

Hubbard Brook Information Management Documents

Table of contents

Preface	5
1 Data Management Plan	7
1.1 History of IM at the Hubbard Brook Experimental Forest	7
1.2 Governance	8
1.3 Research Approval Committee	8
1.4 Personnel	8
1.5 Data Packages	9
1.5.1 Overview	9
1.5.2 Data Holdings	9
1.5.3 Metadata Standards	9
1.5.4 Data Package Development	10
1.5.5 Data Quality Control	10
1.5.6 Access Policy	11
1.5.7 Data Access	11
1.6 IM Software/Hardware	12
1.7 IT Support	12
1.8 Account access	12
1.9 Website	13
1.10 Sample Archive	13
1.10.1 Samples	13
1.10.2 Sample Archive Subsampling Policy	14
1.10.3 Directions for Sample Submission	14
1.11 Bibliography	15
1.12 Computing Facilities	15
2 Data Package Workflow	16
2.1 Overview	16
2.2 Data package development workspace	16
2.3 Step-by-Step Data Package Workflow	17
2.4 Non-tabular datasets (images, audio, very large datasets, etc)	18
2.4.1 Large Datasets	18
2.5 Notes on revising older datasets.	18
2.6 Forest Service Data Workflow	19

3	Data Catalog Workflow	20
3.1	Overview	20
3.2	Database and File access	20
3.3	Step-by-Step Catalog Workflow	20
3.4	APPENDIX A – Code to build the local HBR data catalog:	20
3.5	FetchSarepointDataInventory.R	21
3.6	dataCat.R	22
4	Data Inventory	27
5	Publication Management	28
5.1	Overview	28
5.2	Zotero - step-by-step	28
5.3	EDI publication-data linkages	29
5.4	Zotpress	29
6	Physical Sample Archive	30
6.1	Overview	30
6.2	Sample archive database	31
6.3	Templates/Forms	31
6.4	Rshiny	31
6.5	Management Plan	31
7	Voucher Specimens	32
7.1	Background	32
7.2	What samples have already been submitted to collections?	32
7.3	Recommendations for future sample submissions	33
8	Personnel Directory	34
9	Hubbard Brook Listservs	35
9.0.1	Establishment, configuration, and status	35
9.0.2	The Lists	35
9.0.3	Use patterns	36
9.0.4	Listserv platforms	36
9.0.5	The Future of Lists	36
10	Website	37
10.1	Content editing access	37
10.2	Adding Feature Stories	37
10.3	Adding News posts	39
10.4	Creating new pages	40
10.5	Adding new People pages	40
10.6	Adding Lessons	41

10.7 Adding content to the media library	41
10.8 General tips on Beaver Builder (BB) layouts	42
10.9 Setting publication status and visibility	42
10.10Theme colors	42
10.11Headings	42
10.12Migrating content	43
11 Online Book	44
11.1 Managing chapter content	44
11.2 Automated graphs	44
12 RAC Proposal Submissions	45
12.1 Overview	45
12.2 IM role in RAC submissions	45
12.3 RAC data publishing instructions	46
12.4 Jotform Tips	46
13 Reporting scripts	47
13.1 Publication Status	47
13.2 Data Status	47
13.2.1 Status of data in EDI	47
13.2.2 Status of All data	47
13.2.3 Data access reports	47
13.3 GitHub Repository for scripts	47
14 Templates	48
14.1 Email	48
14.1.1 New community members.	48
14.1.2 Email to RAC approved investigators	48
14.1.3 Emails to data submitter	48
14.2 Data Submission Template	48
14.3 Sample Archive Submission Template	48
15 Computer Resources	49
15.1 Computer Resources/Facilities (UNH/ESRC)	49

Preface

This collection of documentation contains details on Information Management for the Hubbard Brook Experimental Forest. Data are managed by both the USDA Forest Service (USFS) and the Information Manager for the Hubbard Brook LTER site (one of 25+ sites funded by NSF's Long-term Ecological Research program).

In addition to an overview of information management at HBR, we include supplemental chapters and associated documents covering more dynamic details such as inventory/status of data packages, current IM projects and timelines, a guide to the operation of the HBR website, and a step-by-step guide to the development of HBR data packages and associated local data catalog. This document is revised periodically to reflect changes in IM assets, status, workflows, etc.

Throughout these pages you will encounter a number of acronyms. They are typically identified before first use within a chapter...but just in case:

- EDI: Environmental Data Initiative
- EML: Ecological Metadata Language
- ESRC: Earth Systems Research Center
- HBEF: Hubbard Brook Experimental Forest
- HBR: Acronym for the Hubbard Brook LTER
- LTER: Long-term Ecological Research
- LNO: LTER Network Office
- PASTA+: The software that runs the repository
- RAC: Research Advisory Committee
- RCC: Research Computing Center
- USFS: USDA Forest Service

- UNH: University of New Hampshire
- WMNF: White Mountain National Forest

1 Data Management Plan

This chapter is based on the most recent Data Management Plan (DMP) developed for the 2022 LTER site renewal proposal. Subsequent chapters expand on each element and reflect recent accomplishments and detailed instructions on the various workflows used in the HBR Information Management environment. Also available is the full **2022 HBR DMP** as originally submitted.

1.1 History of IM at the Hubbard Brook Experimental Forest

The Hubbard Brook Experimental Forest (HBEF; USDA Forest Service) was established in 1955 and became an NSF-funded Long Term Ecological Research Site (HBR) in 1988. Information management has been an important component at Hubbard Brook from its inception. Data and documents from 1955 onward have been stored and protected, and although most of these early items consist of physical assets (paper charts, photographic slides, field notes, handwritten data, publications, etc), much of this material has been converted to digital format, with original copies in fireproof storage at the HBEF Pierce Lab and at the Northern Research Station in Durham, NH.

The establishment of the LTER-HBR occurred at a time of rapidly changing technology; desktop computers and email were new, and the internet as we know it was still several years away. The Hubbard Brook community fully embraced these emerging technological resources, and established access to data with the ‘Source of the Brook’, a public access dial-up electronic bulletin board, which allowed easy retrieval of many data sets from the HBR (1990). From a dialup bulletin board and gopher server in the early 1990s, to the World Wide Web in the late 1990s, HBR’s latest technology advances in publicly sharing data and resources has seen a migration of the website to WordPress, a data catalog built on dynamic access to content in the Environmental Data Initiative repository (EDI), bibliography management in Zotero, to name a few...

Until 2012, Information management for HBR was provided through the Forest Service, with John Campbell filling this role from 1997-2012. During this time, the LTER network adopted EML (Ecological Metadata Language) as a metadata standard, and HBR was an early adopter of this standard. In 2003-4 the first EML-based data packages were prepared for HBR with online download access and formatted browser display of metadata.

Funding for the HBR Information Management position was provided in the 2010 renewal of LTER-HBR funding (HBR5), and the position was filled in mid-2012 by Mary Martin (Earth Systems Research Center, University of New Hampshire, Durham, NH).

1.2 Governance

Information management at the Hubbard Brook Ecosystem Study (HBES) is guided by the Information Oversight Committee (IOC), which meets on an ad hoc basis with virtual IOC meetings scheduled accordingly.

1.3 Research Approval Committee

A Research Approval Committee (RAC) has been established to assist the Forest Service and broader HBES community in making decisions regarding what research studies will be allowed. In making its recommendation, the RAC considers a number of factors related to: (1) the relationship of the proposed project to the overall Hubbard Brook Ecosystem Study (how does this project fit into the overall study; why is it important for this research to occur at the Hubbard Brook Experimental Forest, as opposed to some other site); (2) the scientific merit of the proposed research; (3) the integrity of the site (e.g. how will this research impact the Forest or other ongoing research projects); and (4) the extent to which the proposed research compromises or enhances ongoing efforts. The RAC's critical review of proposed research at Hubbard Brook helps ensure that the scientific value of the Hubbard Brook Experimental Forest is maintained for the future.

