



SONOMA ECOLOGY CENTER
GIS & INFORMATION SERVICES

**Metadata Training and Data Coordination
Agreement 05HQAG0142
Final Report**

By

Sonoma Ecology Center

P.O. Box 1486

Eldridge, CA 95431

www.sonomaecologycenter.org

Deanne DiPietro, GIS/IS Program Manager and
Project Leader

707-996-0712 ext. 114

deanne@sonomaecologycenter.org

Prepared for

NSDI Grants Officer

U.S. Geological Survey

NSDI Cooperative Agreements Program

Metadata Training & Outreach Project

FINAL REPORT

Project Narrative

[Summarize the project activities. Include its accomplishments, successes strengths and weaknesses, further challenges, and collaboration activities as appropriate. What are the organizational practices that have developed to support metadata creation and maintenance for the future?]

This project builds upon the establishment of the North Bay Initiative for an Environmental Information Commons, funded by a previous NSDI CAP grant (2004; originally called the San Pablo Bay Watershed Data Consortium).

The applicants and these partners recognize the need to regionalize geographic information management and data development. Multiple projects are creating and using similar types of information, and the sharing and integration of the information is known to be lacking. This lack of coordination hampers efforts in water conservation, water supply, ecological restoration of fisheries and wetlands, sustainable water use, water quality management, flood protection, environmental education, and land use planning. In response to this need, The North Bay Initiative for an Environmental Information Commons had its first official meeting on May 12, 2005 wherein it established an online forum for discussion of issues, a Mission Statement, and draft set of Specifications for a Regional Environmental Digital Library Network (see <http://northbaycommons.net>).

Interest in the San Francisco North Bay region regarding the effective exchange of environmental information is very high, as was reflected in the attendance of the North Bay Initiative's first meeting, with representatives of 24 organizations that create, maintain, and use spatial and other types of environmental data. These organizations have indicated that they intend to collaborate toward improved data discovery, sharing, and access, and one of the next steps is seen as the creation of an open network of standard data catalogs that will be integrated with the CERES Environmental Information Catalog, the NSDI Node for California.

The recent collaborative activities of the Sonoma Ecology Center, the San Francisco office of the US Army Corps of Engineers, the Sonoma County Water Agency, the San Francisco Regional Water Quality Control Board, and the San Francisco Estuary Project and the many others who have joined the North Bay Initiative for an Environmental Information Commons represent the most current work toward the creation of a collaborative environmental data management coalition in our region. The current project strengthened the partnership's ability to make progress in its mission by teaching these organizations to create catalogs of their data holdings that may be shared within the region, at the state level via the CERES Catalog, and at the national level through the CERES NSDI Node.

The current project provided the necessary outreach and education to promote a regional inventory of spatial data and environmental information. To accomplish this, the CERES Environmental Information Catalog is being used for the creation and posting of FGDC-compliant metadata to the state's NSDI Node at <http://ceres.ca.gov/catalog>. This project was coupled with the development of new techniques to promote a live network, and now most of the new catalogs we created are held and

FINAL REPORT

served from the owner organizations and harvested into the collections at CERES. These collections are updated from the source of the information, as the organizations make changes directly to their own in-house libraries. All of these live-updated collections will be maintained at CERES for inclusion in the NSDI.

The software and techniques for this digital library network are under development by a cooperative effort called the California Digital Library Network (CalEDLN), and their implementation is being undertaken in several publicly-funded projects within the North Bay and San Francisco Bay regions. One, a pilot library network specifically meant to serve the North San Francisco Bay Region, has been funded by the Sonoma County Water Agency and will link regional datastores with those at CERES and ICE. Another, funded by the Army Corps of Engineers, developed methods for in-house data management that will result in the creation of public libraries by small as well as larger organizations.

