

NSDI Cooperative Agreements Program

Metadata Implementation Project Final Report

Agreement Number: 04HQAG0148

Organization:

PRBO Conservation Science
4990 Shoreline Hwy
Stinson Beach, CA 94970
www.prbo.org

Project Leader:

Viola Toniolo, (415) 868-0655 x 308, vtoniolo@prbo.org

List (if appropriate) **collaborating organizations:** none

Project Narrative:

We were able to create FGDC-compliant metadata for ten spatial datasets, four from the Terrestrial Ecology Divisions and six from the Wetlands Ecology Division. Metadata files for each spatial dataset were first created using ESRI's ArcCatalog metadata tool. We then used the NPS Metadata Tools extension for ArcGIS (found at http://www.nature.nps.gov/im/units/mwr/gis/metadata/metadata_tools.htm) to examine the metadata for errors and export it to XML format, edited the XML files manually to add the biological portions of the metadata profile (description of geographic extent, taxonomy, and methodology), and then used the NPS extension to examine it again and export it to HTML format. All of the metadata are thus FGDC-compliant and consist of a full biological data profile (where applicable).

Terrestrial Ecology Division databases include data from various standardized bird monitoring protocols, including nest monitoring, point count surveys, area searches, and mist-netting, in habitats ranging from riparian to tidal marsh, grassland, coniferous forest, shrubsteppe, desert, and oak woodland. They are associated with landbird ecology, reproduction, population dynamics, and habitat characteristics across western United States (i.e., California, Oregon, Washington, Nevada, Wyoming, and Arizona). Wetlands Ecology Division databases include data on waterbird numbers and wetland habitat characteristics in San Francisco Bay. They are used to quantify bird populations, model waterbird/habitat relationships, and develop predictive spatial models for restoration and other habitat changes.

This is the first time that PRBO's spatial data have been comprehensively documented on a cross-division basis, and having done so provides considerable momentum for the maintenance of current metadata, the creation of metadata as new data are gathered every

year, and our ability to share our spatially explicit data with our funders, scientific and conservation collaborators, and the general public. Project leaders and other PRBO staff have been informed about FGDC metadata standards, and know that all spatial data distributed must be accompanied by FGDC-compliant metadata. We have drafted a PRBO data distribution policy (attached), which has been distributed to staff, and will be posted on our metadata portal after review by the management committee.

In the future we hope to facilitate the process of metadata creation across the organization so that anyone who collects and maintains spatial data (i.e. almost any project leader within the organization) can easily create, edit, and maintain their own biological metadata. This would require us to create a tool that people who are not necessarily literate in GIS and / or XML programming can use for that purpose. This tool would be analogous to the NPS metadata extension; it would piggyback on ESRI's ArcCatalog metadata tool but also have an editing interface for the biological components of the metadata that is easier to use than the NPS extension.

Measurable Project Results:

- *The number of individuals capable of creating metadata:*
 - We currently have three staff who are capable of creating metadata: two GIS specialists (one from the Terrestrial Ecology Division and one from the Wetlands Ecology Division) and one landscape ecologist.
- *The number of metadata files or datasets documented:*
 - We have documented four major Terrestrial Ecology Divisions datasets (point count, area search, band, and nest plot), and six wetlands datasets (saltpond area surveys, tidal marsh area surveys, tidal marsh aerial photos, tidal marsh channels, tidal marsh ponds/pannes, and Pacific Flyway area surveys).

Describe metadata service

- Indicate how metadata is served or posted:
 - Our metadata is currently posted on PRBO's NBII-registered metadata website (part of the California node):
<http://cain.nbii.gov/prbo/metadata/index.php?module=datasets>.
 - We have also contacted Elizabeth Martin (NBII) for linking our metadata to the NBII bird node.
- Indicate how many metadata entries were created:
 - We created a total of ten metadata entries on the link listed above.
- Indicate if you need assistance in providing for metadata service
 - None at this point.