Proposal submissions to the RAC are made through JotForm webforms on <https://hubbardbrook.org/research/research-proposal-submission>, and are currently being used to develop a project-level database to support the RAC review process and IM tracking of data collections at Hubbard Brook. Researchers approved by the RAC are encouraged to submit data to be included in the Hubbard Brook Data Catalog, regardless of funding source.

1.4 Personnel

Hubbard Brook LTER Information management transitioned from the Forest Service to the HBR LTER grant in the HBR-V funding cycle (2010-2016). The Information Manager is based at the Earth System Research Center (ESRC), University of New Hampshire (UNH), and funded through a subcontract between the Cary Institute of Ecosystem Studies and UNH.

Information Management resources (software, hardware, and personnel) from the USDA Forest Service Northern Research station contribute substantially to the overall data holdings of the

HBR. These include the collection, quality control, and development of core hydrological, meteorological, and water chemistry datasets upon which much of the HBR research relies.

See also IT resources for as-needed project support (Section 4) and Appendix 1 for a description of as-needed support from the ESRC *Laboratory for Remote Sensing and Spatial Analysis*.

1.5 Data Packages

1.5.1 Overview

HBR data packages are prepared for submission to the Environmental Data Initiative (EDI), following best practices developed over 4 decades by the LTER Information Management community (<https://ediorg.github.io/data-package-best-practices/eml-best-practices.html>). These best practices, and the efforts of the EDI, ensure that data are Findable, Accessible, Interoperable, and Reusable, following the principles of the FAIR initiative. The EDI serves as the primary repository for HBR data, and details about the operation of EDI can be found at <https://edirepository.org>.

1.5.2 Data Holdings

The HBR data catalog has had a strong emphasis on long-term datasets and data from the major watershed experiments. Many of these data pre-date the establishment of LTER-HBR, and are now available as a result of a 60+ year culture of robust data curation and sharing. More than 20 HBR data packages have been collected over a period of 50 or more years, with another 30 covering a timespan of more than 20 years. Through close coordination with the Research Approval Committee, the Hubbard Brook Committee of Scientists, and project administration, HBR-IM is able to identify datasets that can be incorporated into our data catalog. Graduate students working at Hubbard Brook are also surveyed periodically to identify forthcoming datasets and they are also trained in the EDI data publication workflow.

1.5.3 Metadata Standards

All HBR data packages are prepared for submission through the development of metadata in the Ecological Metadata Language standard (EML; <https://eml.ecoinformatics.org/eml-ecological-metadata-language>). Basic EML content includes: title, abstract, personnel, contacts, publication date, spatial and temporal coverages, keywords (consistent with LTER controlled vocabulary), project funding, publisher, data access and use policies, and detailed attribute-level metadata. Data download and use is facilitated through the fully described attribute metadata (column names, definitions, units, missing values, and coding). The highest level of EML completion is achieved through the EDI congruency checker with informational,

warning, and error messages that provide feedback to HBR-IM on additional steps that can (or must) be taken to submit to the repository. These congruency checks read both metadata and data, testing a minimum of 40 conditions that can be addressed to insure that data packages are fully capable of integration with other data, and fully operational in higher level workflows and automated data processing.

In 2019, LTER sites began using EML2.2. EML2.2 provides the structure to accommodate a number of advanced metadata elements. Of note, is the ability to annotate data packages at the data package, entity, and attribute level, by linking to persistent identifiers in ontologies, such as those found at <https://bioportal.bioontology.org/ontologies> and elsewhere. As the annotation elements in EML2.2 become populated, both the discoverability of datasets, and the ability to use datasets in synthesis efforts will be enhanced. HBR-IM has been a member of both the EDI Semantics Working Group (formed in January 2019), and the EDI/LTER Units Working group (2023-present), which have a goal of developing best practices and training on the use of the new EML2.2 annotation elements. The Units Working Group is responsive to recommendations from the LTER 40-year review on facilitating synthesis efforts. This has been accomplished by establishing a relationship with the QUDT ontology (<https://qudt.org>), developing code to map ad hoc units to this ontology, and preparing a manuscript on this effort for publication.

1.5.4 Data Package Development

HBR IM has fully adopted the EDI ezEML application for data package development. This cloud-based application was developed by EDI and users are supported by the EDI team and ezEML developer. This application continues to be developed to support emerging LTER/EDI data requirements. Within this system, HBR can store templates for access and re-use of common metadata elements (people, taxonomy, geographic coverage, funding, etc). EzEML also supports a ‘collaboration’ mode, where IM and data creators can work together to complete the full metadata documentation necessary for the repository.

A separate chapter describes the data package development workflow in detail: [HBR Data Package Development](#).

1.5.5 Data Quality Control

The HBR research community is widely dispersed among different institutions and laboratories, and data quality control is implemented primarily by the individual researcher. All data packages include methods, wherein detailed data QC protocols can be documented. The HBR-IM works with research teams to document quality control in the data package metadata as appropriate. This may range from descriptions directly in EML, PDF files uploaded with data packages, or cross referencing to details on data QC available elsewhere. IM provides the data submitter with feedback on a number of QC checks that are implemented during the data

package development workflow. These include value ranges for data table attributes, coding consistency, and additional issues that are flagged within the ezEML environment and the EDI congruency checker (the final checks during repository upload).

1.5.6 Access Policy

The HBR data policy follows that of the LTER Data Access Policy, as updated in 2017. (<https://lternet.edu/data-access-policy/>; Creative Commons license - Attribution - CC BY; <https://creativecommons.org/licenses/by/4.0/>). All pre-existing HBR data packages have been revised to include this new policy, and the policy is linked on the hubbardbrook.org information management page. The policy reads as follows:

This information is released under the Creative Commons license - Attribution - CC BY (<https://creativecommons.org/licenses/by/4.0/>). The consumer of these data (“Data User” herein) is required to cite it appropriately in any publication that results from its use. The Data User should realize that these data may be actively used by others for ongoing research and that coordination may be necessary to prevent duplicate publication. The Data User is urged to contact the authors of these data if any questions about methodology or results occur. Where appropriate, the Data User is encouraged to consider collaboration or co-authorship with the authors. The Data User should realize that misinterpretation of data may occur if used out of context of the original study.

While substantial efforts are made to ensure the accuracy of data and associated documentation, complete accuracy of data sets cannot be guaranteed. All data are made available “as is.” The Data User should be aware, however, that data are updated periodically and it is the responsibility of the Data User to check for new versions of the data. The data authors and the repository where these data were obtained shall not be liable for damages resulting from any use or misinterpretation of the data.

1.5.7 Data Access

The complete inventory of Hubbard Brook data can be browsed, filtered, and searched on the HBR website <https://hubbardbrook.org/d/hubbard-brook-data-catalog>. Data are also discoverable through both the EDI data portal (<https://portal.edirepository.org>), through DataONE (<https://search.dataone.org>), google dataset search (<https://datasetsearch.research.google.com/>), and through DataCite (<https://datacite.org>), the entity providing dataset DOIs to EDI. A separate stand-alone document is generated as needed (for proposals and reviews), and describes all HBR data packages in the EDI (and other) repository – ***Hubbard Brook Data Catalog Inventory***.

1.6 IM Software/Hardware

Table 1 outlines software in use by HBR-IM to manage data package development and the Hubbard Brook website.

Table 1. Features of HBR Information Management System

Feature	Details, software, resources
Website: https://hubbardbrook.org	WordPress, html, css, php, xslt, javascript, apache, piwigo
Bibliography	Zotero, Zotpress wordpress plugin
Data Catalog	EDI data repository, local WordPress gateway to EDI HBR data, EDIutils R
Metadata	ezEML, PostgreSQL, EML R package, EML2.1
Computer Hardware	Dell Poweredge R510, desktop and laptop linux systems.
Backup	BackupPC, rsnapshot, daily, weekly and monthly backups, on and off-site
Data management	R, LibreOffice, QGIS, MySQL, PostgreSQL, git

1.7 IT Support

IT support for the UNH IM team is available through the UNH Research Computer Center (RCC). RCC provides support to researchers in the Institute for the Study of Earth, Oceans, and Space, as well as to the wider University research community, State of NH, and Federal Agencies. ESRC has had a long-standing Service Level Agreement (SLA) with RCC (more than 20 years) which can be provided to reviewers upon request. In addition to overall IT support described in the SLA, RCC also provides the personnel for as-needed project support. This gives the HBR-IM team access to expertise for special projects, without the need to provide ongoing support for personnel on the IM team for programming, web design, etc. HBR-IM has also made use of an *RCC-funded* internship program, wherein computer science undergraduates are paired with researchers in the Institute for Earth, Oceans, and Space.