Products and activities of this project:

1. Two workshops were held to develop software and to teach members of the North Bay Initiative to create libraries of FGDC-compliant metadata and share them using the technology under development by the CalEDLN effort.
2. Ongoing support is made available through the North Bay Initiative's communication portal, by phone, and in person by project staff.
3. A website has been developed to support the ongoing regional library network with standards and instructions for participation (<http://northbaycommons.net>).

FINAL REPORT

Metadata Training and Outreach Assistance

- **[List organizations and number of individuals receiving metadata training and outreach assistance as appropriate**
- **At what level of proficiency are the trainees or training**
- **Indicate the number and character of workshops conducted as appropriate]**

The first metadata training workshop took place over several days in July, 2006 and was conducted as a programming Sprint intended to quickly produce a large amount of metadata in the FGDC standard and share it on the network. Ten programmers and about 8 others attended. Organizations represented included:

San Francisco Bay Joint Venture Project Database (SFBJV)

Watershed Portals,

Klamath Resource Information System,

Sonoma Ecology Center,

North Bay Watershed Association,

MIG, Inc. (Provider of several important regional information portals)

Natural Resource Projects Inventory,

Information Center for the Environment,

CERES (California Resources Agency)

Ten professional programmers representing north bay and state-wide data sources attended to develop techniques for porting their organization's existing environmental project description databases into the CalEDLN project standard XML format, a special standard for project descriptions that is based on the FGDC metadata standard. Techniques included various scripts (Python, etc.) for export from the numerous native formats, as well as a relational database application for data management, import, and export. A variety of thesauri were developed from the source data and posted on the North Bay Information Commons portal (<http://northbaycommons.net/standards/thesauri>). Seven project catalogs of legacy project data are now offered online as XML files conforming to the CalEDLN project content standard (<http://northbaycommons.net/northbay-data-catalog>). Metadata libraries represented and number of records posted include

- North Bay Watershed Association (284)
- San Francisco Bay Joint Venture Project Database (291)
- Natural Resource Project Inventory (5,605)
- North Coast Integrated Regional Water Management Plan Projects (239)
- Sonoma Ecology Center, Marin Audubon Society, and Urban Creeks Council (60)
- Central Valley Watersheds Library (148)

FINAL REPORT

Other products and outcomes from the first workshop, the programming Sprint, included:

- Technology development:
 - Important refinements were made to the standard specification
 - NRPI developed a method to query the XML data using SQLserver, and to parse the data into a database.
 - KRIS and SEC collaborated on a relational database application useful for metadata management and for aggregating data.
 - ICE made strides in thesaurus networking techniques and posted several core thesauri (eg., SFBJV's Project Activities)
 - CERES made progress on search utilities
 - Watershed Portals created utilities in Plone
 - All innovations were posted to a common site for access by everyone as open-source products
- Collaboration:
 - All the major project data managers for the North Bay were present, came up to speed on the CalEDLN and Northbay Commons effort, and got to know one another
 - Several cooperative sub-efforts arose that will get us to solutions faster and benefit everyone

The second metadata training workshop was held during the week of August 25 through August 29, 2006. Representatives from 13 organizations with varying levels of metadata experience received training in several methods of producing and sharing metadata. Organizations and number of attendees include

- Audubon Canyon Ranch (4)
- Bay Area Open Space Council (1)
- Bodega Marine Lab, UC Davis (2)
- California Dept. of Fish & Game (1)
- Circuit Rider Productions (1)
- Laguna de Santa Rosa (2)
- Marin County (2)
- North Bay Watershed Association (1)
- San Francisco Bay Conservation and Development Commission (1)
- Sonoma County Agricultural Preservation and Open Space District (2)
- Sonoma County Water Agency (3)
- Sonoma State University (1)
- US Fish and Wildlife Service (2)
- Ray Carlson & Associates (1)
- Interested students (2)

FINAL REPORT

This second workshop, taught by Sonoma Ecology Center staff Deanne DiPietro and Pat Stiefer with help from FGDC Metadata Coordinator Sharon Shin, was designed to help organizations manage their information resources in-house so they can find and use their own data and information and then, if they choose, share it with others.

