harvard: Collection is arranged by series, box, and folder, and has a collection-level MARC record with separate MARC records for items. All metadata is available online via Harvard's discovery service, HOLLIS.

NYBG: Collection is arranged chronologically by accession. Item-level cataloging in MARC began in 2014, and items have bibliographic records available in OCLC and the Library's online catalog.

Berkeley: Collections are arranged in series by individual and described at the item level. Finding aids are in the Online Archive of California. With support of a 2012 CLIR award, collections were entered in Archivists' Toolkit, and are currently being exported to the Library Catalog.

Yale: Collections are arranged by series, folder, and item and have EAD XML collection-level finding aids and metadata. Metadata is available online through the Yale Finding Aid Database. Item-level description occurred from 2012-2015.

AMNH: Collections are organized by box and folder, and arranged chronologically. Our current CLIR grant is allowing us to write finding aids and EAD metadata. Item-level description is taking place with support from the Leon Levy foundation.

Field: Collections will be cataloged in 2015 in preparation for 2016 digitization. We will use best practices from partners, such as the use of both collection and item level MARC, to guide this process.

MOBOT: Collections are arranged by series, box and folder and have collection level MARC records and folder level descriptions in an Access database. Arrangement and description took place from 1997-2001



**Current condition and housing of materials to be digitized and plans for their conservation and preservation**

Describe the current condition and housing of the materials to be digitized, including the means through which this condition has been assessed. Identify the individual or individuals responsible for this assessment and approximately when the assessment took place. Describe the strategies to be employed for stabilization, conservation, and/or preservation of the materials, including the means through which this work will be supported and sustained long-term. Explain the environmental provisions made for the long-term management of the source materials and the strategy for responding to requests for access to them.

**Current Condition**

Smithsonian: Collections are in stable condition for digitization, having been assessed and treated by conservation staff (Nora Lockshin) in the last five years. Collections are in stable archival and environmental housing with the Archives. All housing, storage, and conservation ensures that physical items are preserved for researchers. Once items are digitized, all efforts are made to have researchers use digital copies to further preserve the structural integrity of physical items.

Harvard Botany: Collections are in stable condition for digitization. Bound volumes have custom archival enclosures (from 2005) and unbound items are being re-housed in archival folders. Collections are located in secure, climate-controlled locations within the libraries. Ongoing conservation review ensures that physical items will be preserved for future researchers. Researchers are encouraged to use digitized copies rather than original materials to maximize their longevity.

MCZ Harvard: Collections are in stable condition for digitization, having been assessed and treated by Preservation Center staff (Christopher Sokolowski) in the last two years. All collections are in monitored archival and environmental housing in the Mayr Library Special Collections/MCZ Archives. All housing, storage, and conservation treatment ensures that physical items will be preserved for researchers. Once items are digitized, all efforts are made to ask researchers to use digitized copies instead of physical to further preserve their structural integrity.

NYBG: The field notebooks selected for this digitization project are in stable condition. All of them requiring conservation have been previously treated and/or rehoused by NYBG conservators supported by previous NY State grant. The conservation staff will review and prepare each item prior to being sent to be scanned.

Berkeley: In 2014, the MVZ received a National Endowment for the Humanities, Preservation Assistance Grant for Smaller Institutions. All materials were then assessed by conservator Kathleen Orlenko. All volumes have also been reviewed by Archivist, Christina Fidler and are suitable for scanning. Materials are housed onsite in the MVZ Grinnell-Miller Library and in the Archival collection space.

Yale: All collections have previously been conserved and prepared for digitization, and all are in stable archival and environmental

housing, either in the Peabody building complex on campus or at Yale's West Campus facility. Efforts are made to ask researchers to use the digital copies instead of the physical copies whenever feasible.

AMNH: The collections have been assessed and treated by AMNH conservation staff member Barbara Rhodes. All collections are in stable archival and environmental housing, either in the research library's rare book room or in the science department's archive rooms. Field: Collections are housed with the Botanical Collections Manager. They are in good condition for digitization, having been assessed by our collections manager. We have approval from Botany to transfer these items to the Library/Archives for permanent physical and digital stewardship prior to the start of the proposed project.

MOBOT: The notebooks are in stable condition for digitization. They are unbound in a range of sizes and shapes. All materials will be carefully handling by trained scanning technicians. Preservation evaluation will be conducted by our Library Conservator throughout digitization.

#### **Representative samples of materials to be digitized (max. 10 pages, 12 MB, .pdf format only)**

Upload a PDF document containing images of up to ten (10) selected items from the collection(s) to be digitized, and provide a brief description of the contents of the file, below. Each image should be accompanied by a description and full citation that includes the name of the holding institution, the collection title, any identification numbers or shelfmarks, and any available information about rights or licensing. The document may contain embedded URLs linking to additional content, such as sample audio or audiovisual files, but must contain samples of no more than ten items.

#### **Description of Samples**

MCZ, Harvard: Journals of William Brewster, ornithologist and Curator of Birds and Mammals. Field Museum: Timothy Plowman field notes containing a record of specimens collected in Peru in 1977. Harvard University Herbaria Libraries. Field notes of Henry Pittier documenting the collection of specimens in Venezuela in the 1920s. Smithsonian: The sample is an excerpt from a field book by Frederick True. It contains daily entries on seals and other wildlife on St. Paul Island, Kelaire, Lukannan, Polovina, Tolstoi, and St. George's Island, Alaska, 1895. NYBG: Field book documenting Percy Wilson's botanical collecting in Honduras in 1903, location "Puerto Sierra."

#### **Samples (.pdf format only)**

SampleImages\_CLIRProposal.pdf

#### **May CLIR excerpt from and display some portion of these representative samples within the Hidden Collections Registry, elsewhere on CLIR's website, or in program-related social media?**

If you select "Yes" but would prefer that some representative samples be displayed but not others, note below which specific samples CLIR has permission to share.

Yes/No

Yes

Selected Permissions (if applicable)

## Rights, Ethics, and Re-Use

All parties to this proposal understand that as a condition of acceptance of any *Digitizing Hidden Special Collections and Archives* award from CLIR, all recipient institutions and collaborating partner organizations will be required to sign and execute the program's [intellectual property agreement](#).

Tick to confirm: Confirmed

All parties to this proposal understand that as a condition of acceptance of any *Digitizing Hidden Special Collections and Archives* award from CLIR, all metadata created in the course of funded project activities must be dedicated to the public domain under a [CC0 Creative Commons license](#).

Tick to confirm: Confirmed

All parties to this proposal understand that as a condition of acceptance of any *Digitizing Hidden Special Collections and Archives* award from CLIR, recipient institutions, including partnering institutions in cases of collaborative projects, must not claim additional rights or impose additional access fees or restrictions to the digital files created through the project, beyond those already required by law or existing agreements. Materials that are in the public domain in analog form must continue to be in the public domain once they have been digitized. CLIR strongly encourages grant recipients to share digital copies as public domain resources and/or with Creative Commons licenses, as appropriate.

Tick to confirm: Confirmed

*Applicants who tick any of the four boxes below should provide details clarifying their responses in the Rights, Ethics, and Re-Use Statement.*

Tick all that apply:

### Rights, Ethics, and Re-Use statement (max. 3 pp, 2MB, .pdf format only)

Upload a description (maximum 3 pages) that summarizes all known rights, embargoes, and access or legal restrictions applicable to the source materials to be digitized and describes how these rights, embargoes, or restrictions will be communicated to the public. Identify and explain any ethical considerations that affect access to or re-use of the digital copies. Explain the basis upon which the proposed activities are justifiably legal and ethical, and the specific terms under which users of the collections will be able to access and re-use the digital copies created through the project. Explain and justify any institutional watermarks incorporated into copies made accessible to users and any fees charged for re-use. Describe any other measures to be taken to restrict access to or re-use of the digital copies in order to comply with the law or with applicable, pre-existing agreements or contracts.

Statement (.pdf format only) BHL Copyright Policy.pdf

## Value and National Significance

**Describe the impact of the proposed project upon scholarship, and explain why digitization is the most appropriate means to maximize the value and significance of the materials to scholars and students.**

Address the importance of the collections to teaching, research, and the creation of new knowledge.

### Value and National Significance Statement

Field notes are the primary source documents for the history of scientific research and discovery. They can provide invaluable, unpublished information about scientific expeditions, landmark scientific publications, and natural history museum collections. They can also provide rich data for researchers to understand how biodiversity has changed over time and space. Field diary entries may describe habitats, meteorological events, personal observations, and emotional declarations. Furthermore, they can enhance information associated with specimens by providing details regarding dates, localities (for geo-referencing), and associated event data. However, these documents, and the specimens and publications they documented and supported, are often dispersed within and between institutions across the United States. This geographic dispersal, combined with the fact that field books may be buried in larger archival collections and, to the public or to scholars from other institutions, often difficult to find at the item-level, makes discovery, access, comparison to related field books and the resulting publications, and other scholarship very difficult.

Collaborative digitization and digital publication offers the only opportunity to reunite complete field book collections from many national institutions. The project would reunite field research documentation online in way impossible to replicate physically. It would allow searching across collections and within collections to a far greater level of detail, detail that allows richer research than a simple finding aid in an online catalog would enable. It would also enable scholars to reunite primary source documentation with the scholarship it produced as both become collocated on the BHL, and enable new critical scholarship around landmark historical publications as well as an improved understanding of historic natural history collections across the United States. The additional data made available through digitization and online publication would allow researchers to use the information contained therein in new and better ways, such as assessing the value and condition of specimens, reconstructing historical ecologies, clarifying specimen and article provenance, and re-discovering species and study localities.

These discoveries in turn transform researcher's understanding of many topics of interest. For instance, a climate change researcher might be able to use field books related to historic biodiversity surveys from multiple institutions and eras, formerly dispersed

among the source archives, to map out the geographic range of a species over time to see how it might have changed to adapt to a changing landscape. A history researcher might be able to search across digitized field books for personal names unrelated to the author to pinpoint, for instance, the presence of native or female voices in scientific history. Finally, a taxonomic researcher might be able to find, at last, the original source for the naming of a species and trace the transformation of the taxonomic classification over time to map how the understanding of the species has developed as scientific knowledge has advanced.

**Upload three letters of scholarly support for your project (max. 10MB, .pdf format only).**

Letter of Support #1	LetterOfSupport_Wood.pdf
Letter of Support #2	LetterOfSupport_Nelson.pdf
Letter of Support #3	LetterOfSupport_Lahmeyer.pdf

## Project Context and Impact

**List and describe all envisioned project deliverables. Explain the means through which each will be available to the public, and any applicable conditions or terms limiting their availability.**

Describe all expected outcomes, how each will be made accessible to others, and under what conditions. Deliverables will include the digital surrogates created during the project and related metadata, and they may also include aggregations of those files and metadata with related collections, catalog records, finding aids, authority files, description and digitization manuals, training materials, new or improved software tools, reference guides, or other outcomes of the proposed project. Address any relevant restrictions or licensing terms not already explained in the intellectual property statement.

Deliverables	<ul style="list-style-type: none"><li>•Metadata: Approximately 4135 items assigned any additional structural, administrative, and technical metadata appropriate for digitization (the majority of the items to be scanned already have item-level description but will require the assignment of structural, administrative, and technical metadata at the page level as they are scanned). All metadata will be converted to a web-ready format and published in a full and open manner on the Biodiversity Heritage Library and Internet Archive.</li><li>•Digitization and Access: Approximately 460,000 pages digitized as high quality TIFFs and JPEG2000, embedded with the appropriate technical and structural metadata, processed into reference-ready PDFs for local requests and web-ready JPEGs for online publication, and delivered online in a full and open manner via the Biodiversity Heritage Library and the Internet Archive, along with all item-level metadata.</li><li>•Digital Preservation: Approximately 460,000 master TIFFs and JPEG2000s stored in secure, geographically distributed, closely monitored digital asset management systems for perpetual preservation and access. The master images may be used to re-create the web- and public-deliverable JPEGs and PDFs if</li></ul>
--------------	--

necessary. An additional copy of all items that go into BHL will be maintained on Smithsonian Libraries' Isilon storage servers.

- Project Documentation: Project workflows, metadata mapping, and digitization documentation released in a full and open manner on BHL public wiki for reuse by other institutions interested in digitizing archival materials at the item-level. May include tutorials for other BHL partners to assist in bringing additional archival items into BHL. May include documentation of other projects, such as transcription of field notes materials and scholarship performed from transcribed field notes content.

**Describe all prior initiatives, experiences, or research activities that have directly informed the design of the project, including any notable innovations or practices that will make the proposed approach particularly efficient, ground-breaking and/or cost-effective.**

#### Project Context

Through the Smithsonian Field Book Project (initially funded by a CLIR Cataloging Hidden Collections grant), the Smithsonian has previously completed five years of field book discovery, cataloging, preservation, and in the last year began testing digitization, directly informing the ability of Smithsonian staff to estimate the care, metadata, technical specifications, and workflows necessary to digitize and publish online archival field book materials. So far, the Smithsonian has successfully cataloged and conserved over 7500 items, and has digitized, created item-level metadata, published online through the Smithsonian Collections Search Center and BHL 548 items, giving them extensive experience with the pace, cost, and design of item-level archival digitization projects. Furthermore, through the prior Connecting Content Project, many of the partners on this BHL Field Notes Project have experience in the cataloging, description, related collections, and other unique challenges and opportunities presented by field books. As a result, the partners have full descriptions for their collections and an awareness of their content and digital imaging needs that prepares them for a digitization project of this magnitude.

Additionally, the workflows developed by Smithsonian through the Field Book project, including the use and timing of the Macaw technology and Internet Archive processing pathway, reduce time and increase digitization throughput capacity by 300% given dedicated staff and equipment, an increase that will be passed along to other partners, making the archival digitization process for this project maximally efficient and cost-effective. The project partners will adopt each other's best practices for scanning quality and metadata format and assignment to better strengthen the coherence of their local practices and to align with the broader community standards. The project will also track digitization work across partners to better improve workflows for digitization and metadata production.