1.8 Account access

A number of accounts require access for HBR Information Management. In some cases, access can be granted through the UNH Research Computing Center (RCC), and in other cloud-based accounts the IM has designated a backup person to have account access.

- IM desktop and server [IM/RCC]
- UNH HB sharepoint [IM/Contosta at unh]
- Zotero [IM/BU]
- Databases [IM/RCC]
- Piwigo [IM/RCC]
- EDI/ezEML [IM/EDI]
- Jotform [IM/Keeling at Cary]
- Bluehost [IM/Post at 6288media]
- GitHub
- hubbardbrook gmail [IM/BU]

1.9 Website

The HBR website (<https://hubbardbrook.org>) is the primary means by which HBR information is disseminated, with additional non-digital data (charts, maps, photographs) made available upon request. HBR completed a website migration from html/php files to a content management system (CMS; Drupal) in 2017. In 2022 the website was further migrated to WordPress. With cloud-based WordPress hosting, it becomes simpler to transfer website management to a different IM and/or institution. Website content is managed by the HBR IM and the Hubbard Brook Research Foundation. The website provides access to data, publications, personnel pages, education and outreach material, and a photo gallery. A recent website content addition is the Hubbard Brook Research Synthesis – an online ‘book’ consisting of 19 chapters to date, covering a wide range of long-term research. With content editors assigned to each chapter, this is meant to be a series of dynamic pages, which reflect the full history of Hubbard Brook research in each topic area (<https://hubbardbrook.org/online-book>). Chapters in the online book contain numerous data figures, many of which have been restructured to read data directly from the most recent data revisions in the EDI repository.

See *Hubbard Brook Ecosystem Study Website Management Guide* for website configuration, access to servers and filesystems, recommended practices for site content, etc.

1.10 Sample Archive

1.10.1 Samples

In 1990, an archive facility was built at the Hubbard Brook Experimental Forest to store samples permanently so that they will be available for future research. The 1860 sq. ft. building consists of two rooms: a larger unheated room (30 x 46 ft.) and a smaller room (16 x 30 ft.) heated to just above freezing in the winter. The larger room is uninsulated and

is subject to large variations in temperature and humidity; the most scientifically valuable samples are stored in the smaller, insulated, heated room.

The archive building now houses approximately 70,000 samples of soil, water, plant tissue, and other materials. Samples are preserved, barcoded, and cataloged with accompanying metadata in a database. This database of bar-coded samples is searchable online at <https://data.hubbardbrook.org/samples/>. In 2024, both the underlying collection and sample database and the search interface are being restructured. This effort is resulting in improved collection organization and more detailed sample-level metadata.

Requests for reanalysis of archived samples (e.g. isotopic analyses, heavy metals, etc.) are received periodically, and have resulted in a number of publications.

A Sample Archive Committee (SAC) was formed in 2013 to address storage space, future direction, priority for continued bar-coding, etc.

1.10.2 Sample Archive Subsampling Policy

HBR shares these archived samples with scientists upon request. As stewards of these samples, our highest priorities are:

1. to maintain the chemical integrity of these samples;
2. to document the use of these samples, and any resulting changes;
3. inform principal investigators of interest in using them;
4. to acknowledge the appropriate funding sources for their original collection.

Details on the subsampling of archived material can be found on the sample request form:

https://hubbardbrook.org/sites/default/files/documents/subsampling_request.pdf

1.10.3 Directions for Sample Submission

Requirements for acceptance of samples into the archive:

1. Adequate documentation must accompany physical samples.
2. Samples are stored in either an unheated large room or a smaller room that is heated to just above freezing. The contributing scientist is responsible for deciding that these conditions are suitable for his/her samples.
3. Soil samples must be air or oven-dried and stored in plastic or glass bottles with screw caps to ensure a tight seal. Cardboard is not permitted.
4. Vegetation samples should be dried, ground and stored in clear plastic or glass containers.

5. Water samples must be stored in plastic bottles and will be accepted either treated, or not. If treated, the investigator must specify the type and concentration used.
6. All tree logs, cookies and cores should be air-dried and can be stored in cardboard boxes or arranged in a manner that will allow for air to flow between individual samples. Tree cores should be mounted or stored in straws.
7. Samples that are considered toxic may be rejected. The data management committee may confer with the SAC about important, but toxic samples requiring storage.

1.11 Bibliography

The current bibliography for Hubbard Brook includes Books, Journal Articles, Conference Presentations, and Theses; currently with more than 2400 entries.

Citations are managed locally in Zotero (<https://zotero.org>). This open source bibliography management software allows for export to standard reference management exchange formats, harvests citation information easily through a browser, and provides cite-as-you-write support for MSWord and Open/LibreOffice.

The Zotero bibliography is mirrored to the cloud, and publications are accessed through a link to this service. The WordPress zotpress plugin provides the functionality of linking each HB researcher's profile page to their Hubbard Brook publications.

The Hubbard Brook bibliography is also mirrored to the LTER Network Communication Office (NCO) central LTER bibliographic database.

1.12 Computing Facilities

This section is now in a [separate chapter](#).

2 Data Package Workflow

2.1 Overview

The purpose of this document is to capture details of the data package development workflow that is currently in use at Hubbard Brook. HBR data is published in the [Environmental Data Initiative Repository \(EDI\)](#). With the availability of EDI's ezEML data package builder application (adopted by HBR in 2024), this once long and complicated process has been greatly simplified.

In 2024, HBR fully adopted the EDI ezEML workflow for data package development. All data packages are developed under the EDI HBR user account. The division of effort varies with the nature of the data package. Graduate students are encouraged to collaborate with the IM online within the ezEML environment. This serves as a way to directly input metadata without first populating a template, reduces IM time on some data packages, and is an important skill-builder for HBR graduate students. Additionally, some of our investigators have projects beyond the scope of Hubbard Brook and have begun using this workflow to publish data.

EzEML provides the capability to store often-used metadata components in a template. For HBR, there is one master template and additional templates for HB projects that are frequent publishers (MELNHE, HBWater, BIRD, CRCH). Since the master template is very extensive, it is not cloned as a starting point for a new dataset (which might be a common template use), but instead accessed through the import [creator, geographic, keywords, project, funding] buttons where just selected items are brought into the current dataset.

2.2 Data package development workspace

The working directory for package development is on the HBR-IM desktop with the home directory for data package management identified elsewhere as `$DPM_HOME`.

Assets for each data package are in folders named `$DPM_HOME/ezEML/hbr[pkgid]`. While most of the workflow occurs in the ezEML environment, this local filesystem is used to handle dataset assets submitted to IM (metadata templates, datafiles, etc). The completed ezEML packages are downloaded (as zip) to this location for subsequent upload to the EDI staging and production servers.

2.3 Step-by-Step Data Package Workflow

The HBR Data Inventory table is hosted on the HubbardBrook sharepoint site (HBR-IM administrator at UNH). This table contains additional information that we use on our local data catalog to enhance user experience (flagging of significant core datasets, more robust LTER Core Research Area assignments that may be missing in older metadata, and a code to categorize datasets and to sort them in the initial catalog view). Data packages are entered in this table as soon as they are identified (in some cases with very long lead times). Upon becoming aware of a dataset, a package id is assigned and the entry initiated with status=anticipated. As soon as data and/or metadata are in-hand, the status is updated to ‘draft’, ‘cataloged’ when published, and ‘embargoed’ if published with a temporary embargo. The table includes packageID, abbreviated title, contact, notes as needed, flagging as long-term core dataset, and EDI submission status.

