Data Mining Efforts
Central Alaska Network

19 November 2002

Previous Work

- NPSpecies and NRBib populated for all AK parks by AKNHP (see report: Final Report - Compilation of Existing Species Data in Alaska's National Parks-March 2002) – Summer 2001. The AKNHP entered or edited 457 records in NatureBib and 4,358 records in NPSpecies for CAKN parks. The entries/edits in both databases were categorized as plant, fish, bird, mammal or amphibian.

- Dataset Catalog and NatureBib population for YUCH – Summer 2001. Resource reports and selected publications along with associated datasets from 1975 to 2000 were compiled for YUCH. Two hundred twenty nine reports and 16 datasets were cataloged for YUCH. Park-generated reports were copied and bound into a multi-volume set for archival in the YUGA office. This data mining effort included visits to Eagle and YUGA office library to search old filing cabinets and interview personnel regarding datasets and reports.

- Existing effort to catalog WRST reports and datasets into NatureBib and Dataset Catalog. An on-going process to catalog information for WRST includes entering records into the on-line NatureBib and into a local version of Dataset Catalog. Park personnel are carefully recording each report and dataset generated by park projects using an initial list of datasets drafted by key personnel with significant institutional knowledge.

- WRST Procite database. Created in 1992 but has not been maintained. Information from the park library and literature in permanent staff files was entered. The database was sent to Ft. Collins for first edition of NRBib, therefore the records are in the current version of NatureBib. The database includes park reports as well as published literature for various specialists in the park. Some WRST staff created their own Procite databases since the park database used keywords not suited to their focused research. Poor computer networking and application access also contributed to development of multiple Procite databases in the park. Several filing cabinet drawers have been targeted for entry into the Procite database.

- Earlier efforts to populate NatureBib and Dataset Catalog by WASO – Mid 1990's:

NPSpecies

WASO is currently data mining for NPSpecies vouchers by contacting museums, herbaria, and universities across the U.S. (and some outside the U.S.). Early WASO data mining for NPSpecies consisted mainly of data from the Biota of North America Program (BONAP) county species database and the BONAP floristic data. No records for CAKN parks were found in the BONAP county database. The BONAP floristic database appears to have yielded about 900 records for DENA, about 850 records for WRST, and about 450 records for YUCH.

NatureBib

The initial effort to populate NatureBib began in 1994. WASO sent NPS and contract personnel to each of the I&M parks to conduct citation searches of park libraries and offices and train local staff to continue the effort. Approximately 52,000 records from a previously populated NPS bibliographic database for geologic citations, GRBib, were added to NatureBib on October 16, 2002.

Dataset Catalog

No program has been initiated by WASO to populate Dataset Catalog.

Central Alaska Network Strategy

Stage I

- Obtain latest desktop version of NatureBib, Dataset Catalog and NPSpecies each populated with records for the CAKN.

- Obtain list of users authorized to login to the on-line NatureBib and NPSpecies (Dataset Catalog is not web-enabled, although the new NR-GIS Metadata browser is nearly complete at http://www3.nature.nps.gov/im/metadata/). The list can be obtained from Simon Kingston who maintains login rights for both databases. The list helps illuminate who in the network may be entering data and therefore provides greater control over the data entry process.

- Conduct a network-wide white literature search. This effort is carried out by one staff member assigned to search on-line databases for appropriate information. White literature searches focus on biology since many physical science records, primarily geology, were added during earlier efforts to populate NatureBib. Also, it is clear that an understanding of biological work conducted in the CAKN is critical to development of ecological conceptual models--an important step to designing a network-wide monitoring program. Results from on-line searches are compiled into a text file for later entry into a local version of NatureBib.

- Grey literature for WRST. The existing process for cataloging WRST reports and datasets (described above) is allowed to carry forward without change.

- Evaluate for completeness the previously done cross-check of the DENA Procite database with that park's library. Complete the cross-check if necessary and expand it to check against current NatureBib records.

- Examine the DENA Research Permit Database for information leading to reports and datasets appropriate for inclusion in NatureBib or Dataset Catalog.

Stage II

- Literature search for reports and pertinent documents generated by federal, state and local agencies and organizations with lands adjacent to CAKN parks and preserves. Possible non-NPS sources include Pitman Robinson Reports, BLM, and DNR.

- Organize collected information in NatureBib and Dataset Catalog and submit to Ft. Collins for up load into master databases.

- Map distribution of current and historical monitoring in support of I&M Phase I report.

- Identify cataloged datasets that merit enhancement for purposes of development of the vital signs monitoring program.

Changes/Decision Points

- A quick check of grey literature compiled in 2001 for YUCH revealed that not all citations had been uploaded into NatureBib. The white literature search was suspended and a thorough review of the compiled YUCH grey literature initiated. Citations not found in NatureBib were added to the on-line version.

- Citing database access problems, staff at DENA assigned to cross-checking the Procite database with the park library and examining the Research Permit Database (RPD) instead focused efforts on searching the research permit files for data and publications relating to invertebrate work in DENA.