The workshop included two sessions, each presented twice from August 21st through August 25.

Session 1, "Cataloging and Sharing Information" taught easy techniques to create an in-house library catalog that an organization could use for all its information resources such as projects, maps, reports, monitoring data, photo collections, and books.

Session 2, "Documenting and Sharing Spatial Data", demonstrated how spatial data managers can create basic metadata necessary for organizing, finding, and sharing spatial datasets; how to document the data so others can use it; and suggested methods for organizing the data files on inhouse servers. Sharon Shin, Metadata Coordinator for the Federal Geographic Data Committee, described the FGDC's Geospatial One-Stop and NSDI programs and helped answer questions about using the FGDC Metadata Standard.

Both sessions taught how to publish standard metadata on the Northbay Environmental Information Commons (<http://northbaycommons.net>), where it becomes available for search and discovery by a network of environmental libraries including CERES and the NSDI.

An evaluation form (Appendix A) was distributed to attendees at each session. Thirteen participants returned forms covering a total of 21 sessions. Responses to the first four questions were generally positive, with variations in details and enthusiasm. The fifth question about what should be done differently next time naturally generated the most varied responses. Generally it seems the class thought this was a lot of information to learn in at the sessions (there were several requests for detailed handouts for note-taking and for future reference). There were also several requests to spend more time on various sub-topics covered in the presentations. After the last session, the curriculum outline, powerpoint presentations, and a sample EndNote library with user guide were posted on the North Bay Information Commons portal (<http://northbaycommons.net/training/metadata-training-workshop-materials-1/>).

FINAL REPORT

Status of Metadata Service

- **[Site names where metadata is served; clearinghouse node or Geospatial One-Stop harvestable web folder.**
- **Approximately how many metadata entries have resulted from this project?**
- **Do you need assistance in providing for metadata service to organizations you have assisted?]**

Metadata is served by the owners at their own web-accessible sites (see them all linked at <http://northbaycommons.net/northbay-data-catalog>), where they are harvested by CERES, which is in turn harvested by NSDI. Approximately 6,696 project metadata records have been created and posted at the Northbay Commons as a result of this project. Additional spatial and nonspatial metadata records may have been posted directly to CERES by organizations after the second metadata workshop.

There is a need for assistance in providing continuing education and support services to the organizations that we trained and engaged in these first two metadata workshops. The organizations are now facing the task of delving into their information resources, some for the first time, with the goal of organizing it and cataloging it so it may 1) be found and used again in-house, and 2) shared with other organizations. The general message to us from them is “we can use help!”. The task is big, especially for organizations that have existed for many years and have accumulated a lot of information resources. And in most cases, the work of cataloging has to be done in between all the day-to-day work. Consequently, it goes slowly, and the people trained need help remembering the techniques we taught them and getting the process underway in their own settings. Our sense is strongly that this cannot be the end of the story, or our investment will only be partially returned with successful catalogs that are kept up to date. We should keep offering workshops to increase our audience’s skill and confidence, continue answering questions that come up, and to have the opportunity to update them on new software and techniques released by the CalEDLN project.

FINAL REPORT

Next Steps

- **[Will this project's activities continue in the future**
- **Describe the next phase in your project**
- **Are there issues in metadata management and service**
- **Requirements (more technical assistance, software, other?)**
- **What areas need work?]**

As mentioned, this project was part of the ongoing efforts of the CalEDLN and Northbay Commons programs. We are working now on the Northbay Commons access site to the data for a regional view of the catalogs from our area and the data relevant to our region from the state-wide catalogs in CERES. This “aggregation site” is a library made from distributed libraries that anyone can harvest, subset, and interface for regional or thematic presentations. The intention is that there will be multiple aggregation sites, and regional “Commons” across California.