The steps are as follows:

- In most cases, a very preliminary zoom with data show-and-tell occurs to iron out any major issues and to develop a plan for the composition of the data package.
- Provide submitter with metadata template using [email template](#) to outline the process and description of best practices, such as:
 - titles with what where when, with the latter YYYY, YYYY-YYYY, or YYYY-ongoing
 - Date formats as YYYY-MM-DD, or timestamps as YYYY-MM-DD hh:mm
 - Make sure submitter understands tidy tables!
 - Confirm that submitter has done rigorous QC and help with this if necessary
 - Authors in the order they should show up on dataset citation
 - Keywords from the LTER controlled vocabulary (and additional if necessary)
 - Note that commonly used chunks of metadata can be briefly described to IM without detail, such as author details, geographic locations, taxonomy. These will be pulled from the template.
 - funding sources go in the project section, and if a new sNSF source is supplied the full details can be fetched by just entering the grant number
- Data package assets are provided to the IM through a sharepoint dropoff (data tables as csv, and ‘other entities’ as necessary)
- ezEML collaboration is established if desired for the dataset.
- A data package is initiated in ezEML
 - naming convention of hbr[pkgid]-[shortname]
 - use document/new or document/new from template
- The HBR Master Template (stored in EDI) is used to import people, geographic areas, keywords, projects, and funding. Add items to the template if they are likely to be used in the future. One-off items can be added directly to this new dataset.

- Data tables are loaded from the author-provided csv.
- Data table attributes are documented either directly on the online forms or through the ezEML table entity templating feature (a great timesaver for complex datatables!).
- Note that in 2025, QUDT units are added automatically by ezEML (you can turn this off, but no advantage to that. In ‘check metadata’ there is a link to review all assigned QUDT annotations, and it is a good idea to quickly scan that.
- When the dataset is ready for staging, that can now be done within the ezEML environment. Enter the packageID and step through the evaluation, staging, and view
- Share the URL for the staged package with the data creator for review
- Subsequent edits are made with creator feedback
- Data package is approved by the creator
- Data package is uploaded to the live repository (again right within the ezEML portal!)

2.4 Non-tabular datasets (images, audio, very large datasets, etc)

Hubbard Brook has published a number of datasets that contain zip files of pdfs, images, audio files, etc. Guidance for preparing these special case datasets can be found in the [EML Best Practices document](#).

2.4.1 Large Datasets

Datasets exceeding the 100Gb threshold are deemed “too large for HTTP” and must be prepared as offline data entities. At the current time, large datasets are staged on a UNH server with the ezEML distribution URL set to that location. In some cases we develop packages using a smaller placeholder file so that we do not overload storage on the ezEML platform or portal-s staging area.

For acceptable size datasets that are nevertheless large, the HBR IM typically checks in with the ezEML developer to make sure that disk space is not an issue at that time. If ezEML upload (or data at a URL) is acceptable at the time, HBR IM can stage and publish and notify ezEML that the files can be deleted from staging.

2.5 Notes on revising older datasets.

When earlier data packages are revised, the starting point is an ezEML fetch of the published data package. Steps are similar to those used for a new dataset, but remember to increment the revision number. Assets for earlier data packages developed can be found in either the ‘EMLassemblyline’ or ‘legacy’ folders, although those files should rarely be necessary once a data package is published in EDI.

An EDI fetched dataset may have been developed with a non-ezEML workflow. If that is the case, items requiring attention will be:

- Creators – delete and import from the template to be sure to get ORCIDs and institution RORs for each person.
- Project – EMLAL datasets may have funding in ‘related funding’ or a text string in project abstract. Delete these. Import all funding from the template as primary or related. The template will have enhanced information to include grant url, funding agency ROR, etc.
- Intellectual rights will be correct for all older datasets, but it is best to reset that in ezEML to CC-BY selection.
- Increment the packageId revision number.
- Use re-upload datatable if revision includes new or modified data. This all goes well if the table is identical. If there are new columns, upload as a new table and clone metadata from the original, then define any new columns. If the dataset was prepared in EMLAL, clear min/max bounds.

2.6 Forest Service Data Workflow

The staff at the Hubbard Brook Experimental Forest manage the entire data lifecycle for many of the long-term datasets. These include hydrology, meteorology, phenology, and others. These data are prepared for the EDI repository using EDI’s EML Assemblyline workflow.

3 Data Catalog Workflow

3.1 Overview

The purpose of this document is to capture details of the workflow to build a data catalog table that is used on the [Hubbard Brook Data Catalog](#) displayed on our website. This workflow has been reconfigured from earlier versions to now read only from publicly available sources - EDI and a public sharepoint file with enhanced data package details. The latter improves the user experience when searching for data by categorizing HBR data and adding robust LTER core area tags.

3.2 Database and File access

Access to dataset details is provided by the LTER PASTA API via the EDIutils R package. The enhanced table resides on the HubbardBrook sharepoint site (HBR IM admin, UNH). The sharepoint data inventory file is maintained to track the status of each dataset and to provide additional information that is not included in the formal metadata or may be lacking in very old datasets but is useful in our data catalog (LTER Core Research Area, HBR significant data status, categorization to organize the catalog browse listing)

3.3 Step-by-Step Catalog Workflow

- Run the code in Appendix A (dataCat.R)
- log in to the wordpress site
- open the data catalog table
- upload wptablefeed.csv to replace existing version

3.4 APPENDIX A – Code to build the local HBR data catalog:

Will be setting this up in github to replace this snapshot here

3.5 FetchSharepointDataInventory.R

```
#####  
# FetchSharepointDataInventory.R  
#####  
  
library(httr)  
library(readxl)  
  
# Function to read Excel file from SharePoint URL  
read_sharepoint_excel <- function(url) {  
  
  # Download the file  
  response <- GET(url, config(ssl_verifypeer = FALSE))  
  
  # Check if the download was successful  
  if (status_code(response) != 200) {  
    stop("Failed to download the file. Status code: ", status_code(response))  
  }  
  
  # Create a temporary file  
  temp_file <- tempfile(fileext = ".xlsx")  
  
  # Write the content to the temporary file  
  writeBin(content(response, "raw"), temp_file)  
  
  # Read the Excel file  
  df <- read_excel(temp_file)  
  
  # Remove the temporary file  
  unlink(temp_file)  
  
  return(df)  
  
}  
  
# URL of your SharePoint Excel file  
url <- "[https://universitysystemnh.sharepoint.com/:x:/t/HubbardBrook/EXFAJdG37Vx0srGI5jMhxn...]"  
  
# Read the Excel file  
df <- data.frame(read_sharepoint_excel(url))
```

```
# subset to just cataloged datasets, save as df 'ps' (for package state)
```

```
ps=df\[which(df$status=="cataloged"),\]
```

```
}
```

3.6 dataCat.R

```
#####
```

```
# dataCat.R
```

```
#
```

```
# 20241016
```

```
# Purpose: build a datatable for the HB website local data catalog
```

```
# This is a revised script to generate the data catalog file on the  
wordpress site
```

```
# this revision now runs on all publicly available data - no pw  
protected databases for local info
```

```
# Inputs are - sharepoint.xlsx file that tracks data status locally and  
provides additional info to enhance user
```

```
# experience in searching for data
```

```
# Requirements: The following script is sourced and should be located  
in the same folder as this main script: FetchSharepointDataInventory.R
```

```
# Usage: Run this script, then move the wptablefeed.csv file to the  
website and updated the table to refresh to this new version
```

```
# Note: once this runs a few times and I gain a little confidence, I  
will consider adding a command to ftp directly to wordpress to refresh  
the site
```

```
#####
```

```
library(EDIutils)
```

```
library(tidyverse)
```

```
library(httr)
```

```
library(\"stringr\")
```

```
library(xml2)
```

```
# set the working directory to the location of the script
```

```
setwd(dirname(rstudioapi::getActiveDocumentContext())$path))
```

```

# source the script that gets local package info from sharepoint
spreadsheet

# returns dataframe ps(package state)
source(\"FetchSharepointDataInventory.R\")

# Fetch the basic eml info directly from EDI for each package
# consider getting abstract and making that display as a hoverover on
the website table

res\<-search_data_packages(

query =
\"q=*&fq=-scope:ecotrends&fq=scope:knb-lter-hbr&fq=-scope:lter-landsat*&fl=id,packageid,do

names(res)=c(\"id\", \"PackageId\", \"doi\", \"Title\", \"pubdate\", \"begindate\", \"enddate\")

res\${Title}=gsub(\"\\[\\r\\n\\]\", \" \", res\${Title})

res\${Title}=gsub(\"\\\\\\\\s+\", \" \", res\${Title})

# extract begin and end YEAR
res\${startYear}=as.POSIXlt(as.Date(res\${begindate}))\${year} + 1900
res\${endYear}=as.POSIXlt(as.Date(res\${enddate}))\${year} + 1900

# if dataset has start/end dates, create column that shows them with
dash separator
res\${yearrange}=\"NA\"
res\${yearrange} = paste0(res\${startYear}, \" - \", res\${endYear})

index= grep(\"NA\", res\${yearrange})
res\${yearrange}\[index\] = \" \"

#create the edilink
res\${edilink}=paste0(\"https://portal.edirepository.org/nis/mapbrowse?packageid=\", res\${Package

# Fetch the pesky keywords and authors as xml so you can insert a
separator

k\<-search_data_packages(

query =