This project will aggregate metadata for related field book collections across all eight partner institutions using the Internet Archive and BHL repositories and web platforms, describing all of the items in the widely used and easily mapped standard MARC XML and offering the metadata with an open API key on BHL to fit with best

professional practices for interoperable and re-usable metadata. It will expand the BHL to include, for the first time, a significant and dedicated volume of primary source materials alongside existing secondary sources. It will strengthen the coherence of local efforts to scan archival materials at the item level to ensure that these unique materials can be promoted for use by the public in a digital, online way that protects the original items from over-handling or misuse.

Furthermore, all partners involved in the proposed project have previously received training on the Macaw and Internet Archive import process, or will receive training prior to the start of the proposed grant, and will therefore be familiar with the procedure. Partners will also work to discover and share ways in which the created digital files may be embedded into their original finding aids to improve discovery and contextual understanding.

**Describe all outreach and marketing activities planned to maximize the impact of the project, for both scholarly and professional communities of interest.**

Outreach and Marketing

The project will follow the recommendations of the broader BHL Outreach and Communications Plan. The first objective of the BHL Field Notes Project outreach plan will be to increase awareness of the project and grow the audience for the resulting digital materials, which will be accomplished by coordinated interfacing with various audiences through social media outlets and listservs, attending and presenting at relevant conferences and meetings such as the Society of American Archivists (SAA), Association of College and Research Libraries (ACRL), Society of the Preservation of Natural History Collections (SPNHC), and Biodiversity Information Standards (TDWG), and the creation of content highlighting the work of the project and its scholarly value, to be distributed on the project blog, BHL blog, and related institutional blogs and websites. The second objective will be to keep audiences informed about project activities, which will be accomplished by the publication of quarterly project updates on the various partner blogs and newsletters, maintaining an online bibliography of publications and presentations given by project staff about project work, and highlighting major project news and innovative breakthroughs with articles and presentations for appropriate outlets. The third objective of the communications and outreach plan will be to foster dialogue with project, BHL, and scientific and professional communities about biodiversity topics and research opportunities highlighted or made possible by the project, which will be accomplished by creating a calendar of events throughout the year that will be used to organize outreach events and identify social media topics, featuring creators, collections, image highlights, and biodiversity news topics in social media coverage while highlighting project content related to those topics, encouraging crowdsourcing and citizen science activities that enhance the project collections, and featuring scientists currently working with project collections on partner institution blogs, newsletters, and social media outlets.

**Describe any collections related to the materials nominated for digitization and describe any plans for creating meaningful linkages to those collections.**

Applicants should be as specific as possible in naming these related collections. Related collections may be held by the applicant or partner institutions, or by other institutions not participating directly in the project.

**Related Collections**

The goal of the project is to digitize and create meaningful linkages between field book collections at multiple institutions that relate to each other through creators, expeditions, geographies, time periods, and species covered. Often, collections that connect to each other through authors, correspondence, expeditions, or species and specimens covered have become dispersed over multiple departments within each partner institution and even between geographically distributed institutions as staff and specimens changed locations. A primary goal of this project is to reconnect digitally, through publication on BHL, collections which are thus related but physically separated. Meaningful linkages will be created by aggregating the metadata for these related collections on the well-known research portal, BHL, along with their digital images; by adopting the common MARCXML standard for describing the items and publishing on BHL; and by partnering with BHL to promote the use of this collection as soon as it begins to populate (that is, as soon as items begin to be digitized). These collections will also relate to any previously digitized field book collections from each institution by virtue of the extensive overlap between creators and research areas. Please see the Smithsonian Field Book Collection on BHL for related field book collections already in BHL. Related collections may also be discovered through the publication of items on the Smithsonian Transcription Center, where crowdsourced transcription has led to discoveries of new connections across items. An additional related collection is that of the published scholarly articles that emerged from the documented field work now available on the Biodiversity Heritage Library. Due to the ability of BHL to digitally publish the original items alongside their related papers, part of the project will be the ability for researchers to search a name or expedition and see the links between published articles and archival documentation on BHL.

**Describe any future initiatives that would be made possible by the completion of project work.****Future Initiatives**

Completion of project work will make available for scholars a broad selection of the primary scientific source notes for research and study alongside related publications and images in the Biodiversity Heritage Library. The volume of field notes scans is also anticipated to enable a transcription and data discovery initiative, whereby the field book images will be entered into various transcription platforms (such as the Smithsonian Transcription Center), the completed transcriptions imported alongside the original field book images in BHL, OCR'd for full-text searching, and then exported again into projects similar to Purposeful Gaming and iDigBio to experiment with computer extraction of specimen information into tables,



spreadsheets, relational and semantic databases, and visualizations for increased analysis and comparison to contemporary survey and specimen data. It is anticipated that this future project will provide unprecedented support to scientists studying biodiversity, geology, species growth, expansion, depletion, and extinction, and other ecological issues over time.

## Project Design

**Upload a project plan with timeline that includes all major project activities and deliverables, including a project timeline with deliverable deadlines (max. 3 pages, 2MB, .pdf format only).**

The timeline for the project should be as explicit as possible, identifying major activities to be undertaken during each quarter of the proposed grant term and naming the parties who will participate in those activities.

Project Plan (.pdf format only)

BHLProject\_Design\_ProjectPlan.pdf

### **Project plan appendix (max. 2MB, .pdf, .xls or .xlsx format only)**

Upload a project plan appendix\* specifying in greater detail the collections to be digitized for the project, the relevant holding institution, and the partners who will be responsible for digitization, description, and the long-term maintenance of the digital files that will be created (no page limit, .pdf, .xls or .xlsx format only).

**NOTE: The appendix must follow CLIR's template, which can be found [here](#). Non-compliant appendices will be removed prior to review and may result in the application being rendered ineligible for review and funding.**

\*This document is required for collaborative projects; optional for all other projects.

Project Plan Appendix (.pdf, .xls, .xlsx)

CLIR\_ProjectPlanAppendix.xlsx

### **Technical Plan (max. 4 pages, 5MB, .pdf format only)**

Upload a document detailing your proposed technical approach, including a description and justification for the proposed workflow that clearly identifies all tools, systems, standards, and technologies to be employed in the project.

Technical Approach (.pdf format only)

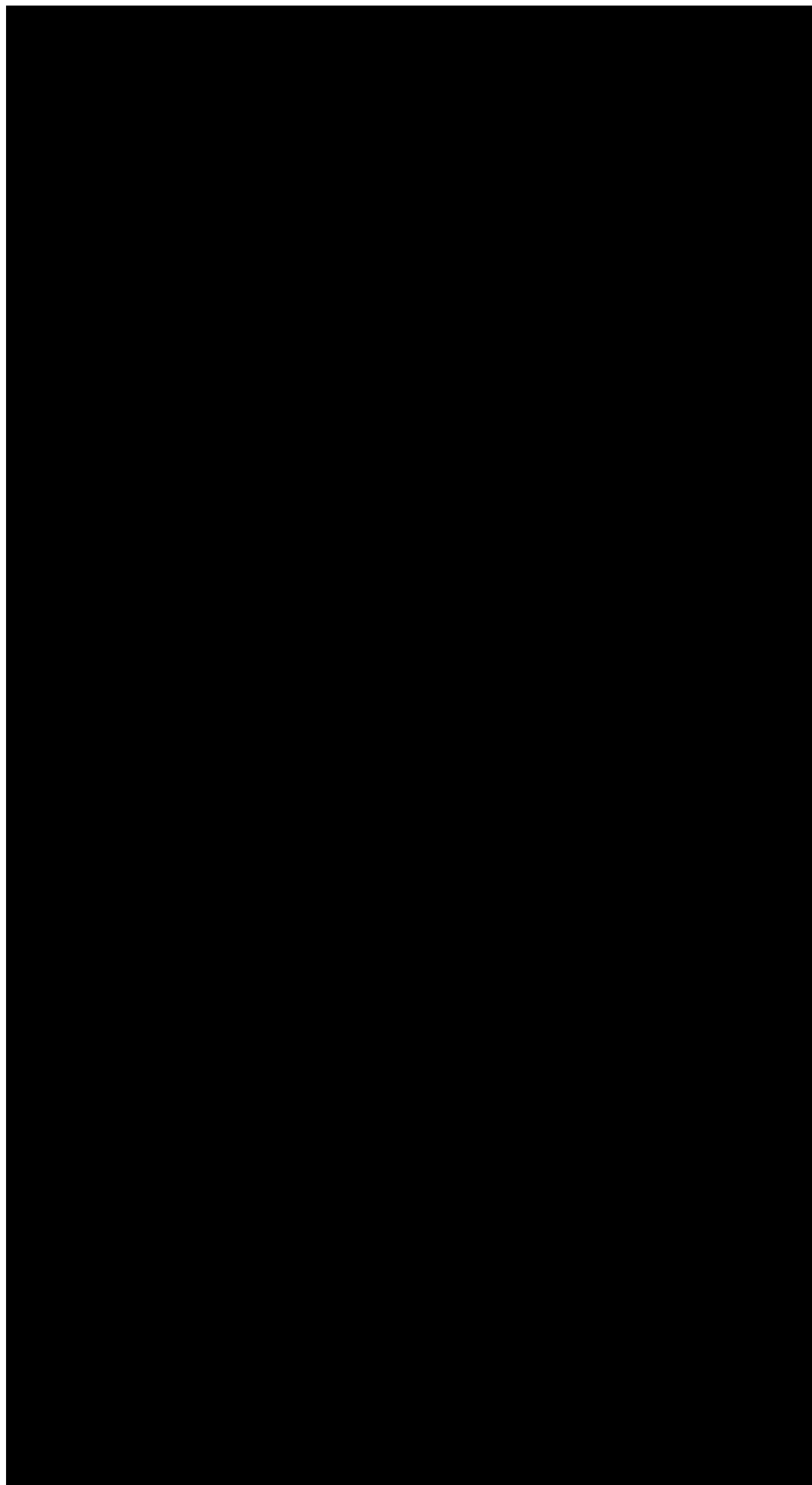
BHLProject\_Design\_TechPlan1.pdf

**Principal Investigators/Primary Staff**

Briefly describe the relevant qualifications of up to three individuals who will contribute substantially to the project. The qualifications of all named Principal Investigators (PIs) must be included here. If the project includes fewer than three PIs, applicants may optionally use this space to describe other important staff members' qualifications.

Upload resumes/CVs for these individuals below (.pdf, .doc, or .docx format only). In the event that a person holding major responsibility for the project has not yet been identified, applicants may upload a job description for that unidentified person's role.

Staff Qualifications



Resume #1 (.pdf, .doc,.docx)

Resume #2 (.pdf, .doc,.docx)

Resume #3 (.pdf, .doc,.docx)

### Staffing

How many staff will be assigned to this project? You may include students and volunteers in this list.

Staffing

Staff assigned to the project and funded by the grant include a Project Manager, Coordinator, and 6 scanning techs. Percent time per position varies by project year. Staff working in an advisory capacity, providing assistance in planning, scanning, processing, preservation, and publication, include one project coordinator from each partner and a series of advisers for a total of 17. All staff members are listed below. Technical staff remain unassigned pending hiring after project is funded.

### Will special skills or training be required?

Explain the nature of any required skills or training to undertake the project and how the applicant institution intends to solicit or provide it.

Training

All staff assigned and/or hired specifically for the project have or will have prior familiarity working with all assigned standards, equipment, and technologies. Currently, training materials on project technologies (Macaw, IA, BHL) are being developed as part of the BHL Africa Project, and all project partners and staff will complete training as needed with these materials prior to the start of the grant period. No additional training will be required during the grant period.

## Sustainability

### Digital Preservation and Discoverability Plan (max. 2 pages, 2 MB, .pdf format only)

Upload a digital preservation and discoverability plan explaining how project deliverables will be made secure and discoverable for the long term.

Preservation/Discoverability Plan  
(.pdf format only)

BHLProject\_DigPresAndDiscoverPlan1.pdf

## Institutional Capacity

Upload a letter of support from the head administrator of the applicant institution.

Letter of Institutional Support

InstitutionalLettersofSupport1.pdf

## Institutional Strengths

Describe the institutional strengths that justify the undertaking of the proposed project by the applicant (and any collaborating partners). Strengths may include existing infrastructure, partnerships, professional associations, staff experience, access to local expertise (scholars, volunteers, students), financial or other resources, etc.

### Institutional Strengths

The Smithsonian has prior experience in digitizing large quantities of archival materials at a rapid pace without compromising quality. The Smithsonian Libraries and Archives both have full-time scanning facilities, staff, and workflows both on-site and off-site at the Smithsonian Institution Research Library Annex (SILRA) that are used to contribute material to the Biodiversity Heritage Library and Smithsonian Collections Search Center. Custom software (Macaw) serves these workflows and will also be used to support the workflows for the proposed BHL Field Notes Project. Staff expertise will also be used to inform the design of the BHL Field Notes Project workflows, and will be on hand to assist the dedicated project staff in case of any unforeseen slowdowns during the process. Staff at all partner institutions have at least five years or more experience as well as related infrastructure for managing any and all project operations, from item level metadata assignment to scanning to digital publication, and the majority of staff belong to professional associations such as the Society of American Archivists and Society for the Preservation of Natural History Collections that give them a network of colleagues to reach out to should any question need additional insight. Furthermore, the Smithsonian and the proposed project partners have prior experience working together on the scanning and full and open online publication of digitized items through their partnership in the Biodiversity Heritage Library. The Smithsonian and partners also have access to a steady stream of graduate student interns and community volunteers with demonstrated interest and ability to assist in portions of the digitization workflow, such as the physical handling of the materials and QA of the metadata. Partners have also agreed to share expertise and even scanning facilities across institutions, offering existing scanning facilities and digital publication (metadata prep, assignment, quality review, BHL ingest support) to partners who have materials, but no facilities of their own. The partnership with the Internet Archive provides additional support through their generous provision of their expertise and economies of scale in digitization and digital image and metadata processing, the use of their Tabletop Scribe digital imaging machines, and their in-kind IT support for the Biodiversity Heritage Library import process, taking on the work of processing master images and metadata into web-deliverable formats. Finally, each project partner is financially able to contribute the time of a local Project Coordinator to constantly monitor work for quality and efficiency, and to document project workflows for publication and the hopeful re-use by other institutions after the completion of the project.

