

# HIGHLANDS BIOLOGICAL STATION DATA MANAGEMENT POLICY AND GUIDELINES

1. Highlands Biological Station is a year-round field station with a strong tradition of scientific research, teaching, and training. Data collected at HBS fall broadly into two categories. The first consists of data related to the history and operation of the field station, including weather data (1961-present), site history and descriptive data, scientific publications, and biological collections. These data are updated annually and are managed and archived in Excel Workbooks by the HBS Executive and Associate Directors. Some of these data, such as HBS weather records (highlandsbiological.org/weather-data) and publications (highlandsbiological.org/publications) are currently made available on the HBS website.

2. The second category of HBS data consists of data collected by independent investigators pursuing research based at HBS or through HBS on nearby National Forest, National Park, or private lands. A broad spectrum of field and laboratory research is conducted at HBS, supported by private, State, and Federal sponsorship. Researchers from institutions across the country and internationally collect, analyze and publish their HBS research according to disciplinary standards. The PIs conducting this research are in the best position to organize and archive the resulting data, and evaluate data sensitivity and appropriate data sharing.

2a. HBS requires that researchers supported with Grant-in-Aid funding from the Highlands Biological Foundation, Inc. submit a final research report that summarizes their findings and includes research/field site descriptors, maps / plots, data set descriptors, and specimen descriptors (species, samples, physical repository). Hard copies of these records are secured in file cabinets located in the Archives Room in the Reinke Library at HBS, and, since 2007, they are also archived as PDFs by the Executive and Associate Directors.

2b. HBS will seek support for scanning and databasing its decades of hard-copy research reports, with the goal of summarizing metadata (researchers, dates, locality information, plots, investigations, study organisms, etc.) in Excel spreadsheets. These will ultimately be made available on the HBS website for consultation by future researchers.

2c. HBS will work toward databasing and disseminating site and collections data. The HBS herbarium has already been digitized and the collections data curated through the Southeastern Regional Network of Expertise & Collections (SERNEC; sernecportal.org). A longer-term goal is generation of databases pertaining to sites and project/organism metadata, ultimately made available via the HBS website.

3. While HBS does not currently require independent PIs working out of HBS facilities to archive copies of their data and metadata, as a matter of policy PIs are encouraged to comply with NSF and OBFS guidelines and recommendations for the archiving and sharing of datasets and metadata. Accordingly, HBS researchers are asked to submit their raw and metadata to an appropriate archival repository within a year of publication, following the standards recommended by Michener et al. (1997). PIs are further asked to notify HBS of their chosen repository, in addition to providing the Station with copies of any publications stemming from their work at HBS. (See Data Management Guidelines, below.)

3a. It is understood that, according to disciplinary norms, many PIs submit their data to a public repository (*GenBank, TreeBASE, Dryad*) or as online Supplemental Materials to their scientific publications, as a matter of course. PIs working in disciplinary areas where this is not the norm will be provided with information on data archiving information and resources, such as *DataOne* (www.dataone.org), Environmental Data Initiative (environmentaldatainitiative.org/) and *DMPTool* (www.dmptool.org). The HBS budget does not allow for a dedicated Data Manager, so **researchers are responsible for the quality of their data, and submitting their data and sharing this information with HBS in a timely manner.** In keeping with standard practice, PIs are the best judges for the sensitivity or confidentiality of their data, and it is understood that withholding data or metadata is sometimes necessary (e.g., specifying localities of threatened and endangered species).

#### Data Management Guidelines for Highlands Biological Station PIs

With thanks to the Organization of Biological Field Stations (OBFS.org) and Pymatuning Laboratory of Ecology (www.biology.pitt.edu/facilities/pymatuning)

Every scientific discipline has protocols for conducting research, and the following items are offered as guidelines to consider when conducting research at or through HBS. These guidelines constitute part of HBS's Data Management Policy (DMP), the goal of which is to emphasize reproducibility, transparency, and accountability in key aspects of HBS-based research.

## 1. Recording and maintaining data

1a. Research activity and data collections should be kept in a single, waterproof laboratory notebook. Entries should consist of experimental design, calculations, data records, experimental notes, and summaries of discussions concerning the research. All entries should be legible and clearly marked with date of entry. Undergraduate assistants are allowed to record in this notebook under strict monitoring by the PIs and co-PIs for clarity, legibility, and accuracy. No revision of entries should be permitted, although future entries may reference mistakes in any past entries with clear explanation for the reason of the mistake. Backup photocopies of the notebook entries should be made on a regular basis in accordance with lab protocol.

1b. Frequently, raw data measurements will be taken in the field on waterproof paper. Such datasheets should be kept in a binder designated for experimental measurements. Datasheets should be referenced in the lab notebook whenever taken. Back-up hard and digital copies of datasheets should be made on a weekly basis in accordance with lab protocol.

1c. Within a reasonable amount of time, all raw data should be added to a single Microsoft Excel worksheet with clear identifiers regarding the type of data and when the data were collected. Digital backups should be made on a regular basis via an external hard drive. Prior to being archived, digital and hard copies of all data should be held in at least two different locations for security and redundancy.

1d. It is recommended that research notebooks, datasheets, and all back-up copies are retained for a minimum of seven years after final reporting or publication, consistent with Federal and State policy.

### 2. Notification of publication

2a. When a publication results from a research project, the PI should contact HBS personnel (Director or Associate Director) in a timely manner and provide a PDF copy of the publication. \*<u>This is required of PIs receiving HBS Grant-in-Aid funding for their research.</u>\*

2b. Students are encouraged to provide a *bound copy* of their thesis/dissertation resulting from research based at HBS, to be placed in the Reinke Library, and **are required to provide a digital (PDF) copy for the HBS publications webpage.** 

### 3. Archiving data

3a. Any set of archived data includes both primary data and metadata. Following the standards recommended by Michener et al. (1997), the metadata for each study should include the following:

- Data set descriptors (i.e. abstract, key words)
- Research origin descriptors (i.e. PIs, funding, objectives, methods, experimental design)
- Data set status and accessibility (i.e. date of last modifications, data use restrictions)
- Data structure (i.e. file type, description of response variables, data type)
- Supplemental descriptors (i.e. data forms, quality control procedures, publications)
- Location(s) of physical specimens (preserved organisms, DNA, etc.)

3b. Within one year of publication, PIs should archive all data and metadata in an appropriate archival repository (e.g., *Dryad*, *GenBank*, *TreeBASE*, *NCEAS Data Repository*, Environmental Data Initiative etc.) or as a supplement to the manuscript.

3c. Once archived, HBS requests that **researchers send the HBS Associate Director a summary and links to the archived data.** These links will be placed at the end of each publication's entry on the Research/publications page of the HBS website.

3d. As with all data archiving, consideration must be given to the dissemination of any confidential or private data.

### 4. Voucher specimens

4a. Where relevant, voucher specimens, along with associated label data, should be prepared following standards established by repository museums. DNA vouchers should be preserved in -80°C freezers dedicated to specimen preservation. Such specimens may be stored in HBS ultralow freezers initially, but should be transferred to the PI's home institution for long-term storage in a timely manner.

Michener, W. K., J. W. Brunt, J. J. Helly, T. B. Kirchner, and Susan G. Stafford. 1997. Nongeospatial metadata for the ecological sciences. *Ecological Applications* 7(1): 330-342.