```

```

\"q=*&fq=-scope:ecotrends&fq=scope:knb-lter-hbr&fq=-scope:ltter-landsat\*&fl=id,packageid,ke

as = \"xml\"

)

kw \<- data.frame(
  PackageId=character(),
  Originators=character(),
  Keywords=character(),
  stringsAsFactors=FALSE)

count=1

for (doc in xml_find_all(k, \".//document\")) {

  print(doc)

  # Get the keywords from the current doc
  keyword_elements \<- xml_find_all(doc, \".//keyword\")
  keyword_strings \<- xml_text(keyword_elements)

  # This doesn't get the authors where HBWATER and USFS are institution
  authors.
  # see solution below to get those from cite.edirepository.org

  kw[count,3\] \<- paste(keyword_strings, collapse = \";\")

  # Get the authors from the current doc
  author_elements \<- xml_find_all(doc, \".//author\")
  author_strings \<- xml_text(author_elements)

  kw[count,2\] \<- paste(author_strings, collapse = \"; \")

  \# get the packageid
  pid \<- xml_find_all(doc, \".//packageid\")
  pidstring \<- xml_text(pid)
  kw[count,1\] \<- pidstring
  count=count+1

}

# tidy up keywords where some strings have newlines or consec white

```



```

spaces
kw\$Keywords=gsub("\\[\\r\\n\\]", " ", kw\$Keywords)
kw\$Keywords=gsub("\\\\\\s+", " ", kw\$Keywords)

# merge the tidier keywords with the main table
resj=merge(res,kw, by = "PackageId")

# get the dataset citations so that you have a nicer listing of authors
# you could probably do that where I do the keywords now from xml, but
it doesn't

# populate the records where author is an institution (USFS and
HBWaTER)

for(i in 1:dim(resj)[1]){

print(i)

CMD=paste0('GET
("https://cite.edirepository.org/cite/',resj[i,1],'\\"')\')

# sleep can be removed if you are whitelisted to make rapid EDI queries
Sys.sleep(0.5)

print(CMD)

c=content(eval(parse( text = CMD )))

print(c\$authors)

resj[i,"Originators"]=c\$authors

}

# FetchSharepointDataInventory returns ps dataframe (aka package state
in the originalmetabase database)

# add data category based on sort order codes in ps (local package
state table)

# apply category name to sort order values

ps\$category=0

```

```

index=which(substr(ps$pub_notes,1,1)==1)
ps$category[index\]="Hydrometeorology\"
index=which(substr(ps$pub_notes,1,1)==2)
ps$category[index\]="Water Chemistry\"
index=which(substr(ps$pub_notes,1,1)==3)
ps$category[index\]="Soils\"
index=which(substr(ps$pub_notes,1,1)==4)
ps$category[index\]="Vegetation\"
index=which(substr(ps$pub_notes,1,1)==5)
ps$category[index\]="Heterotrophs\"
index=which(substr(ps$pub_notes,1,1)==8)
ps$category[index\]="Documentation\"
index=which(substr(ps$pub_notes,1,1)==9)
ps$category[index\]="Spatial Datasets\"

# subset out the columns that are to be used in the datatable
pscat=ps[,c(\"dataset_archive_id\", \"category\", \"ltercore\", \"pub_notes\")\]

# merge the EDI query dataframe with ps
m=merge(resj,pscat, by.x=\"id\",by.y=\"dataset_archive_id\")

##### Write out wordpress data table #####

# pull out the columns needed for website table
wptablefeed=m[,c(\"PackageId\", \"Title\", \"Originators\", \"yearrange\", \"ltercore\", \"edili

# sort packages based on pub_notes
wptablefeed.order=wptablefeed[order(wptablefeed$pub_notes),\]

# write out the table
write.table(wptablefeed.order, \"wptablefeed.csv\", row.names=FALSE, sep=\",\", na=\\
\\")

```

4 Data Inventory

A spreadsheet listing a complete inventory of HBR data, including datasets anticipated and in draft format, is maintained on the UNH HubbardBrook sharepoint site. Access to this sheet is shared by request. This is used to build the local data catalog by merging with content from the EDI PASTA+ API. It is also used by IM as a local record of what is published, as well as those datasets anticipated, in draft format, or staged for review. The emlWorkflow is used to identify datasets developed by earlier workflows that could be updated to include eml elements not available at time of publication (improved funding metadata, general annotations, qudt unit annotations). The checkbox columns for HBR_V were used to modify our data table listing (proposal supplemental document), by highlighting data used in papers during that funding cycle and data used in the top10 publications that we highlighted in the proposal.

Shared access to this data inventory table is view-only. Edits to the table are limited to the site IM team. In view-only mode, some useful search/filter includes filtering by project ID (used in the 'nickname'), by dataset status, core research area, etc. This provides a comprehensive look at our data holdings to include datasets currently under development as well as those anticipated on a longer timeline.

- DataSetID: just the pkg number
- dataset_archive_id: packageId (knb-lter-hbr.XXX) assigned by IM
- rev: revision number
- nickname: a short name for the dataset
- status: dropdown choice of anticipated, draft, staged, cataloged, deprecated, embargoed
- emlWorkflow: legacy, EMLAL, ezEML, mmb (minimetable)
- notes
- pub_notes
- dbupdate: carryover from mmb
- update_date_catalog: date of last published revision
- who2bug: contacts
- in_pasta: binary 0/1 to indicate published
- signature: do we consider this a core HB longterm dataset
- ltercore: core research areas (DP(disturbance process),IM(inorganic matter),OM(organic matter),PP(primary production),PS(population study)
- hbr_vcited: data cited in HBR-V publications
- hbr_vtop10: data cited in HBR-V top 10 publications
- longterm data category: to further organize core longterm data. helpful when evaluating status and reaching out to investigators

5 Publication Management

5.1 Overview

Hubbard Brook publications are managed in Zotero. The primary database is on the IM desktop where publications are initially ingested. The full bibliography can be found [here](#). Note the use of tags below for further exploration of this library.

5.2 Zotero - step-by-step

- By using Zotero on the UNH network or via VPN, there is full access to (most) publications through the UNH Library and those are saved with each citation.
- publications are identified through self-reporting to IM, google scholar searches, Hubbard Brook Monthly newsletters, and HBR LTER Annual Reports.
- each update period is managed in a dated folder in zotero
- by clicking zotero icon on pub landing page, it is ingested into Zotero (full citation plus accessible pdfs). Some publications require InterLibraryLoan, which typically takes < 24hours to receive the pdf. Those are then added to the citation with rightClick/addAttachement.
- tagging identifies 'LTER', 'data-only', and last_firstinitial tags for each HB author (the latter facilitates Zotpress links to each person's profile on the website). Additional tags are HBR_dataonly, HB_related, and archived_samples.
- once tagged, the new pubs are copied to the HBR-LTER group folder. At one level above this is 'Hubbard Brook Ecosystem Study' - pubs that are not HBR-LTER are added to the group at that level.
- the group folder is synced to the hubbardbrook Zotero cloud account. Syncing is set *not* to upload the pdfs, as this can violate copyright. Publications since ~2010 contain DOI links.
- an ongoing effort is underway (2025) to add DOIs and attached PDFs to the Zotero records. DOIs were issued beginning in 1997, but publishers have retroactively assigned DOIs to older publications. For example, Leonard, R. E. (1961), Net precipitation in a northern hardwood forest, *J. Geophys. Res.*, 66(8), 2417–2421, doi:[10.1029/JZ066i008p02417](https://doi.org/10.1029/JZ066i008p02417).