**Institutional Priorities**

Describe the applicant's (and partners') institutional priorities for digitization, digital collection development, maximizing access, and supporting scholarship, learning, and/or the public good.

**Institutional Priorities**

Beginning with the Smithsonian Strategic Plan of 2015, and in addition its continuing mandate to work for the increase and diffusion of knowledge, the Smithsonian also operates under the four Grand Challenges, one of which is to advance cross-disciplinary, integrated scholarly efforts which support the goal of "Understanding and Sustaining a Biodiverse Planet." This Challenge encourages Smithsonian staff to advance research, broaden access, revitalize education, strengthen collections and encourage new ways of thinking that involve emerging technology, all of which would be addressed by the BHL Field Notes Project, where the Institution seeks with partners to expand access to and scholarly efforts related to the history of scientific understanding of biodiversity. Furthermore, the Smithsonian operates under an additional Digitization Strategic Plan that emphasizes the necessity of making highest-quality digital surrogates of museum objects, libraries, and archives available to the public to broaden access, enrich context, and support education, which are again specifically relevant to the BHL Field Notes Project, which seeks to follow that mandate with the Smithsonian and partners' scientific field note collections. Project partners share strategic goals that relate directly the BHL Field Notes Project' emphasis on digitization and digital publication of archival field materials to increase accessibility and scholarship. The Field Museum strategic plan emphasizes maximizing the accessibility of collections via digitization. The Internet Archive mission includes offering permanent access to researchers, historians, scholars, people with disabilities, and the general public by supporting digitization, conducting digitization, publishing online digital materials, and committing to long-term storage and preservation. The first priority in the Mertz Library at NYBG strategic plan is to make more resources available in electronically, and to push them beyond the physical walls of the Library into by undertaking major digitization and online publication efforts. The AMNH strategic plan offers a digital strategy that emphasizes increasing access through new technologies-using the web to open collections for research and learning to scholars, lifelong learners, and the general public. The Museum remains strongly committed to stewarding the irreplaceable legacy these collections embody, as well as making them widely available for research and study. Priorities in the Ernst Mayr Library of the MCZ include facilitating access to and preservation of resources in collections as well as providing teaching and learning opportunities and services to support the activities of Harvard University and the global community. The Harvard Botany Libraries share the strategic goals of the Mayr Library and also prioritize the digitization of archival materials based on two main criteria, use and condition. A core element of the Yale Peabody Museum's 2013 strategic plan is devoted to digitization and archive management. Using the Imaging Lab at the Yale Institute for Preservation of Cultural Heritage, the Peabody has developed work flows for

digitizing field notes and related materials, which is part of the Peabody mandate for the curation of original documentation. MOBOT also shares an institutional mandate to make their archival materials as widely available and usable to the broader professional and public communities through digitization and digital publication.

## Prior Initiatives

Provide up to three examples of prior initiatives that demonstrate preparedness of the applicant institution(s) to undertake project work.

Initiative #1	Connecting Content: Connecting Content involved the California Academy of Sciences, Missouri Botanical Garden, Smithsonian, Academy of Natural Sciences of Philadelphia, Harvard Botany Libraries, Harvard Museum of Comparative Zoology Library, and the New York Botanical Garden. This cooperative project improved access to primary source and published biodiversity research materials through cataloging and metadata to create linkages between collections that are physically dispersed.
Initiative #2	Smithsonian Field Book Project: The Field Book Project began in 2010 as a joint initiative to discover, catalog, preserve, and make available online the Institution's previously un-cataloged field book collections, initially supported by a CLIR Cataloging Hidden Collections grant. The project has cataloged and published online over 7500 field book records, and has recently begun to digitize field books ad-hoc to satisfy the increased reference requests from the professional and academic communities due to the presence of the catalog records online.
Initiative #3	Biodiversity Heritage Library: The Biodiversity Heritage Library was launched in 2007. It is a consortium of natural history and botanical libraries that successfully cooperate to digitize the legacy literature of biodiversity and to make that literature available for open access and responsible use. In partnership with the Internet Archive and through local digitization efforts, the BHL has digitized 45+ million pages of literature.

## Funding

### Budget Documents

CLIR requires all applicants to complete and upload two budget documents:

- Budget Narrative ([click here](#) for detailed information about the Budget Narrative). *No page limit, max. 2MB, .pdf format only.*
- Budget Detail (must be submitted using [CLIR's Excel form](#); [click here](#) for detailed information about the Budget Detail). *Max. 2MB, .xls or .xlsx format only.*

Budget Narrative (.pdf format only)      BHLProject\_BudgetNarrative.pdf

Budget Detail (.xls or .xlsx format only)      CLIRGrant\_Budget.xls

**Subcontracts (No page limit, max. 5MB, .pdf format only)**

Provide below any subcontracts for work associated with this project. Quotes from vendors may be provided in lieu of more formal contract documents as necessary, as long as the relevant work to be conducted and costs incurred are clearly delineated.

If your project will involve multiple subcontracts, combine them into one PDF document for upload into the system.

Subcontract(s)

**Funding Allocations**

Will funds from the grant be applied to current staff salaries or infrastructure? If so, justify the need to cover such costs with external funds and their relevance to the proposed project.

Yes/No

Yes

Justification

Funding for existing staff are to extend current soft money appointments in the case of the scanning technicians. Project Manager position is currently vacant at the Smithsonian. Quarter-time funding for Project Coordinator at Berkeley is to fund time dedicated to the project due to the large collections at this institution.

**Funding Justification**

Explain why this project is uniquely suited to this funding program and how it reflects the program's [core values](#). What prior efforts, if any, has the applicant made to identify other sources of funding for this project?

Justification

The BHL Field Notes project is uniquely suited to the CLIR Digitizing Hidden Collections program because of its size, scope, and emphasis on enabling new scientific and historical scholarship through collaborative digitization in a way that could not be replicated across institutions by any other funding program. Though other efforts have been made to identify sources of funding for the digitization of archival field book items, including multiple efforts by each partner individually and by subsets of the partner group together, no program shares the BHL Field Note Project partner's interest in digitization of entire collections, providing full access to digitized collections online, encouraging and engaging in new forms of scholarship with the digital collections (such as the partner's interest in continuing to use these collections to inform new access and database tools, such as transcription and consequent full-text analysis and research), and supporting the full, online, perpetual access to these items through robust digital preservation. Furthermore, the project could only be possible with multiple, strategic partners across the United States, and no other funding program provides such strong support for collaborative digitization efforts. The partners involved in the BHL Field Notes Project strongly believe that the CLIR Digitizing Hidden Collections program most

closely aligns with their own goals and missions and that this program is therefore best suited to fund such an ambitious digitization initiative.

## Applicant Information

### Applicant Institution Address

In the event this proposal is approved for funding, payment will be made to the applicant institution as named at the beginning of this application form, and will be mailed to the address listed here.

Address

A large black rectangular box redacting the address information.

### Proof of Nonprofit Status

Nonprofit Status (Max. 2MB, .pdf format only)      NonProfitProof\_ALL\_small.pdf

### Board/Trustee List (not required for colleges or universities; required for all other applicants)

Upload a current list of board or trustee members. The list must be on the applicant institution's letterhead.

Board/Trustee List (Max. 2MB, .pdf format only)      DirectorsTrusteesRegents\_ALL.pdf



**Contact Information**

Provide the contact information for the proposed project's primary Principal Investigator(s) (PIs). The PI(s) will take direct responsibility for completion of the project, should funds be awarded. He or she must be significantly involved with the project's direction and execution and will be responsible for submitting required narrative and financial reports to CLIR.

*Note: Applicants may propose up to three PIs for their project. All applicants must assign at least one PI. Information for co-PIs should be included below as applicable.*

PI #1: Mr./Ms./Dr.	Mr.
PI #1: First Name	Martin
PI #1: Last Name	Kalfatovic
PI #1: Title	Associate Director, Smithsonian Libraries and Program Director, The Biodiversity Heritage Library
PI #1: Institution	Smithsonian Libraries
PI #1: Email	
PI #1: Address	
PI #1: City	
PI #1: State/Province	
PI #1: Zip/Postal Code	
PI #1: Country	

PI #2: Mr./Ms./Dr.	Ms.
PI #2: First Name	Anne
PI #2: Last Name	Van Camp
PI #2: Title	Director, Smithsonian Institution Archives
PI #2: Institution	Smithsonian Institution Archives
PI #2: Email	
PI #2: Address	
PI #2: City	
PI #2: State/Province	
PI #2: Zip/Postal Code	
PI #2: Country	

PI #3: Mr./Ms./Dr.	
PI #3: First Name	
PI #3: Last Name	
PI #3: Title	

PI #3: Institution

PI #3: Email

PI #3: Address

PI #3: City

PI #3: State/Province

PI #3: Country

### **Application Contact**

If CLIR's point of contact during the application period should be someone other than the proposed Principal Investigator(s) (e.g. a grants administrator), enter the name and contact information for the relevant individual here.

Application Contact: Mr./Ms./Dr.

Application Contact: First Name

Application Contact: Last Name

Application Contact: Title

Application Contact: Institution

Application Contact: Email

Application Contact: Address

Application Contact: City

Application Contact:

State/Province

Application Contact: Zip/Postal  
Code

Application Contact: Country

1890

March 24

Suwannee River, Florida.

Clear and rather cool with stiff N. wind.

After breakfast we all started out in our canoes, Chapman leading. He went down river about a mile and had the rare good luck to kill an Ivory-billed Woodpecker with his first shot hearing it pounding and calling in the cypress swamp only a few rods from the river and stalking it easily enough.

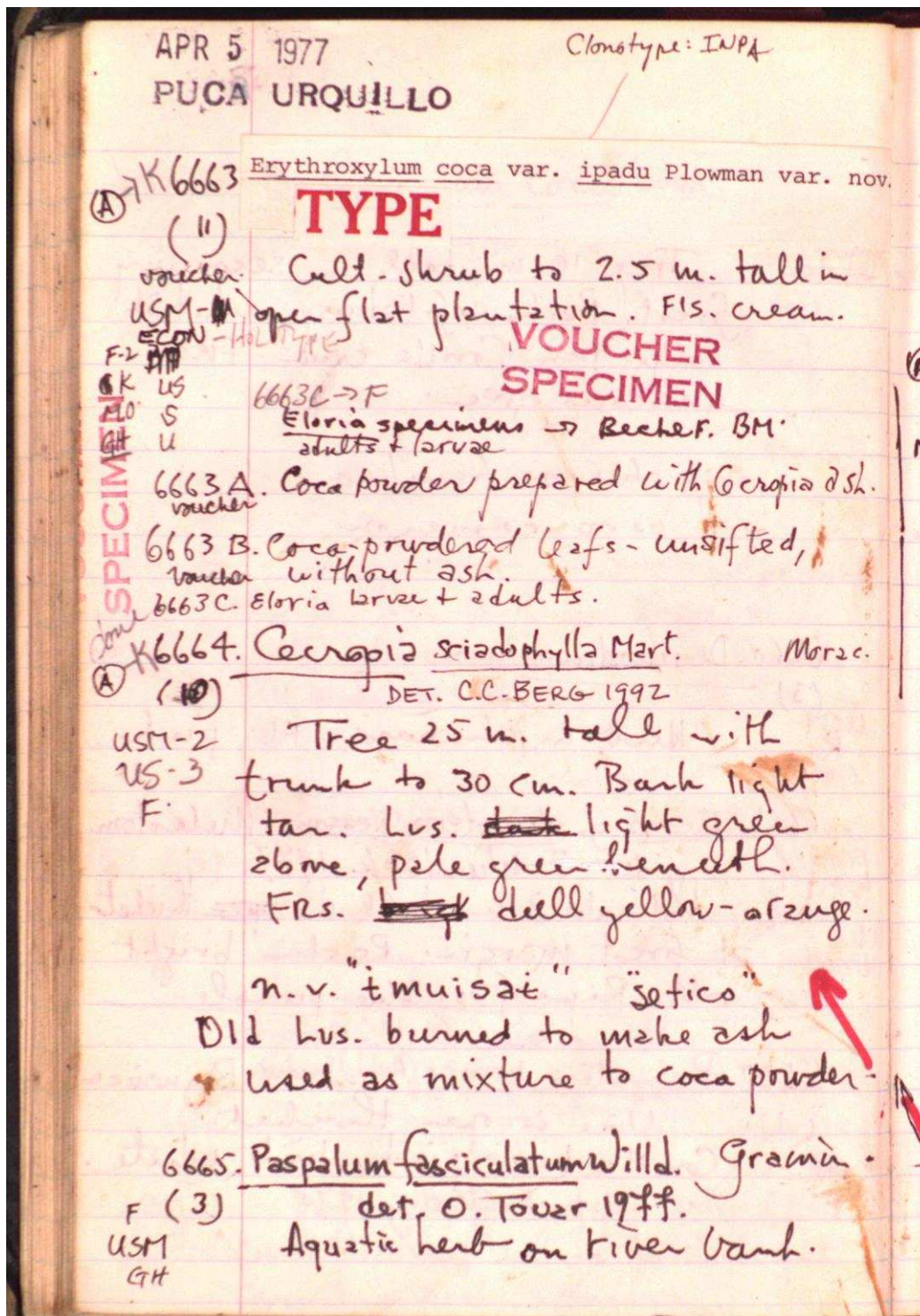
I landed about half a mile down and found a number of Warblers among which I shot a ♀ Bachman's. *Prothonotaria* were rather common along the river but I heard none singing.