The Northbay Commons committees continue to meet, plan activities, and write grant proposals to fund our ongoing work. Recently the Steering Committee decided to re-name the regional effort to the San Francisco Bay Area Conservation Commons. This reflect the fact that the group has adopted the Principles of the Conservation Commons (see <http://conservationcommons.org>) and its philosophy about information access, and also the desire to include the entire Bay Area in its scope. We have discussed the creation of a California Conservation Commons and have received permission and encouragement from the Conservation Commons to do so. A new domain, calconservationcommons.net, has been registered and our regional Commons will be the first component of what will hopefully catch on across the state. Our website is currently being re-named and will soon be found at <http://sfbayarea.calconservationcommons.net>.

A proposal has been submitted to the National Science Foundation to fund the technical development needed to support a live, integrated network of metadata that is managed at all our many organizations. Another proposal is being considered by the Sonoma County Water Agency to support the regional effort of the SF Bay Area Conservation Commons, primarily the coordination of the committee meetings and the ongoing work to build the regional data network and aggregation library. CERES and UC Davis Information Center for the Environment have committed staff time for another year to support the CalEDLN effort, which develops the technology we implement for building the regional network. We anticipate that our efforts will grow to include participation by all environmental organizations in our region, and so we will need to continue to teach metadata management workshops and improve the tools for metadata management in the context of the live network. One metadata workshop is planned for 2007.

FINAL REPORT

Feedback on Cooperative Agreements Program

- **[What are the program strengths and weaknesses?**

Strengths: Staff are willing to help, as evidenced for us when Sharon Shin traveled to our metadata workshop on her own budget to help us teach metadata techniques and tell people about the NSDI. Contracting is easy and gives the recipient a lot of flexibility, and the money can be requested ahead of our incurring costs, which is important for small non-profits with cash-flow challenges (don't change that!).

Weaknesses: Grant categories are not consistent from year to year, making it hard to provide continuity of support. Example: now that we have completed the "Getting People Together" and the "Metadata Training" projects, we were hoping to set up a map server under the category that used to support that, but it's not there this year.

- **Where does the program make a difference?**

The two grants we received helped us launch our effort. It is hard to get seed money, and the "Getting People Together" grant was instrumental in getting us started. The "Metadata Training" grant was just what we needed to follow through and train people, after getting them hooked on the idea of data sharing in the first round.

Another way these grants have made a difference for us is by strengthening our effort by associating it with the NSDI and FGDC. We are very happy to mention that our work is supported by the NSDI, as it lends credibility.

- **Was the assistance you received sufficient or effective?**

Yes; the grant was enough to conduct two workshops, which was enough to do in one year. We feel like we could do it again now much more easily and perhaps we'll repeat it more often in coming years. The help from Sharon was also great- again, this helped us not only because she knows a lot about the FGDC and metadata, but also her presence at our workshop lent it that much more credibility.

- **What would you recommend doing differently?**

Nothing really, just doing it again.

- **Are there factors that are missing or need to consider that were missed?**

It would be great to keep in touch with the program somehow, even when we aren't funded. I will try to participate in the email discussions.

- **Are there program management concerns that need to be addressed? Time frame?**

I like it the way it is!

- **If you were to do this again, what would you do differently?**

FINAL REPORT

Nothing; it worked out great. Thank you for the opportunity, and I hope we continue to work together in some way.

FINAL REPORT

APPENDIX A

Evaluation Form
for the
Environmental Information Cataloging Workshops
August 21-25, 2006

Please help us improve our efforts to promote information management and sharing in our community by answering the following questions:

1. What is your overall rating for the training workshop?
(On a scale of 1-5: 1=Poor, 3=Pretty Good, 5=Excellent)

Part 1: Cataloging and Sharing Information _____

Part 2: Documenting and Sharing Spatial Data _____

2. Was the information useful to you?

3. Was the material presented clearly?

4. Do you think you'll apply what you learned in your work?

5. Is there anything you think we should do differently next time?

Thank you for participating and best of luck in your information management efforts!
Sonoma Ecology Center