- This article was the motivation to capture digital articles. We do have paper copies at Pleasant View Farm from inception through 2000. By archiving digital copies, these older publications become more accessible to our research community - they can be provided by the HBR IM upon request. <https://www.nature.com/articles/d41586-024-03842-z>

5.3 EDI publication-data linkages

The EDI repository facilitates the linking of publications with underlying data through the ‘add journal citation’ feature. For HBR we add all known linkages. The EDI team has a semi-automated process where they identify new data DOI’s in the literature and then they add the pubs to the datasets. This is not 100%, as some papers still slip up on proper data citation, or data publication lags paper publication. In this case, IM is on the lookout for instances where manual additions need to be made. Carefully documenting the use of data in publications in this way provided important metrics on the value of our data. Each paper added to the bibliography is reviewed for underlying data and IM confirms that the pub is associated with the dataset in EDI, and if not, adds the article.

5.4 Zotpress

This wordpress plugin on our website is linked to the cloud-based HB Zotero account and to displays publications on the people profile pages on website through tagging of last__firstinitial for each author on a publications.

6 Physical Sample Archive

6.1 Overview

The Hubbard Brook Experimental Forest (HBEF) is operated and maintained by the USDA Forest Service, Northern Research Station. Research at HBEF began in the 1960's and the research community consists of investigators from numerous institutions across the US (and world). Collectively, this research has been funded by the USDA Forest Service, National Science Foundation, Department of Energy, US Geological Survey, other federal grant programs, and private foundation funds. Research projects have generated a wealth of samples, preserved in either the on-site sample archive facility located at HBEF and/or numerous institutional collections (biological voucher samples).

The Hubbard Brook Physical Sample Archive was completed in 1990 to permanently store research samples for future use in Hubbard Brook studies. The building provides a special and unique opportunity to maintain a research memory for long term ecosystem studies. To date there are 70,000 barcoded and cataloged samples ranging from water samples to soils and plant parts. Samples from ongoing, long-term collections and shorter-term projects are included. Cataloged samples are accessible on-line in a public searchable database and are available for subsampling in support of new studies and additional analyses. Stored and cataloged samples have been used to address scientific questions and have resulted in additional publications featuring archived sample analysis. [https://www.zotero.org/groups/4549695/hubbard_brook_ecosystem_study/tags/archived_samples/library]

The archive is managed to accommodate future sample submissions. Samples from Hubbard Brook are accepted into the archive system if they are likely to be of future use and are accompanied by adequate documentation. Detailed metadata is submitted with each sample and stored in a relational database. The Hubbard Brook website provides browse/search/filter access to the sample database at both the collection and sample level.

External collections have been used extensively since the 1960's for the preservation of biological voucher specimens. A 2024 assessment of HB voucher samples identified more than 2000 samples in 40+ collections in the US and Canada. With the growing acceptance of digital (image/audio) vouchers, this number grows to several hundred thousand. Management and policies for samples and specimens in external collections are addressed in a separate document. See [voucher sample chapter](#) for more details.

6.2 Sample archive database

The sample archive database is managed by the Hubbard Brook USFS. Sample metadata are processed from submission spreadsheet templates (L0) to harmonized tables (L1) to appended files for accessions, collections, and samples (L2). The workflow and products occurs in the USFS Pinyon (BOX) environment. L2 files are staged with URL access for the Rshiny sample search app.

6.3 Templates/Forms

Templates for submitting a new accession can be found [here](#)

Sample subsetting request form can be found [here](#)

6.4 Rshiny

The interface for searching and downloading samples of interest is developed in Rshiny. This app allows search at the collection level and full sample search across all collections. The R shiny code is in a separate [GitHub repository](#)

6.5 Management Plan

Under development and maintained in a seperate document, to cover:

- Archive Location and Infrastructure
- Sample Accession – Policies and Procedures
- Sample Identification and Documentation
- Data Management
- Metadata lifecycle
- Access and Permissions
- Disposal and Deaccessioning

7 Voucher Specimens

7.1 Background

There has been a decades-long history of voucher collections at Hubbard Brook. These were typically ad hoc collections, with no coordinated effort in the destination collections, or reporting to HB leadership. The end result is that several thousand voucher samples can be found in over 40 collections in the US and Canada.

The [2024 LTER solicitation](#) includes specific language requiring a plan for handling and potentially archiving specimens collected by LTER sites.

“For proposals that include specimen collection: If collecting or generating specimens (e.g., organisms, parts of organisms, fossils including trace fossils, microbial isolates, etc.) is proposed, the Data Management Plan must include a description of how the specimens and associated data will be accessioned into and maintained in an established biological collection.”

Although current HB projects not generate many grant-related voucher specimens, this was a good nudge toward determining what had been collected, and to develop recommendations in the event that a grant were to be submitted that required voucher language in the data management plan.

7.2 What samples have already been submitted to collections?

This opened a huge can of worms! Searching was done in a variety of manners - known collections, collection clearinghouses (iDigBio, Symbiota thematic portals), personal outreach, etc.

This spreadsheet contains those samples found to date: [Voucher sample draft inventory](#) This is just a draft of ‘notes’ on existing collections to be reformatted and enhanced to be more of a ‘finding guide’ to samples that may exist now in collections. It is evident that there are some key collections that house Hubbard Brook samples and some of these have well-established relationship with Hubbard Brook. This makes them likely collection candidates for future samples to be collected. This is a work-in-progress and more details are forthcoming as we develop best practices recommendations for the site.

7.3 Recommendations for future sample submissions

**** under development****

8 Personnel Directory

Describe tracking, grad/postdoc lists, updates to LNO

9 Hubbard Brook Listservs

9.0.1 Establishment, configuration, and status

Updated: April 30, 2025

In the early 2010's COS members requested the establishment of listserv(s) for the Hubbard Brook Community. These notes describe the configuration of the lists.

9.0.2 The Lists

Three primary lists were established in 2014.

- HubbardBrook – a general list comprising HB investigators, grad students, techs, and those generally interested in HB work
- HubbardBrookCOS – all members of the COS
- HubbardBrookGrads – all graduate students/postdocs, SCC chairs, LTER leadPIs, HBR IM
- Several project-level lists have been established – for instance, for the Ice Storm Team

Membership

- Control of list members is by the admin – no self-addition to the list (that opens up a world of spam). Only members can post to lists. If a non-member posts, it is sent to admin for approval and release.
- Addition of list members is by request to the list admin.
- The general list contains everyone and there has been no removal (or very little) as folks have moved on from HB
- Grads are kept up-to-date with additions/removals on both the general and grad list.
- IM reaches out to COS several times a year to review grad list and provide updates.
- SCC informs the admin of new members when they are voted in to COS.

- Some requests for addition come from undergrads, and admin has been adding notation of ‘undergrad’ and ‘year’, so that these can be removed periodically.

9.0.3 Use patterns

In general, the lists have been used for announcements and there has been very little 2-way communication. Initially, this type of list was chosen, since 2-way communication was a requirement in the request from COS to establish lists. Is this something we still need....

9.0.4 Listserv platforms

The lists were established and managed by the HBR IM on the mailman listserv platform hosted at UNH-RCC: <https://list.org/> <https://lists.sr.unh.edu/mailman/listinfo> Tech support has been provided by UNH Research Computing Center. Hubbard Brook used this until 2025 until the service was no longer available at UNH. See below for current status

9.0.5 The Future of Lists

In 2024, a decision was made at UNH to discontinue the mailman list service and all those with lists were asked to migrate to a new system.