I then went further down landing by chance at the very place where Chapman killed his Woodpecker. There were many birds here, also, but I found nothing worth shooting. In the swamp I heard a loud, harsh cry repeated every few seconds. I suspected that it might be an Ivory-bill and stole cautiously in keeping a sharp looking out in the trees but at length discovered the author of the sound at my feet. It was a small frog which a small garter snake was trying to swallow feet first. Every little while the snake would open its mouth wider and try to suck the frog in further when he would cry out as already described. I finally put my foot on the snake when he let go his hold and the frog went off with long and doubtless joyful jumps.

Chapman joined me in time for lunch which we ate on a point at the mouth

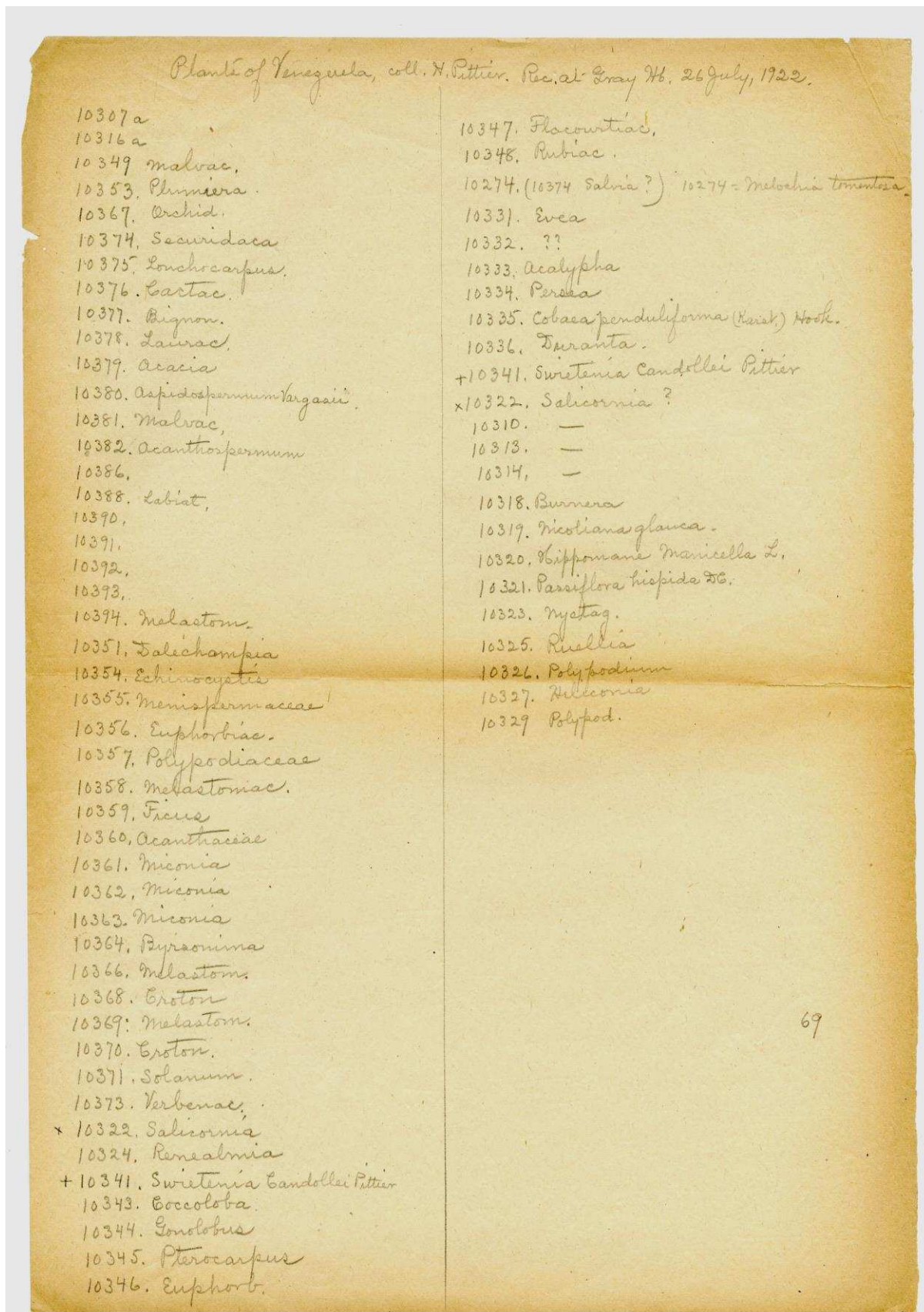




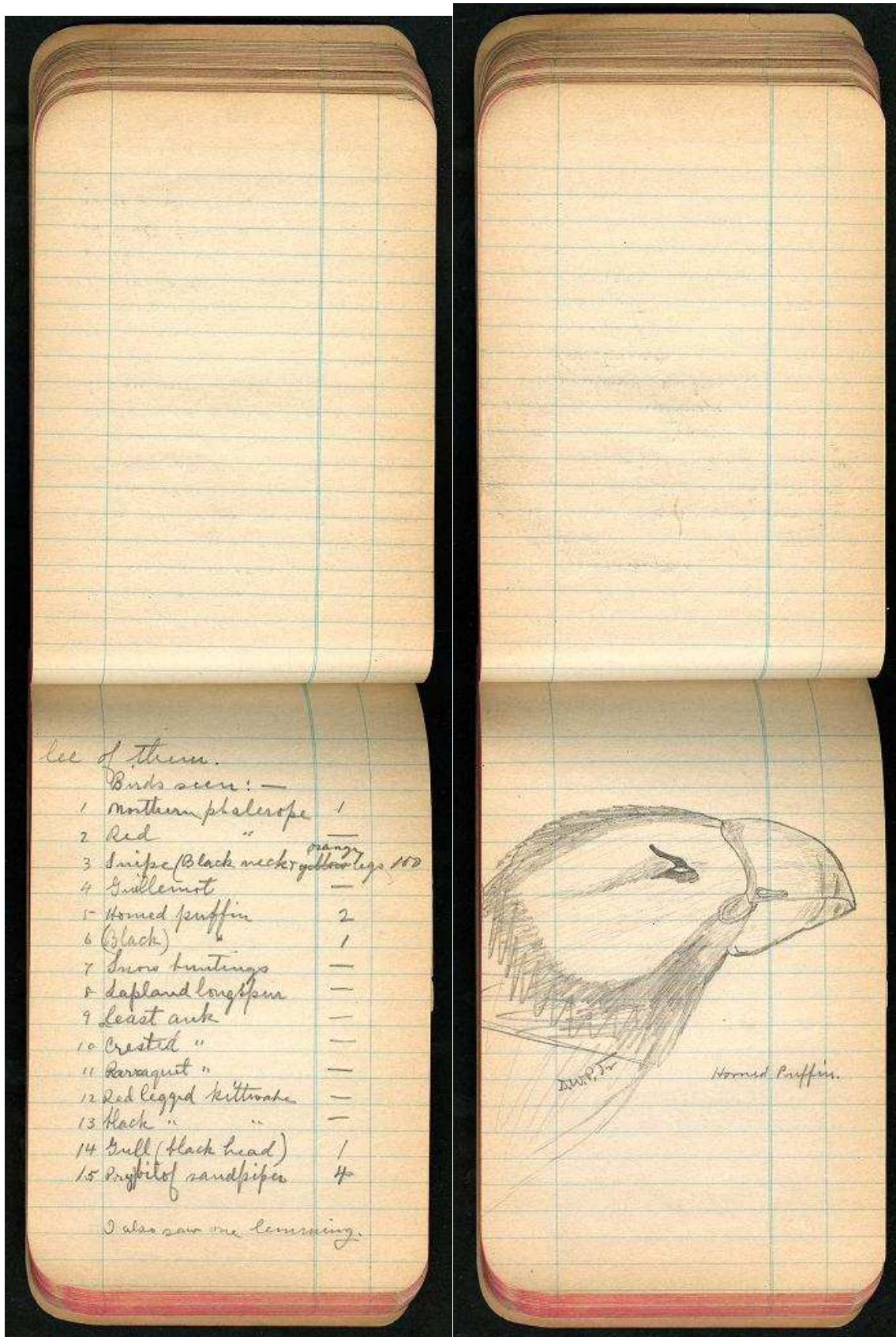


Field Museum of Natural History, Libraries and Archives. (1977). Collection Title: Timothy Plowman Field Notes. Rights and Licensing: Creative Commons Attribution-NonCommercial-ShareAlike 4.0