Feedback on Cooperative Agreements Program:

- What are the program strengths and weaknesses?

- Unfortunately we were unable to attend the live training, so we can't provide adequate feedback on the training process. The metadata resources provided by the program were of great value to us, though it was sometimes difficult to navigate through the vast amounts of available information.
- Where does the program make a difference?
 - This program enabled us to receive the proper training and thus gain considerable momentum in our ability to create and maintain metadata throughout the organization. Due to the lack of other funding for metadata creation it would otherwise have been very difficult for us to be able to dedicate adequate time to this task.
- Was the assistance you received sufficient or effective?
 - The assistance was sufficient to receive training in metadata creation and create metadata for our existing datasets, but the funding and time budget were not sufficient for us to create custom tools that would enable staff who have no GIS expertise to create their own metadata.
- What would you recommend doing differently?
 - Provide training in ArcGIS or similar programming so that people can create their own metadata tools and / or extensions.
- Are there factors that are missing or need to consider that were missed?
- Are there program management concerns that need to be addressed? Time frame?
- If you were to do this again, what would you do differently?



PRBO Data Sharing Policy November 2003

It is PRBO policy that data should be shared whenever possible while considering the tradeoffs between widespread distribution of data sets, the need to assure data quality and validity, and the privacy rights of land owners, as applicable. The general rule is that data should be released as soon as possible, along with documentation that can be used to judge data quality and potential usefulness.

PRBO believes that the intellectual investment and time committed to the collection of a data set entitles the investigator to the fundamental benefits of the work. Therefore, publication of descriptive or interpretive results derived immediately and directly from the data is the privilege and responsibility of the investigators who collect the data. This is one of the primary motivations for scientific data collection, and the scientific method is supported by this principle.

PRBO encourages its scientists to anticipate that most data collected will eventually become part of the public domain, and recommends that procedures be employed to facilitate this sharing, such as concurrent metadata generation, data verification, data entry, and data backup.

Data for which there are no associated metadata, or which have not been verified, should not be shared. Data that have not been used for peer-reviewed publication should be shared with caution. Data used in publications should be ready for public distribution at the same time as the actual publication. Data relevant to public policy should be shared as quickly and widely as possible. All PRBO data that have been shared should be maintained in a fashion that facilitates its continued public accessibility. This includes the need for data versioning, informing users when there have been important changes to public datasets. Data collected on private lands may be subject to more extensive sharing restrictions, depending on agreements between PRBO and landowners.

PRBO maintains numerous datasets, including several with hundreds of thousands of records covering 20 or more years and involving multiple investigators. PRBO recognizes that these datasets can be used for many purposes, either on their own or through combination with other data. Any researcher making substantial use of a dataset should communicate with the investigators who acquired the data, or those to whom responsibility for the data has been entrusted, prior to publication. In most cases these researchers should anticipate that the data collectors or current managers would

be co-authors of published results. PRBO will encourage and facilitate the ethical and courteous use of data within its archives.

PRBO will maintain a list of all data access and will notify those who access the data that they may be the intellectual property of the collecting scientists. PRBO requests that it be acknowledged in any publication, report, or presentation using data that have been collected by PRBO scientists.

The requesting party(s) and the appropriate PRBO Division Director should sign a data sharing agreement, and a copy should be filed with the Division Director. Data may be posted on the web with approval in advance from the collecting scientists and the Division Director. Such public posting of data must be permissible under relevant contracts.

Any costs that PRBO incurs in the process of providing or maintaining public access to its datasets should be reimbursed. These costs should be considered at the project proposal stage for new data, and at the data request stage for legacy data. Costs include hourly wages for persons responsible for data maintenance and retrieval, as well as indirect costs for maintenance of computer equipment and backup systems.

Developed by Grant Ballard, revised per feedback at PRBO's 2003 Science Retreat and with additional revisions and approval by the PRBO Management Team, November 2003.

PRBO reserves the right to modify, revise or supplement this policy, as it deems appropriate.