Transitioning to Hubbard Brook Lists 2.0

UNH RCC is recommending that those with lists transition to UNH Outlook distribution lists. The lists were migrated to Outlook and this is not ideal - due to an issue with DMARC security configuration, some recipient institutions reject messages. This is not an issue if messages are sent from a UNH address. This has limited the functionality of the list for two-way conversations. This is a configuration that would need to be changed at a high UNH IT level. A ticket has been submitted, but no resolution after a year.

Other options have been investigated - using google groups through the HBRF workspace. The limitation here is that the address is either lastname@hubbardbrookfoundation.org or lastname@googlegroups.com, neither is ideal. Although googlegroups does not require a google account to be a member, there can be some resistance to this. Using the HBRF domain is not ideal as they manage a number of different mailing lists/services and there is already some confusion between list and other mailings.

For now, it is UNH Outlook with its limitations. There is some possibility of a UNH revival of the Mailman service which would be ideal, as that worked well for a decade.

10 Website

The Hubbard Brook website is developed in WordPress and hosted on Bluehost. Admin access is granted to the HBR IM and to the WordPress consultant. Details on the setup of the site, hosting account access, domain name management, plugin accounts and codes are all maintained off-line and can be accessed by the HBR IM as necessary.

This document is limited to tools and tips for maintaining the current content and for adding new content.

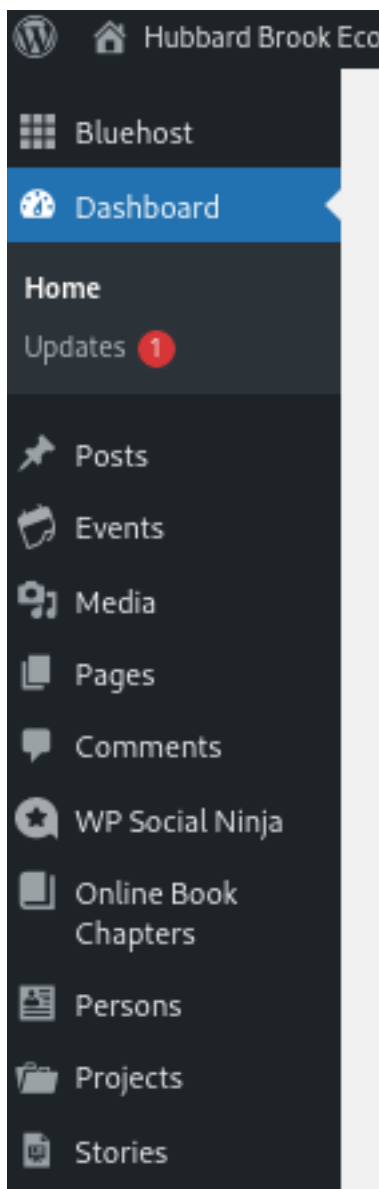
Instructions for Adding Website Content

10.1 Content editing access

Log in to the site at hubbardbrook.org/wp-admin. HBR IM is the admin for this site, other content creators can be added as necessary

10.2 Adding Feature Stories

Feature stories are added as a custom post type - locate 'Stories' in the left sidebar, and 'add new'.



In the add new window, enter in a title for the story.

Also enter a featured image. This can be an existing element in the media library or new image that you upload. See notes on general tips for adding [media](#). Since multiple stories are displayed in a ‘feed’, this will look best if you use an image with the general aspect of 2wide x 1 high.

Featured Image ^ v ▲
[Set featured image](#)

Add New Story

Add title

Text Editor

Beaver Builder

Add Media

Pods Shortcode

Paragraph ▼

B

I

≡

≡

“

≡

≡

≡

🔗

≡

📄

<> ▼

🔧

Once the title and featured image are entered, select the ‘Beaver Builder’ tab. Then [see section](#) on general instructions for Beaver Builder layouts.

10.3 Adding News posts

News posts are usually pretty straight forward and simple format. Sometimes just pointing to something in external media. These do not use Beaver Builder, just the regular text editor. If the item is more involved and media-rich, consider making this a ‘Feature Story’.

- Add new post.
- Add title
- Add featured image - [see image instructions](#)

- Enter news article in the text editor, not beaver builder.

NOTE: when transferring older news articles from the drupal site, use this checklist to indicate status (y=now on wordpress; changed to a feature story; save archive, but do not put on wp site; etc)

[https://docs.google.com/spreadsheets/d/1yPRhX9DHN1tL8_EQAt9WCQloV5xwf4tymI009C2Pfic/edit?usp=](https://docs.google.com/spreadsheets/d/1yPRhX9DHN1tL8_EQAt9WCQloV5xwf4tymI009C2Pfic/edit?usp=sharing)

10.4 Creating new pages

Pages are typically used for more static website content. Pages are typically accessed through the high level menu, or from links within existing pages

Add new page

Add title

Enter Beaver Builder mode to construct page

10.5 Adding new People pages

10.6 Adding Lessons

Add new page

Set page type to k-12Lesson

Set parent to k-12 Education Program

Enter Beaver Builder

Select Templates/SavedTemplates/LessonTemplate

Enter descriptive text in left column

Enter lesson assets in right column

10.7 Adding content to the media library

Media can be added directly to the media library by going to media on the left sidebar dashboard, or it can be added at the time of page/post editing by using the ‘add media’ button in the beaver builder text editor.

JPG are preferred format

When uploading most types of graphics and photos, use the JPG/JPEG file format as it uses the least amount of space. If you take a screenshot, which creates a PNG, you can use any basic photo editor to change to a JPG.

Naming of media files

Try to use a filename that is descriptive, but not too long. This can be super helpful when working on the site and searching for photos to use in layouts. For example, if I’m working on a page and need a photo of a person, it’s super easy to find those photos via the media library search when the filenames are something like “jane-smith-2022portrait.jpg”.

Uploading PDFs for general use (not using PDF viewer)

PDFs uploaded via ‘media’ button in the beaver builder text block will display image of pdf as a pdf viewer. If you just want a link to the pdf to to the Wordpress Dashboard, go to the Media menu item, select ‘Add New’ and upload. Then in text block, just add it as a link (do not use media button).

10.8 General tips on Beaver Builder (BB) layouts

BB General Documentation

<https://docs.wpbeaverbuilder.com/beaver-builder>

BB Rows

<https://docs.wpbeaverbuilder.com/beaver-builder/layouts/rows/work-with-rows>

BB Modules

<https://docs.wpbeaverbuilder.com/beaver-builder/layouts/modules/module-overview>

BB YouTube

<https://www.youtube.com/c/BeaverBuilderWP/featured>

10.9 Setting publication status and visibility

10.10 Theme colors

10.11 Headings

When designating heading sizes on pages, there should only be one H1 heading per page/post as the H1 is designated for the page title. This is automatically done in the title/BCs/search bar, so when working on pages, you should only be designating H2 or smaller (and you can use multiple H2s and smaller per page/post). Here is an example:

<https://hubbardbrook.org/policy-briefings-youth-climate-forums/>

- Policy Briefings & Youth Climate Forums in the title/BCs/search bar gets the h1 designation

- the Policy Briefings text and the Youth Climate Forums text in their sections gets h2 because those are the second most important headings
- other headings can get h3 or smaller depending on their importance

An article about this:

<https://yoast.com/how-to-use-headings-on-your-site/>

10.12 Migrating content

When cutting and pasting text from documents or other websites, it is important to switch to the ‘text’ tab (not ‘visual’) in the text editor. This will paste the content in plain text, and not drag along any formatting bits from the originating document.

11 Online Book

This document is managed elsewhere - migration to this quarto site is pending

11.1 Managing chapter content

Describe how the chapters are configured in Wordpress, how updates are made.