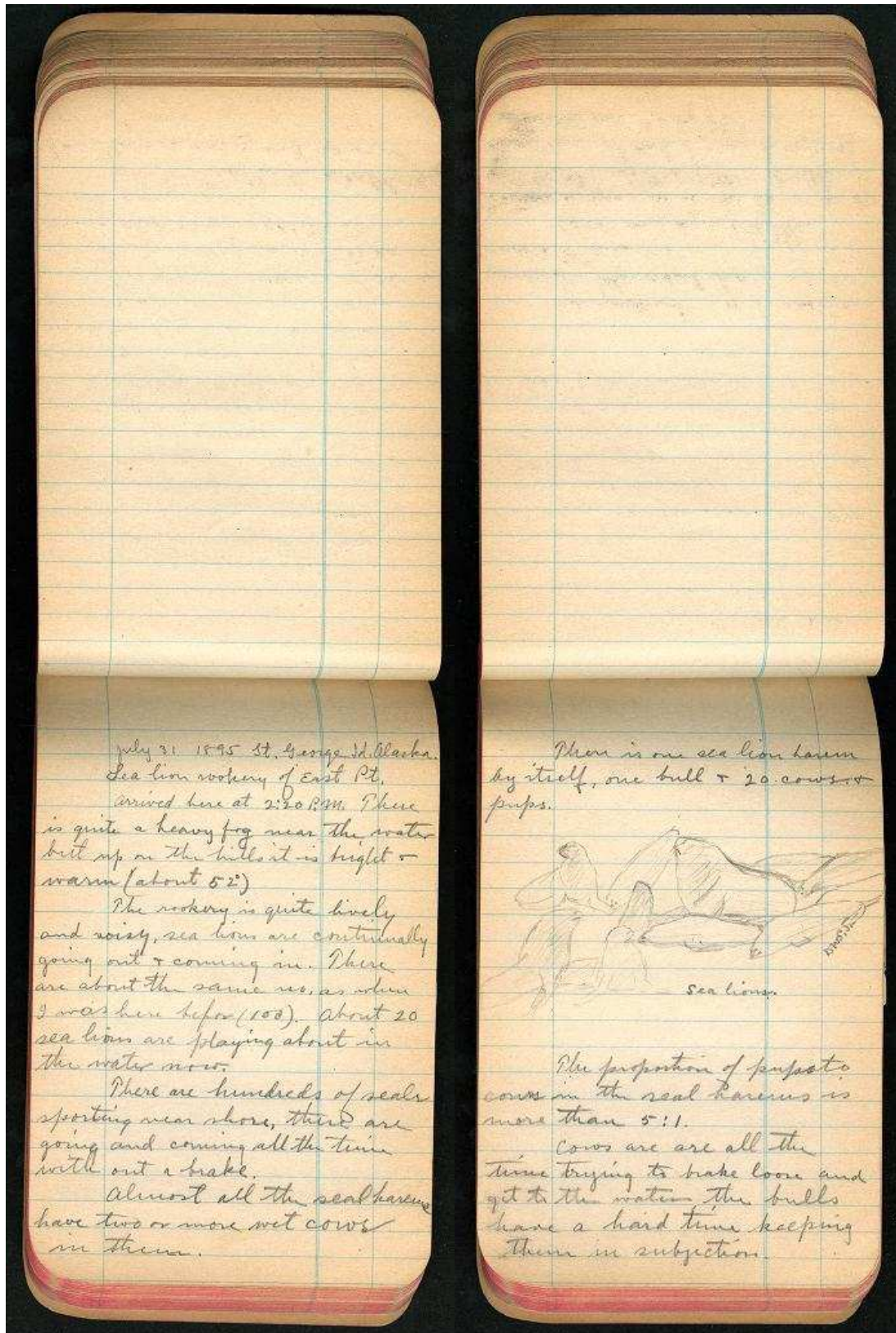




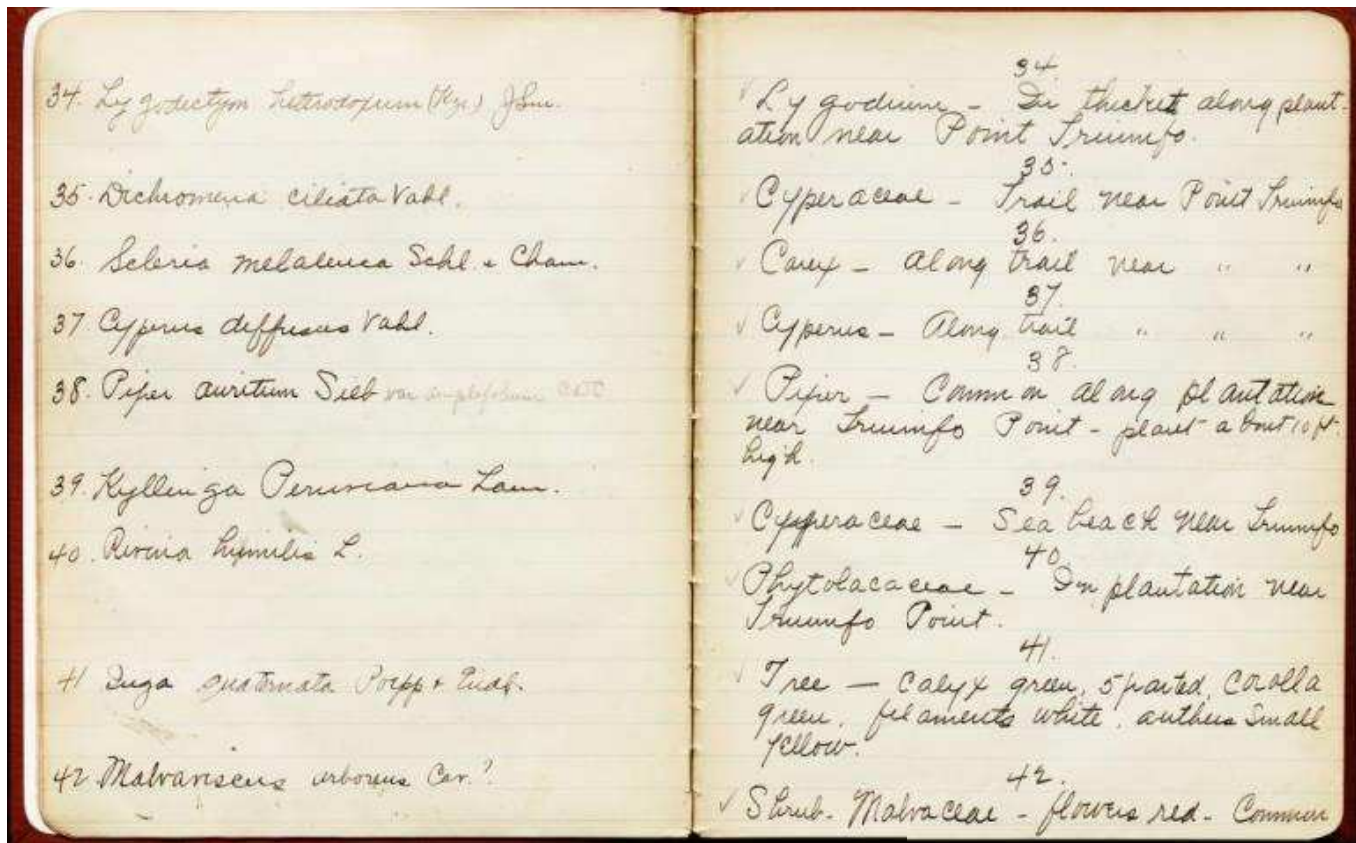


Smithsonian Institution Archives. (1895). Title: Two notebooks of Frederick William True containing field notes and daily entries on seals and other wildlife on St. Paul Island, Kelaire, Lukannan, Polovina, Tolstoi, and St. George's Island, Alaska. Rights and Licensing: No restrictions. The Smithsonian makes its content available for personal and non-commercial educational uses consistent with the principles of fair use.





Smithsonian Institution Archives. (1895). Title: Two notebooks of Frederick William True containing field notes and daily entries on seals and other wildlife on St. Paul Island, Kelaire, Lukannan, Polovina, Tolstoi, and St. George's Island, Alaska. Rights and Licensing: No restrictions. The Smithsonian makes its content available for personal and non-commercial educational uses consistent with the principles of fair use.



LuEsther T. Mertz Library, New York Botanical Garden. (1903). Title: Wilson Field Notebook: Puerto Sierra and Honduras, Volume 28, 1903. Rights and Licensing: The Mertz Library does not hold any copyrights on any files. If the original works from which the digital versions were produced are in the public domain, so too are the digital files. All information currently in the public domain remains in the public domain. The Library will not seek to assert any intellectual property rights over public domain materials.



## BHL Copyright Policy\*

BHL will only accept materials into its collection that fall under open access principles and thus encourages the digitization of:

1. Public domain materials
2. Materials free of any known copyright restriction
3. In-copyright content for which express permission to digitize has been received from the copyright holder

It is the responsibility of the scanning institution to ensure that all materials digitized are legally permitted to be available as open access content. Where in-copyright content is scanned for the BHL collection, it is recommended that the scanning institution apply a Creative Commons Attribution-Non Commercial-Share Alike 4.0 license (CC-BY-NC-SA 4.0 license) to the content. The BHL on behalf of consortium members and affiliates may pursue license agreements with copyright holders to digitize in-copyright materials for the collection. You may use BHL's standard licensing agreement form to secure agreements with copyright holders for digitization.

Generally speaking, participating BHL institutions that are in the United States or fall under U.S. jurisdiction may digitize and make openly accessible publications published prior to 1923 as these works are considered to be in the public domain. On occasion, items published in or around 1923 have been erroneously digitized by BHL libraries. Some early digitization procedures interpreted the public domain cut off as up to and including 1923. BHL will not proactively remove these items but will honor the terms of our "Take Down Guidelines" should we receive notice of a potential copyright violation.

Works that may be in the public domain under U.S. law may not be in the public domain under the laws of other countries. If your institution does not fall under U.S. law, you will have to research your own country's laws to determine appropriate copyright compliance. The BHL cannot and will not advise you or any other party about interpretation of your country's copyright laws.

**\*Note: For the purposes of the BHL Field Notes Project, all contributing partners have agreed to assume responsibility for ensuring that their contributed materials are in the public domain. BHL only accepts materials for this project that are in the public domain with no embargoes, access, or legal restrictions on use and re-use. All partners understand that they also assume responsibility for ensuring that there are no ethical considerations that would prohibit the online publication, use, and re-use of the materials digitized during this project. All materials digitized by this project and accompanying metadata are therefore assumed to be public domain with no restrictions on use, and all metadata will be published online under BHL's standard CC0 (public domain) certification.**

## Copyright Metadata

BHL uses 4 fields to document metadata regarding copyright issues for EVERY SINGLE item digitized. See BHL's template for Internet Archive's Partner Meta App for detailed instructions on how to fill out these fields. The fields relate to:

1. Copyright Status
2. License
3. Rights
4. Due Diligence

BHL requires the use of the following language standards. If your institution has already established language standards for copyright metadata then you may use your institution's standard in lieu of the BHL standard.

### Permissions Items/Volumes

IN-COPYRIGHT ITEMS / VOLUMES			
Required Values	Field Names	BHL Display	IA Partner Meta App
In copyright. Digitized with the permission of the rights holder.		Copyright status	possible-copyright-status
<a href="http://creativecommons.org/licenses/by-nc-sa/4.0/">http://creativecommons.org/licenses/by-nc-sa/4.0/</a>		License type	licenseurl
<a href="http://biodiversitylibrary.org/permissions">http://biodiversitylibrary.org/permissions</a>		Rights	rights
[Leave Blank]		Due diligence	duediligence

### IMPORTANT NOTES

- Add appropriate (c) metadata for your own home institutional publications where applicable
- For a current list of permissions titles see Permissions
- BHL provides access to in-copyright content under the Creative Commons Attribution Non-Commercial Share-Alike license <http://creativecommons.org/licenses/by-nc-sa/4.0/>

## Due Diligence Items/Volumes

DUE DILIGENCE ITEMS / VOLUMES			
Required Values	Field Names	BHL Display	IA Partner Meta App
No known copyright restrictions as determined by scanning institution.		Copyright status	possible-copyright-status
[Leave Blank]		License type	licenseurl
[Leave Blank]		Rights	rights
<a href="http://biodiversitylibrary.org/permissions">http://biodiversitylibrary.org/permissions</a>		Due diligence	duediligence

## Public Domain Items/Volumes

PUBLIC DOMAIN ITEMS / VOLUMES			
Required Values	Field Names	BHL Display	IA Partner Meta App
Public domain. The BHL considers that this work is no longer under copyright protection.		Copyright status	possible-copyright-status
[Leave Blank]		License type	licenseurl
[Leave Blank]		Rights	rights
[Leave Blank]		Due diligence	duediligence

**IMPORTANT NOTE:** Content under United States' public domain stays as U.S. public domain wherever possible. ONLY IF public domain is NOT an option, the BHL will go with the Creative Commons license that provides "maximum dissemination", the CC-BY (Attribution) license  
<http://creativecommons.org/licenses/by/4.0/>

# Project Design

## Project plan:

### i. Project Activities:

- Acquisition and/or set-up of staff and equipment for digitization. Q1. All partners.
- Metadata preparation. Includes creation of item-level records for those partner organizations that do not yet have item-level metadata created and converted to MARC XML. Q1-Q2. Involves Harvard MCZ, NYBG, UC Berkeley MVZ, and MOBOT.
- Digital imaging (scanning). Includes creation of high quality preservation TIFF images, processing of images into web-deliverable JPEGs, assigning technical and structural metadata such as local image numbers and page order, quality review of images and metadata, and storage of master TIFFs in local digital asset management system for preservation purposes. Begin Q1 or as soon as each partner has completed metadata conversion. Will continue throughout Q3, Q4, Q5, Q6, Q7, and halfway through Q8, reserving the last month of the program for project closeout analysis and documentation. The Smithsonian, MOBOT, Yale Peabody Museum, MVZ at Berkeley, AMNH, and both Harvard Institutions will be scanning throughout the entire process, Q1-Q8. The Field Museum will be scanning in Q1-Q4.
- Digital publication. Includes loading of materials into shared porting software Macaw, exporting items from Macaw to Internet Archive, and further processing of metadata by IA and export/import to Biodiversity Heritage Library. Begins and runs simultaneously with digitization, Q1-Q8. Partners who complete scanning in Q1-Q4 or who have smaller collections (such as Yale and NYBG) to scan will assist other partners with digital publication processes (metadata assignment, review) and, if time and resources allow, will assist in the ingest of as many digitized field books as possible from partner institutions into the BHL through the Macaw and Internet Archive process.
- Project closeout. Includes retrospective analysis of baseline goals for digital pace and quality, assessment of project plan and both successes and areas for improvement, preparation of workflows and tools documentation for open publication to the library and archives community via the BHL public wiki, and closeout report with lessons learned. Q8. Involves all partners.

### ii. Project Staff Responsibilities:

- A Project Manager will serve as a central coordination point for project planning and direction and will work out of the Smithsonian Libraries. He or she will resolve minor issues, assist in training on digital publication technologies, coordinate metadata workflows and digitization standards, and maintain clear and regular project communications between partners and up to PIs and CLIR.
- Each partner will be responsible for designating a Project Coordinator to manage on-site digitization work, ensure compliance with digitization standards, and report progress to the Project Manager. The Project Coordinator will also be responsible for ensuring that all items digitized are in the public domain. The Project Coordinators may also assign additional, in-kind contributed staff resources to the metadata assignment and digital ingest processes.

- Each institution that will be conducting local scanning will be responsible for the hiring and supervision of a full- or part-time Scanning Technician, depending on the amount of items they have identified for digitization. The Scanning Technician will be responsible for ensuring that their equipment is up to standard, their images are being scanned at the proper resolution, and that all metadata is mapped and imported into BHL for publication.

### iii. Project Deliverables:

- Approximately 50 entire collections of field note materials digitized, for a total of approximately 499,150 master TIFFs and JPEG2000 images created along with associated item-level metadata. All partners involved – each partner serves a role in digitization by engaging in digitization and/or metadata creation and quality review. Due 12/2017.
- Approximately 499,150 digital page images (TIFFs and JPEG2000) processed into web-deliverable formats (JPEGs) and published online for researcher access through the Internet Archive and Biodiversity Heritage Library along with all associated collection level, item level, and technical metadata for discovery, access, and display in a full and open manner. All partners involved, with additional support of BHL staff and IA staff for processing. Half due 12/2016, the remainder due 12/2017.
- All master digital images and metadata stored in a redundant, geographically distributed, fixity monitored digital asset management database for preservation by the Internet Archive. All partners involved, with significant contributed support of the Internet Archive digital asset management teams. Due date ongoing as files are created.
- Additionally, Smithsonian Libraries will maintain an archive of all files created during the project on Isilon high-density online storage with built in redundancy and high availability. Due date ongoing as files are created.
- Project workflows and documentation available for review by funding agency after project closeout; relevant workflows, metadata maps, and other documentation made available online via the BHL wiki for use and re-use by other institutions undertaking similar projects. Due 12/2017.

### iv. Equipment, technologies, standards:

Each partner is responsible for either providing local equipment or sending items to other partners for scanning. Currently, the Field Museum and Museum of Vertebrate Zoology at UC Berkeley would be sending their items to other partners for digitization, and all other partners will be scanning in-house. All partners will capture item level metadata using local standards and convert all metadata into MARC XML format, using Macaw software, for import to the Internet Archive and Biodiversity Heritage Library. After importing items to Macaw, partners may use Macaw to create page level technical and structural metadata.

PARTNER	RESPONSIBILITY	EQUIPMENT	TECHNOLOGIES	STANDARDS
Smithsonian	Scanning, local; digital publication support	Konika-Minolta book scanner, camera copystand for oversized items.	Adobe Photoshop, Macaw	MODS, MARC, XML

PARTNER	RESPONSIBILITY	EQUIPMENT	TECHNOLOGIES	STANDARDS
Internet Archive	Image and metadata processing; scanning, partners; online display; preservation.	Scribe and Tabletop Scribe; large image overhead capture for oversized items.	IA Biblio tool, Scribe Software and Republisher	MARC, XML
American Museum of Natural History (AMNH)	Scanning, local; digital publication support.	Tabletop Scribe	Freeflow software, Macaw	MARC
Yale Peabody Museum	Digital publication, scanning, local	Kirtas robotic book scanner; cameras and copystand/vacuum table for oversize items	Adobe Photoshop, Macaw, KE EMu, Kirtas proprietary software.	EAD, XML
Botany Libraries, Harvard University Herbaria	Scanning, local; digital publication support	IA tabletop scribe at MCZ (two Nikon 1 J3s); Internet Archive: Canon model 5D, Mark II, and Canon EF 100mm f and a variety of state of the art equipment at Harvard Library Imaging Services as needed.	Macaw; Internet Archive Harvest Tool (Harvard-developed, for ingestion into Harvard DRS); Adobe Creative Suite	MODS, METS, MARC, XML
Ernst Mayr Library of the MCZ at Harvard	Scanning, local; digital publication support	IA tabletop scribe. WideTek 25 flatbed scanner for oversized items up to 17.7 x 25" as needed.	Adobe Photoshop, Macaw, LuraTech JP2 Command Line Tool	MARC
LuEsther T. Mertz Library, New York Botanical Garden	Provision of field notes, digital publication support	n/a, outsourcing scanning to IA	See Internet Archive	EAD, MARC, MARCXML
Museum of Vertebrate Zoology at UC Berkeley	Provision of field notes, digital publication support	n/a, outsourcing scanning to IA	See Internet Archive	MARC, MARCXML
Field Museum, Chicago	Provision of field notes, digital publication support	n/a, outsourcing scanning to IA	See Internet Archive	MARC, MARCXML
Missouri Botanical Garden	Scanning, local; Digital publication support.	Copy stand with Leaf Aptus camera	Capture 1 (Leaf Aptus specific image capture and editing software), Adobe Photoshop. MARCEdit, MarcXMLImport, Botanicus.	MARC, MARCXML

Collection Title	Collection Size	Holding Institution	Institution/Staff Responsible for Digitization	Institution/Staff Responsible for Description	Institution/Staff Responsible for Quality Control	Institution/Staff Responsible for Long-Term Maintenance	When in the course of the project will this collection be processed?
George Engelmann Botanical Notebooks	58 items, 28000 pages	Missouri Botanical Garden Archives	Missouri Botanical Garden Archives	Missouri Botanical Garden Archives	Missouri Botanical Garden Archives	Internet Archive	Q1-Q4
Timothy Plowman (Curator, Economic Botany) field notes	35 items, 3800 pages	Field Museum	Field Museum	Field Museum	Field Museum	Internet Archive	Q2-Q5
MVZ Archival Field Notebooks	1500 items, 165000 pages	Museum of Vertebrate Zoology at UC Berkeley	Internet Archive	Museum of Vertebrate Zoology at UC Berkeley	Museum of Vertebrate Zoology at UC Berkeley	Internet Archive	Q1-Q8
Collectors' Field Notes, Archives, The New York Botanical Garden	230 items, 12000 pages	New York Botanical Garden	Internet Archive	New York Botanical Garden	New York Botanical Garden	Internet Archive	Q1-Q4
Archbold Expeditions Field Books	14 items, 1912 pages	American Museum of Natural History, Mammalogy	American Museum of Natural History	American Museum of Natural History	American Museum of Natural History	Internet Archive	Q2-Q5
Whitney South Seas Expedition Field Books	54 items 10,900 pages total	American Museum of Natural History, Ornithology	American Museum of Natural History	American Museum of Natural History	American Museum of Natural History	Internet Archive	Q2-Q5
Mss .H45 1918, The diary of Edmund Heller, October 9, 1917-January 12, 1918 : covering his return trip from the First Asiatic Expedition led by Roy Chapman Andrews of the American Museum of Natural History.	1 item, 132 pages	American Museum of Natural History, Research Library	American Museum of Natural History	American Museum of Natural History	American Museum of Natural History	Internet Archive	Q2-Q5
Yale Peabody Museum, Archival Field Notes	400 items, 44,000 pages	Yale Peabody Museum Archives	Yale Peabody Museum Archives	Yale Peabody Museum Archives	Yale Peabody Museum Archives	Internet Archive	Q1-Q8
RU007179; Edmund Heller Papers, circa 1898-1918	138 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
SIA Acc. 12-300; Edmund Heller Field Notes, 1909-1910, Birds	3 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
SIA Acc. 12-212; Mammals, Edmund Heller Field Books, 1898-1915	13 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
SIA RU007186; United States Exploring Expedition Collection, 1838-1885	35 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
12-443; Mammal Section, Bird and Mammal Labs, Fish and Wildlife Service, US Department of the Interior	161 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
RU007264; Collected notebooks, manuscripts, drawings, photographs, and correspondence on vertebrate and invertebrate paleontology, United States National Museum, circa 1850-1940 and undated.	17 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
12-314; Palearctic Migratory Bird Survey (PMS) field catalogs, 1966-1973	24 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
RU007242; August F. Foerste Papers, 1887-1933 and undated	33 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
07-041; Geological Survey (U.S.), field notebooks, 1881-1915 and undated	30 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
12-346; Edward Palmer Field Notes, Botany	26 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q4-Q8
RU000245; National Museum of Natural History, Pacific Ocean Biological Survey Program, records, circa 1961-1973, with data from 1923	1115 items	Smithsonian	Smithsonian	Smithsonian	Smithsonian	Internet Archive	Q1-Q5
Field notes and plant identification records, 1850-1990	38500 pages	Harvard Botany Libraries, Harvard University Herbaria	Harvard Botany Libraries, Harvard University Herbaria	Harvard Botany Libraries, Harvard University Herbaria	Harvard Botany Libraries, Harvard University Herbaria	Internet Archive	Q1-Q8
Diaries (enclosures for vols. 13 & 15), bBr 97.41.1	5 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Journals, sBr 97.41.2	1094 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Journals (transcripts of selections), bBr 97.41.2a	1220 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8

Collection Title	Collection Size	Holding Institution	Institution/Staff Responsible for Digitization	Institution/Staff Responsible for Description	Institution/Staff Responsible for Quality Control	Institution/Staff Responsible for Long-Term Maintenance	When in the course of the project will this collection be processed?
Birds of Lake Umbagog, Maine, sBr 97.41.5	2147 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Notes on birds, arranged systematically, sBr 97.41.9	7597 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Notes on birds, eggs, & nests, sBr 97.42.1	126 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Field notes on New England birds, sBr 97.42.2	366 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Ornithological notes on New England birds, sBr 97.42.3	362 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
[Measurements of birds] , sBr 97.42.4	176 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
New England birds. General notes, sBr 97.42.5	34 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Concord field list, sBr 97.42.7	194 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Concord field lists and notes, sBr 97.42.8	444 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
New England birds. Systematic notes, sBr 97.42.19	892 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Arizona Field Notes, sBr 641.42.1	187 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Correspondence,	25388 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Photographic prints (EML),	830 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Original manuscripts of papers, sBr 97.41.3	95 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Lake Umbagog Itinerary, sBr 97.41.8	54 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Record of pigeon shoots, sBr 97.42.6	18 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Record of guests, sBr 97.50.5	35 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Callers at the log cabin, Ball's Hill, Concord, Massachusetts, June 1904-October 1910, sBr 97.50.6	52 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Bibliography of the writings of William Brewster, sBr 97.50.7	88 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Photo Album: Gulf of the St. Lawrence; Mt. Carmel, Ind. ca. 1882; Florida & Georgia; Concord, Mass.; Lake Umbagog, Maine, sBr 97.70.1	106 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Photo Album: Concord & Cambridge, Mass.; Florida & Georgia; Maine; Europe, sBr 97.70.2	52 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Photographs: Interior views of Brewster's home in Cambridge, Mass., sBr 97.70.17	24 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8
Prints of Glass Plate Negatives, sBr 97.70.17	25 pages	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Ernst Mayr Library, MCZ, Harvard	Internet Archive	Q1-Q8



# BHL Field Notes Project

## Project Design

### Technical Plan

#### i. Project Workflow:

1. Metadata prep: Local item prep completed by each partner. Includes assignment of item-level metadata, where absolutely necessary for administrative purposes, in accepted standards such as MODS and MARC, as well as mapping of preexisting item-level descriptive metadata into MARC XML for web publication through BHL. No metadata will be restricted in any way – all will be published online in the Internet Archive and BHL in a full and open manner.
2. Digitization: Local scanning completed on partner equipment, such as Konika-Minolta book scanner for Smithsonian, Scribe or Tabletop Scribe for Internet Archive, Botany Libraries at Harvard University Herbaria, and Ernst Mayr Library of the MCZ at Harvard, and copy stand with Leaf Aptus digital camera at the Missouri Botanical Garden. Items from the Field Museum, the NYBG, and the MVZ at Berkeley will be sent to the Internet Archive for digitization. “Digitization” includes imaging (as TIFF or JPEG2000), cropping, color adjustment, and QA of images using Adobe Photoshop. Images will be scanned at 300 dpi minimum and 24 bit RGB color depth.
3. Local image prep: Images given local ID numbers. Master TIFFS embedded with technical, structural, and administrative metadata and sent to local digital asset management systems for preservation per partner institution workflows.
4. Macaw import: Images loaded into Macaw along with associated metadata in MARC XML through Macaw CSV import process. As part of import process, Macaw converts images into web-ready uncompressed JPEGs.
  - MACAW is an open source tool with all code available on GitHub (<https://github.com/cajunjoel/macaw-book-metadata-tool>). MACAW performs three tasks in the process of digitizing a book-like item: import of the images from the scanner, camera, or existing files, collection of the page-level metadata that describes the physical aspects of the page, and post-processing and exporting/uploading the digital book to other systems.
5. Structural metadata assignment: Images assigned technical and structural metadata by local staff using Macaw interface. Metadata includes page numbers, dates, and type of page image (text, illustration, fold out, cover, blank).
6. Internet Archive import: Images exported to Internet Archive by Macaw. During export, Macaw embeds images with the assigned structural metadata.
7. Internet Archive processing: Internet Archive processes images into additional web-deliverable formats (JPEG, full color PDF, black and white PDF, EPUB, Kindle, plain text).
8. Digital preservation: Internet Archive exports master copies of images and metadata to long-term digital preservation solution. Internet Archive is responsible for long term access to images and metadata even after project conclusion.
9. BHL import: Internet Archive exports all derivative (web-deliverable) images with associated item and page level metadata to BHL.

10. BHL processing: BHL processes images and metadata for web platform display. Once per week, BHL updates the public site with all newly imported items. All newly imported items are then assigned to one central collection to assist researchers in finding the materials. All items are then also accessible through the standard, BHL-wide search and browse functionality via the item-level indexed metadata fields.

#### ii. Relation to current local practice:

All project workflows and metadata standards align with current local best practices for description and preparation of archival materials for item-level digitization. All digital files will be linked to their catalog records at their local institutions, either as links to EAD finding aids or other collection records (for instance, the Smithsonian connects all digital files to an NCD collection record instead of an EAD finding aid).

#### iii. Quality control of deliverables:

Materials are reviewed for quality at several stages. All scanning technicians review each image for scanning quality and to verify image metadata before loading it into Macaw. All project staff working in Macaw (one per institution, contributed time) review the images and item-level metadata again as they assign page level metadata, and review the page-level metadata before marking each item "Review Complete" and initiating the export to the Internet Archive. Project staff (contributed time) at the Smithsonian monitor the daily addition of new materials to the Internet Archive to ensure that no image or metadata issues occur during import. Finally, project staff (Project Manager) completes a weekly review of new project items loaded into BHL to record their BHL unique identifiers and, during the process, complete a final quality review of the published materials' image and metadata quality.

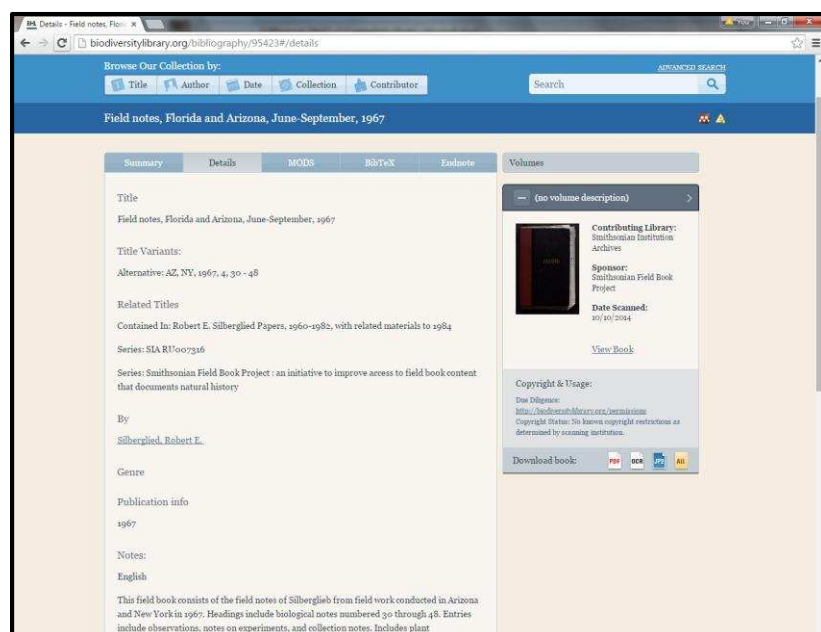
#### iv. Project staffing detail:

Staff assigned specifically to the project include one Project Manager and four scanning technicians for a total of 5. Staff expected to assist in planning, metadata, scanning, processing, digital preservation, and digital publication, include one project coordinator from each partner and several advisers for a total of 17 (contributed time). Staff members are listed below. Technical staff names remain unassigned pending hiring for the BHL Field Notes Project, should funds be awarded.