11.2 Automated graphs

Describe how this is done and point to the repository where the graphs live and are linked to the book

12 RAC Proposal Submissions

12.1 Overview

Proposals submitted to the Hubbard Brook Research Advisory Committee (RAC), are managed through the JotForm account at the Cary Institute. Calls for proposals go out to the research community for March and September deadlines. Information on the process is found on [proposal submission page](#) on the Hubbard Brook website.

12.2 IM role in RAC submissions

Beginning in Spring 2023, the RAC proposals have been submitted through the Cary institute Jotform. Prior to this, the submission pathway was via website forms hosted at UNH (2008-2014) and google forms (2015-2023). The structure of the form was migrated to JotForm, and from 2023-2024 a number of enhancements have been made. The Hubbard Brook sharepoint RAC folder contains a backup of all proposals. Admin access to the JotForm can be obtained from IT support at the Cary Institute.

Form related tasks are carried out by the HBR IM (under direction of RAC chair) to include:

- Add enhancements as requested by RAC chair
- Monitor incoming proposal
- Prepare proposals for review by RAC (export individually as pdf, and in an appended pdf format).
- HBR IM reaches out to accepted proposal leads to let them know how we archive data in the repository at HBR, what the IM needs to know about their data collection, and what assistance can be provided to them by the IM. See [email template](#)
- JotForm technical details in appendices below
- NOTE: Access to all submissions in the JotForm environment can be made available to the RAC chair. In this scenario, there are enhanced features where the proposals could be flagged as ‘under review’, ‘accepted’, etc. And the chair could manage their own downloads of proposals.
- The new RAC chair in 2025 has monitored JotForm and handled proposal exports.

12.3 RAC data publishing instructions

Although not in the form itself, the following text is used in the RAC proposal acceptance letter:

The Hubbard Brook Ecosystem Study preserves and shares data through the Environmental Data Initiative Repository (EDI). Projects supported through NSF LTER funding, follow the LTER data access requirement. In order to maintain a one-stop data catalog for all site research, we encourage all researchers, regardless of funding source, to preserve their data in a similar manner. Support for this process is available through the Hubbard Brook Information Management Team. Please contact Mary Martin about preparing your data. Please let me know once you have met with Mary and addressed any data format questions that you have.

12.4 Jotform Tips

- Edit form allows for insertion and editing of form blocks as necessary.
- More/submissions provides access to all submitted proposals
- On the top left, the pdf editor allows you to check all fields to be included in the report. By default, this does not include the file upload links. Scroll down to this, check file upload and this will now be inserted in each proposal pdf output.
- The links to uploaded assets that are included in with the submission are publicly accessible by link (uploaded images are displayed in the pdf), so reviewers do not need any special permission to view these uploaded files. Uploaded images are displayed in the pdf, other filetypes by link.
- Download All will create a zip file with each proposal in a separate file.
- To append all files, use: `pdfjam file1 file1 file3 ... -o RAC-proposals_YYYY[spring,fall].pdf`

13 Reporting scripts

13.1 Publication Status

This section will describe the reporting we do on publications and link to GitHub repo with scripts as well as current versions of output. Full chapter content pending.

13.2 Data Status

13.2.1 Status of data in EDI

Scripts are available to generate an automated table of datasets in EDI. The basic configuration is that of the supplemental table required in LTER renewals. Modifications can easily be made for other purposes/formats.

13.2.2 Status of All data

This status report is derived from our Data Inventory. A script produces an output table that removes some of the columns in the table that are for day-to-day IM data management.

13.2.3 Data access reports

A series of reports are generated through EDIutils to show datasets in ranked order of downloads and a timeline of daily downloads. This report can be generated for all of Hubbard Brook, as well as for individual projects (CCASE, MELNHE, Bird, etc).

13.3 GitHub Repository for scripts

Data scripts primarily use the PASTA+ API. Publication scripts are based on the Zotero bibliography and the EDI Journal Citation feature that links publications to data.

14 Templates

14.1 Email

14.1.1 New community members.

Confirm addition to listserv, invite to have a profile page, introduce IM support

- Access [Email template](#)

14.1.2 Email to RAC approved investigators

- IM introduction following RAC proposal acceptance

14.1.3 Emails to data submitter

- Initial instructions
- Invitation/instructions to review draft dataset
- Confirmation of published data with information on citing and revisions

14.2 Data Submission Template

- add link to template

14.3 Sample Archive Submission Template

- add link to template

15 Computer Resources

A section in the Data Management Plan describes the software/hardware regularly used by HBR IM. It also covers IT support and a list of contacts who have access to accounts used by IM (UNH systems, EDI, cloud-based accounts. [Link](#)

15.1 Computer Resources/Facilities (UNH/ESRC)

The Earth Systems Research Center's (ESRC) Science Computing Facility (SCF) has a wide range of computer servers, printers, plotters, archiving systems, software, data archives, and web based data distribution systems that are integrated using several internal networks and connected to the outside world through a high speed pipe. The overall SCF administration is provided by the Research Computing Center (RCC) located in the Institute for the Study of Earth, Oceans and Space (EOS). Scientific data processing and analysis support is distributed throughout workgroups within the center with additional centralized expertise provided by ESRC's Laboratory for Remote Sensing and Spatial Analysis. RCC's Lenharth Data Center was upgraded in September 2011. This upgrade brought in new APC UPS units, energy efficient in-rack cooling systems, and monitoring hardware to provide more space and capability for future growth. Within this proposal, we take advantage of this existing computer infrastructure, to meet our anticipated computational needs.

The main ESRC servers consist of high-end, multi-processor computing systems manufactured by Dell. The Dell systems run Linux and are used for CPU intensive jobs, parallel modeling, and storage. Backups and archives are done using BackupPC a disk-to-disk based system. Most of the main servers share a gigabit (Gb) switch with the archive/backup system for high-speed communications. All of this equipment is kept within a physically secured, humidity and temperature controlled data center complete with closed circuit video surveillance and a remotely monitored security alarm system. Final data and image products are produced from several ink-jet plotters and laser printers within the department. Additionally, several CD/DVD writers are used for data distribution.

EOS includes a CRAY XE6m-200 with 132 nodes, over 4000 processors and 160Tb of storage. In addition, our infrastructure has been strategically upgraded to provide gigabit networking to desktops.

Individual scientists and research groups have additional computing resources at their disposal. These include dedicated servers, individual workstations, and various peripheral devices. The

group servers and individual workstations include: Linux servers and workstations, Windows workstations, Apple Macintosh desktop and laptop computers. All servers, user systems and networked peripheral devices are accessible within EOS through a gigabit ethernet network. Wireless networking is available throughout the building, including access to EduRoam. These systems also have access to both Internet 1 and Internet 2.

ESRC currently houses a 65TB+ geographically referenced data archive used for spatial data processing and analysis. This archive stored on RAID5 data disks served by a series of data servers, houses global, regional and local, Landsat, MODIS, IKONOS, Hyperion, ASTER, and SPOT satellite imagery, land cover classified products, vegetation and other indexes (EVI, LSWI, NDSI, NDVI, NDWI, LAI), aerial photography and GIS vector data layers for use by all projects within the department. Portions of these data, processed data products, and project results archive are disseminated and distributed through several dozen regularly updated and maintained ESRC operated websites. These websites are served through a variety of web servers running Apache web server software supported by other applications and libraries such as Tomcat, Web Mapping Server (WMS), OpenLayers and other geographically enhanced libraries such as GDAL, PROJ4, and GCTP.

ESRC also leverages the center's Laboratory for Remote Sensing and Spatial Analysis, a spatial information processing, analysis and distribution research laboratory. This laboratory provides geographic information system (GIS), Web Mapping, spatial data archiving, data distribution, remote sensing, image processing, cartography, large format printing and scanning support to several ESRC and EOS research projects. Staffed by professional geo-spatial information technicians, computer programmers, and graduate and undergraduate university students, the laboratory houses a multiple seat Linux, PC, and Mac OS computer cluster supplied with a variety of open source Remote Sensing, GIS, web mapping, image processing and cartography software and ESRI ArcGIS, Leica ERDAS Imagine, and IDL/ENVI, commercial site, block, and individually licensed GIS and Image processing software.