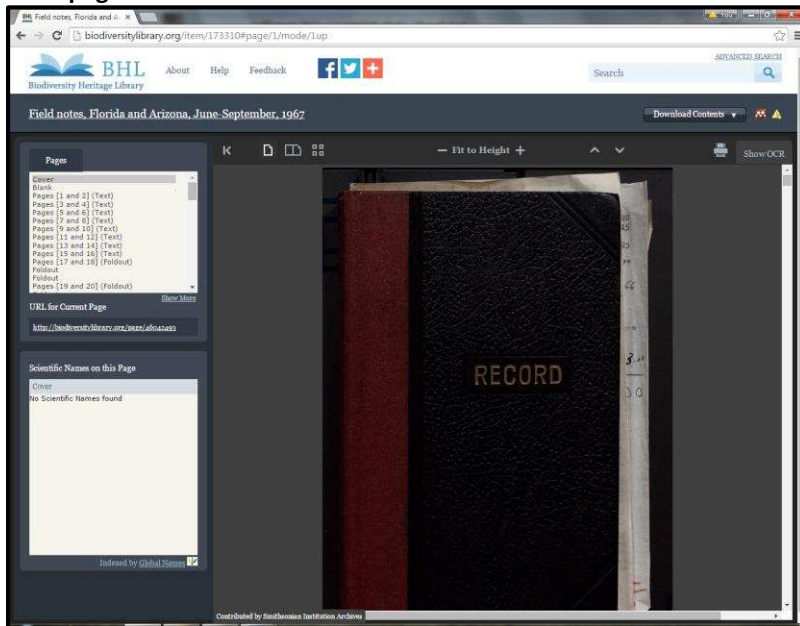
Name	Project Title (Staff Title)	Duty station	Funding
Martin Kalfatovic	Principal Investigator (Associate Director, SIL)	Smithsonian Libraries	Contributed time
Anne Van Camp	Project Coordinator (Director, SIA)	Smithsonian Archives	Contributed time
<b>Julia Blase</b>	<b>Project Manager (Field Book Project Manager)</b>	<b>Smithsonian Libraries</b>	<b>CLIR grant award</b>
Robert Miller	Project Coordinator (General Manager, Digital Libraries)	Internet Archive	Contributed time
Judith Warnement	Project Coordinator (Library Director)	Harvard University Herbaria	Contributed time
Constance Rinaldo	Project Coordinators (Library Director)	Ernst Mayr Library of the MCZ at Harvard	Contributed time
Joe DeVeer	Project Coordinator (Head of Technical Services)	Ernst Mayr Library of the MCZ at Harvard	Contributed time

Name	Project Title (Staff Title)	Duty station	Funding
Susan Fraser	Project Coordinator (Library Director)	LuEsther T. Mertz Library, NYBG	Contributed time
Thomas Baione	Project Coordinator (Library Director)	American Museum of Natural History	Contributed time
Tim White	Project Coordinator (Director of Collections)	Yale Peabody Museum	Contributed time
Larry Gall	Project Coordinator (Head of Computer Systems)	Yale Peabody Museum	Contributed time
<b>Samantha Wentworth</b>	<b>Scanning technician</b>	<b>Yale Peabody Museum</b>	<b>CLIR award (.5 FTE)</b>
Christine Giannoni	Project Coordinator (Library Director)	Field Museum	Contributed time
Christina Fidler	Project Coordinator (Archivist)	Museum of Vertebrate Zoology, UC Berkeley	Contributed time; <b>.25 FTE CLIR award</b>
Doug Holland	Project Coordinator (Library Director)	Missouri Botanical Garden	Contributed time
<b>TBD upon hiring: 4 scanning technicians</b>	<b>Scanning technician</b>	<b>Smithsonian, Harvard (shared by MCZ and Botany Libraries), MOBOT, AMNH</b>	<b>CLIR grant award</b>
Bianca Crowley	BHL Advisor (BHL Collections Coordinator)	Smithsonian Libraries	Contributed time
Joel Richard	Macaw Advisor (Lead Developer, SIL)	Smithsonian Libraries	Contributed time
Ricc Ferrante, Kira Sobers	Digitization Advisors (Digital Services Director, Digitization Specialist SIA)	Smithsonian Archives	Contributed time

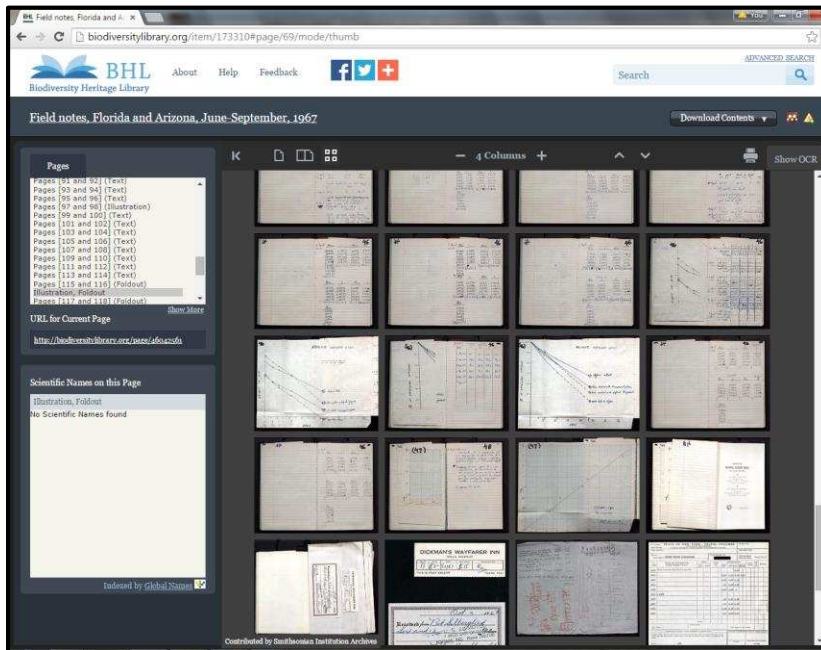
**v. Mock-up image, title page, in BHL:**



## Book page:



## Grid view:



## Metadata

Title: Field notes, Florida and Arizona, June-September, 1967  
Collection name: Robert E. Silberglied Papers, 1960-1982, with related materials to 1984  
Physical Description: 1 folder  
Physical Location: Smithsonian Institution Archives  
Record type: Fieldbook record  
Object Type: Field notes, Graphs, Charts (graphic documents)  
Place: United States, New York, Arizona  
Date: 1967  
Topic: Entomology  
Accession #: SIA RU007316  
Access Information: Many of SIA's holdings are located off-site, and advance notice is recommended to consult a collection. Please email the SIA Reference Team at [osiaref@si.edu](mailto:osiaref@si.edu).  
Abstract: This field book consists of the field notes of Silberglied from field work conducted in Arizona and New York in 1967. Headings include biological notes numbered 30 through 48. Entries include observations, notes on experiments, and collection notes. Includes plant identifications for plants used in experiments. Labels for collections made in Arizona in 1967 are also included. Includes data charts, tables, and sketches.  
See more records from this collection: [Robert E. Silberglied Papers, 1960-1982, with related materials to 1984](#)  
See more records associated with this person: [Silberglied, Robert E.](#)  
Data Source: Smithsonian Field Book Project  
Record ID: fbr\_item\_MODSI5926

# BHL Field Notes Project

## Sustainability

### Digital Preservation and Discoverability Plan

All images created by the project partners will be sent to the Internet Archive as part of the BHL import process. The Biodiversity Heritage Library, in partnership with the Internet Archive, will be responsible for long term accessibility to the digital files and metadata, as it is also the digital publication venue and a widely-used service with a user base comprised of the scientists, students, and researchers that would be the primary users of the collections. The BHL platform also maintains, updates, aggregates and publishes project metadata for external harvesting via open API keys. The BHL technical team, housed at the Missouri Botanical Garden, currently provides technical support services for BHL and will continue to do so, even after project conclusion. As part of their partnership with the BHL Field Notes Project, and in a reflection of the technical support services they already provide to the BHL which includes providing web-ready images and metadata from their secure servers, the Internet Archive will take responsibility for long term digital preservation of all images sent to it, during the project and after project conclusion. Storing the Archive's collections involves parsing, indexing, and physically encoding the data. The Internet Archive's hardware consists of PCs with clusters of IDE hard drives. Data is stored on DLT tape, hard drives, and the Petabox system in various appropriate formats, depending on the collection.

Preservation is the ongoing task of permanently protecting stored resources from damage or destruction.

Discoverability over time results from robust digital preservation, such that the digital images and all associated metadata can be maintained, web access copies can be re-generated if necessary, and content can be migrated across platforms as technologies change to ensure they are always available to the public. The main issues are guarding against the consequences of accidents and data degradation and maintaining the accessibility of data as formats become obsolete. Multiple copies of the digital images ingested during the BHL Field Notes project will be maintained at multiple sites to help alleviate the risk of any accident happening to a particular storage site, tape, or disk. Over time, storage media can degrade to a point where the data becomes permanently irretrievable. The industry rule of thumb is to migrate data every 10 years. The Internet Archive uses a Petabox™ system for long term digital storage and preservation. The PetaBox™, custom-designed by Internet Archive staff, was originally created to safely store and process one petabyte (a million gigabytes) of information. The goals and design points were:

- \* Low power: 6kW per rack, 60kW for the entire storage cluster
- \* High density: 100+ TB/rack
- \* Local computing to process the data (800 low-end PC's)
- \* Multi-OS possible, linux standard
- \* Co-location friendly
- \* Shipping container friendly: Able to be run in a 20' by 8' by 8' shipping container.
- \* Easy Maintenance: One system administrator per petabyte

- \* Software to automate full mirroring
- \* Easy to scale
- \* Inexpensive design
- \* Inexpensive storage

The Internet Archive data center now houses ~3PB of PetaBox storage technology and is expanding steadily. All images, with their associated metadata, created during the BHL Field Note Project will be stored on the Petabox system, fully mirrored, monitored for fixity, and with plans to migrate or emulate outdated files and technology every 10 years, or as otherwise dictated by industry-wide best practices.

Furthermore, certain partners will maintain their own backup copies of all master TIFF images created by their local scanning operations over the course of the project. The Smithsonian images are sent to the Smithsonian Digital Asset Management System for secure, permanent storage as part of the standard digitization workflow. Furthermore, all images imported to BHL through this project will be stored again on the Smithsonian Isilon server. All AMNH materials will have a further copy stored on a local DSpace repository. All digitized images created by the Harvard Botany Libraries will be stored in the Harvard Digital Repository Service (DRS) (<http://hul.harvard.edu/ois/systems/drs/>). Similarly, Yale Peabody assets are sent to a campus-wide Digital Asset Management System as part of workflow, and public access derivatives are automatically generated and managed through an Amazon S3 service/API. All MOBOT items will have a further copy stored in the local digital repository. Finally, the BHL web platform continues to go through development phases to add functionality that will increase the discoverability and sustainability of the materials published therein, including a recent assessment of accessibility to populations such as the hearing and vision disabled. These and other discoverability improvements will continue throughout the BHL lifecycle to ensure that all of the materials it makes available will remain discoverable to the widest audiences and for the longest time possible.

# BHL Field Notes Project

## Budget Narrative

### Summary:

The total budget for the BHL Field Notes project is \$491,713.

The primary costs of the project are salaries and wages for scanning technicians at each institution and the processing service costs of digitization, minus overhead, through and in partnership with the Internet Archive, who will not only contribute their processing expertise but also the platform and tools or materials for publishing items online through the BHL research portal. The partnership and contributed expertise of the Internet Archive enables economies of scale on their end and significantly lowers institution-by-institution digitization costs, saving both time and funding.

Notice that some partners request scanning technicians full-time and some part-time, some for one year and some for both years – each partner has evaluated the number of and type of materials, as well as other existing resources such as equipment, space, and support, and determined individually what arrangement of scanning technicians is best for their institution.

Three partners are outsourcing work entirely to Internet Archive (Berkeley, NYBG, and the Field Museum), and therefore are not requesting any funding for scanning technicians or image processing service costs. Three partners will be using Internet Archive Tabletop Scribe machines locally and request funding for technicians to operate the machines (Harvard MCZ, Harvard Botany Libraries, and AMNH). Three partners will continue to use their own machines and request scanning technician salaries (Smithsonian, MOBOT, and Yale Peabody). The Internet Archive, as a full partner, requests funding to support per-page image processing costs for full, open, high resolution digital publication of images and import of images to the BHL; Internet Archive will be requesting these costs for all pages scanned either by them or using one of their Tabletop Scribes, in line with existing institutional and BHL procedures. Internet Archive will be processing images from Berkeley, NYBG, Field Museum, both Harvard partners, and AMNH.

All scanning technicians will also provide digital image processing support for their own institutions. All scanning technicians, as well as project coordinators, can and will provide support (including taking on scanning or metadata assignment and review) to other partners should project work face any unexpected slowdown due to employee issues (extended sick leave, for instance) or technical issues (local equipment failure for any reason).

### Line Items:

- **Line items, Year 1: Salaries and Wages**
  - Project Manager: The salary for a full-time project manager is included to facilitate coordination of work amongst the nine project partners. The project manager will be responsible for communications, coordination of work between partners, maintaining accurate and regular project



progress reports and workflow and progress documentation, and sharing those reports with the funding agency and other executive level project stakeholders.

- Scanning Tech, Smithsonian: The salary for a part-time scanning technician for the first year for the Smithsonian is requested to support the scanning of approximately 531 items in the first year of the project (the beginnings of the Pacific Ocean Biological Survey collection). A full-time scanning technician is not requested because the available equipment time for 2016 only supports a part-time position. Beginning in 2017, equipment is available full-time and a full-time scanning technician will be requested.
  - Project Coordinator, MVZ Berkeley: Berkeley has the most materials to scan of any partner except the Smithsonian and therefore requests a quarter-time paid project coordinator to assist in the movement of materials between Berkeley and the Internet Archive scanning facility, and also in the reviews of digitized items for quality control.
  - Project Coordinator/scanning tech, Yale: Yale will be using its own, local Kirtas robotic scanner and requests a half-time project coordinator and scanning tech to manage local workflows.
  - Scanning tech, Harvard MCZ: Harvard will be using an Internet Archive Tabletop Scribe machine to scan materials and therefore requests a part-time scanning technician to operate the Scribe. This technician will be shared, part time for MCZ's materials and part-time for Harvard Botany Libraries Materials. The division of time will be carefully documented to ensure that equal time is spend on the collections from both Harvard partners.
  - Scanning tech, Harvard Botany Libraries: As above – both Harvard partners will be sharing an Internet Archive Tabletop Scribe machine for scanning, as well as a scanning technician, who will devote part of their time to Harvard MCZ and the other part time to Harvard Botany Libraries.
  - Scanning tech, AMNH: AMNH will be purchasing a Tabletop Scribe machine but requests funding for a part-time scanning technician to operate the machine and direct scanning operations.
  - Scanning tech, MOBOT: MOBOT will be using their own Kirtas Robotic Scanner but, due the fragile nature of some of their materials, requests the salary of a full-time scanning technician.
  - Fringe benefits: Fringe benefit rates vary by institution. The attached budget has added prefixes to each “% of” column (column B) to indicate which institution's rate is being cited.
- **Line items, Year 1: Supplies and materials**
    - Tabletop Scribe: Most institutions either already have, will be cost-sharing the purchase of, or will not be needing additional scanning equipment. However, the AMNH does not have existing equipment with the capacity to scan at the speed (quantity) targeted by the project. Therefore, the project requests \$5,000 in the first year to support the purchase of a Tabletop Scribe at AMNH. This machine costs \$10,00 and the other \$5,000 will be cost-share by the AMNH. After the completion of the project, the machine will be dedicated to scanning additional materials from AMNH and other partners for the BHL, with the goal of specifically targeting materials that are significantly related to the digitized field note collections.
  - **Line items, Year 1, Services**
    - Isilon backup storage: Additional storage is requested for Smithsonian to support the permanent preservation storage of all items digitized by the project and added to IA and BHL through Macaw.
    - Internet Archive processing fees: These line items reflect the Internet Archive's standard \$.10 per page charge for image processing. Each item reflects the number of pages that will be processed, per institution, by Internet Archive, in year 1.
    - Internet Archive item set-up fees: This line item reflects Internet Archive's standard \$3 per item charge to set up the Scribe machine with the new item and do color and alignment calibration pre-



scanning to ensure the highest possible quality images are captured. In year 1, the estimated items to be scanning on IA machines is 1,274.

- **Line items, Year 2: Salaries and Wages**

- Project manager: As in year 1, the salary for a full-time project manager is included to facilitate coordination of work amongst the nine project partners.
- Scanning tech, Smithsonian: As indicated, in year 2 the Smithsonian will have the equipment available to scan full-time and requests and full-time scanning tech salary to support the digitization of an additional 1,062 items.
- Project Coordinator, MVZ Berkeley: As in year 1, Berkeley has the most materials to scan of any partner except the Smithsonian and therefore requests a quarter-time paid project coordinator to assist in the movement of materials between Berkeley and the Internet Archive scanning facility, and also in the reviews of digitized items for quality control.
- Project Coordinator/scanning tech, Yale: As in year 1, Yale will be using its own, local Kirtas robotic scanner and requests a half-time project coordinator and scanning tech to manage local workflows.
- Scanning tech, Harvard MCZ: As above, Harvard MCZ requires an additional year of a part-time scanning tech to support digitization. As above, this tech will continue to be shared between MCZ and the Harvard Botany Libraries.
- Scanning tech, Harvard Botany Libraries: As above, Harvard Botany Libraries requires an additional year of a part-time scanning tech to support digitization. As above, this tech will continue to be shared between MCZ and the Harvard Botany Libraries.
- Scanning tech, MOBOT: As above, MOBOT requests a full two years of full time scanning technician salary to support digitization.
- Scanning tech, AMNH: As above, AMNH requests a full two years of a part time scanning technician salary to support digitization.
- Fringe benefits: As above, the attached budget has added prefixes to each “% of” column (column B) to indicate which institution’s rate is being cited.

- **Line items, Year 2: Services**

- Internet Archive processing fees: These line items reflect the Internet Archive’s standard \$.10 per page charge for image processing. Each item reflects the number of pages that will be processed, per institution, by Internet Archive, in Year 2.
- Internet Archive item set-up fees: This line item reflects Internet Archive’s standard \$3 per item charge to set up the Scribe machine with the new item and do color and alignment calibration pre-scanning to ensure the highest possible quality images are captured. In year 2, the estimated items to be scanning on IA machines is 1,009.

- **Digitization costs**

- Digitization costs are primarily the salaries of digitization technicians. Digitization costs also include post-scanning image processing, such as the assignment of item- and page-level technical and structural metadata, the embedding of that metadata in the images, the transformation of any collection-level metadata into publication-appropriate MARC XML, and the processing of images into web-deliverable formats. These costs are taken on by the Internet Archive, which possesses extensive experience, expertise, and economies of scale, and can offer these services at a rate of \$.10 per page. This cost will be applied to all images processed by the Internet Archive, either through outsourced digitization directly to them (for Berkeley, NYBG, and Field Museum) or through the use of their machines by local technicians (Harvard MCZ, Harvard Botany Libraries,

AMNH). These costs are covered in the salary estimations of the partners who will be using their own scanning machines and processing software (Smithsonian, Yale Peabody, and MOBOT).

- **Vendors**
  - No vendors will be necessary during this project.
- **Grant management**
  - The grant will be managed as an in-kind or cost-share contribution by the Smithsonian Libraries grants management office.
- **Cost share**
  - Cost share includes the contributed staff time from each institution for a Project Coordinator (nine in total), each of whom is expected to contribute .02 to .25 FTE on project work, depending on the size of their collections. Contributed time also includes that of the Smithsonian advisory staff, each of whom may contribute up to .15 FTE on project work. Cost share includes half of the purchase cost of the Tabletop Scribe, contributed by AMNH. Cost share includes the costs of staff time and equipment to absorb digital preservation, provided by the Internet Archive. The Internet Archive also contributes staff time required to scan items sent to it by partners and manage the technology used in processing images received by it in preparation for their publication on BHL, which also includes server uptime and any software or technology troubleshooting.

Digitizing Hidden Special Collections and Archives  
Application Budget Form

Name of Applicant (Institution):

Smithsonian Institution

Collection Title:

BHL Field Notes Project

Project Period:

1/2016-12/2017

**BUDGET SUMMARY**

*NOTE: The Budget Summary section is populated from information provided in the Budget Detail sections that follow. Please do not attempt to enter figures into the Budget Summary.*

	Requested Funds
Salaries and Wages	\$332,070
Fringe Benefits	\$117,093
Consultant and Training Fees	\$0
Supplies and Materials	\$5,000
Services	\$37,551
Other Costs	\$0
<b>TOTAL</b>	<b>\$491,713</b>



Digitizing Hidden Special Collections and Archives  
Application Budget Form

**Name of Applicant (Institution):**

Smithsonian Institution

**Collection Title:**

BHL Field Notes Project

**Project Period:**

1/2016-12/2017

will be dedicated to the project may be included here.

<b>Item</b>	<b>Basis/Method of Cost Computation</b>	<b>Requested Funds</b>
Tabletop Scribe	Cost listed by partner, Internet Archive	\$5,000
	<b>SUBTOTAL</b>	<b>\$5,000</b>

Digitizing Hidden Special Collections and Archives  
Application Budget Form

**Name of Applicant (Institution):** Smithsonian Institution  
**Collection Title:** BHL Field Notes Project  
**Project Period:** 1/2016-12/2017

**5. Services**

Include the cost of other services (e.g. equipment rental, server time, backup charges) related to project objectives that are not included under other budget categories. Subcontracts should be included in this category.

Item	Basis/Method of Cost Computation	Requested Funds
Isilon backup storage	Approximately \$833 per terabyte * 4 tb	\$3,332
Internet Archive processing fees, Berkeley	.10 per page * 82500 pages from Berkeley MVZ	\$8,250
Internet Archive processing fees, NYBG	.10 per page * 12000 pages from NYBG	\$1,200
Internet Archive processing fees, Field Museum	.10 per page * 3850 pages from Field Museum	\$385
Internet Archive processing fees, Harvard MCZ	.10 per page * 20702 pages from MCZ	\$2,070
Internet Archive processing fees, Harvard Botany Libraries	.10 per page * 19250 pages from Harvard Botany Libraries	\$1,925
Internet Archive processing fees, AMNH	.10 per page * 12944 pages from AMNH	\$647
Internet Archive item setup fees, all	\$3 per item * 1274 items total	\$3,822
	<b>SUBTOTAL</b>	<b>\$21,631</b>

**6. Other costs**

Include any items not previously listed. Please note that "miscellaneous" and "contingency" are not acceptable budget categories.

**Funds to cover indirect costs may not be requested.**

Item	Basis/Method of Cost Computation	Requested Funds
	<b>SUBTOTAL</b>	<b>\$0</b>

**7. Total costs**

(Subtotals of items 1 through 6)

**Requested Funds**

\$232,950

<b>Name of Applicant (Institution):</b>	Smithsonian Institution
<b>Collection Title:</b>	BHL Field Notes Project
<b>Project Period:</b>	1/2016-12/2017

## BUDGET DETAIL: YEAR TWO

**Budget detail for the period**

FROM (mm/yyyy):	01/2017
TO (mm/yyyy):	12/2017

## 1. Salaries and Wages

Provide the names and titles of the principal project personnel. For support staff, include the title of each position and indicate the number of persons who will be employed in that capacity. Unpaid volunteers should not be included.

Name/Title of Position	No.	Method of Cost Computation	Requested Funds
Project Manager			
Scanning Tech, Smithsonian			
Project Coordinator, UC Berkeley			
Project Coordinator/scanning tech, Yale			
Scanning Tech, Harvard MCZ			
Scanning Tech, Harvard Botany Libraries			
Scanning Tech, MOBOT			
Scanning Tech, AMNH			
		<b>SUBTOTAL</b>	<b>\$180,129</b>

## 2. Fringe Benefits

If more than one rate is used, list each rate and salary base.

[illegible]

### 3. Consultant and Training Fees

Include payments for professional or technical consultants.

Name or type of consultant or training	No.	Method of Cost Computation	Requested Funds
		SUBTOTAL	\$0

#### 4. Supplies and materials

Include consumable supplies, materials to be used in the project and items of expendable equipment (i.e., equipment items costing less than \$5,000 and with an estimated useful life of less than a year). The proposed purchase of software which is both essential and

Digitizing Hidden Special Collections and Archives  
Application Budget Form

**Name of Applicant (Institution):**

Smithsonian Institution

**Collection Title:**

BHL Field Notes Project

**Project Period:**

1/2016-12/2017

will be dedicated to the project may be included here.

Item	Basis/Method of Cost Computation	Requested Funds
	<b>SUBTOTAL</b>	<b>\$0</b>



Digitizing Hidden Special Collections and Archives  
Application Budget Form

<b>Name of Applicant (Institution):</b>	Smithsonian Institution
<b>Collection Title:</b>	BHL Field Notes Project
<b>Project Period:</b>	1/2016-12/2017

### 5. Services

Include the cost of other services (e.g. equipment rental, server time, backup charges) related to project objectives that are not included under other budget categories. Subcontracts should be included in this category.

Item	Basis/Method of Cost Computation	Requested Funds
Internet Archive processing fees, Berkeley	.10 per page * 82500 pages from Berkeley	\$8,250
Internet Archive processing fees, AMNH	.10 per page * 6472 pages from AMNH	\$647
Internet Archive processing fees, Harvard MCZ	.10 per page * 20702 pages from MCZ	\$2,070
Internet Archive processing fees, Harvard Botany Libraries	.10 per page * 19250 pages from Botany Libraries	\$1,925
Internet Archive item fees, all	\$3 per item * 1009 items total (year 2)	\$3,027
	<b>SUBTOTAL</b>	<b>\$15,919</b>

### 6. Other costs

Include any items not previously listed. Please note that "miscellaneous" and "contingency" are not acceptable budget categories.

**Funds to cover indirect costs may not be requested.**

Item	Basis/Method of Cost Computation	Requested Funds
	<b>SUBTOTAL</b>	<b>\$0</b>

### 7. Total costs

(Subtotals of items 1 through 6)

**Requested Funds**  
\$258,763

Smithsonian Institution

BHL Field Notes Project

1/2016-12/2017

### BUDGET DETAIL: YEAR THREE

FROM (mm/yyyy):

N/A

---

## 1. Salaries and Wages

Provide the names and titles of the principal project personnel. For support staff, include the title of each position and indicate the number of persons who will be employed in that capacity. Unpaid volunteers should not be included.

Name/Title of Position	No.	Method of Cost Computation	Requested Funds
		<b>SUBTOTAL</b>	<b>\$0</b>

## 2. Fringe Benefits

If more than one rate is used, list each rate and salary base.

Rate		Salary Base	Requested Funds
	% of		
	% of		
	% of		
	% of		
		<b>SUBTOTAL</b>	\$0

### 3. Consultant and Training Fees

Include payments for professional or technical consultants.

Name or type of consultant or training	No.	Method of Cost Computation	Requested Funds
		<b>SUBTOTAL</b>	<b>\$0</b>

#### 4. Supplies and materials

Include consumable supplies, materials to be used in the project and items of expendable equipment (i.e., equipment items costing less than \$5,000 and with an estimated useful life of less than a year). The proposed purchase of software which is both essential and will be dedicated to the project may be included here.

Item	Basis/Method of Cost Computation	Requested Funds
------	----------------------------------	-----------------

## Digitizing Hidden Special Collections and Archives Application Budget Form

Name of Applicant (Institution):

Smithsonian Institution

**Collection Title:**

BHL Field Notes Project

**Project Period:**

1/2016-12/2017

[illegible]

**Name of Applicant (Institution):**  
**Collection Title:**  
**Project Period:**

Smithsonian Institution  
BHL Field Notes Project  
1/2016-12/2017

## 5. Services

Include the cost of other services (e.g. equipment rental, server time, backup charges) related to project objectives that are not included under other budget categories. Subcontracts should be included in this category.

Item	Basis/Method of Cost Computation	Requested Funds
	<b>SUBTOTAL</b>	<b>\$0</b>

## 6. Other costs

Include any items not previously listed. Please note that "miscellaneous" and "contingency" are not acceptable budget categories.

**Funds to cover indirect costs may not be requested.**

Item	Basis/Method of Cost Computation	Requested Funds
	<b>SUBTOTAL</b>	<b>\$0</b>

## 7. Total costs

(Subtotals of items 1 through 6)

## Requested Funds

\$0