Meeting Notes: Northwest Environmental Data-network Steering Committee and PNAMP Data Management Work Group Meeting

Time: 8/2/2006 from 9:00 to 2:30

Location: Conference Room: Northwest Power and Conservation Council
851 S.W. Sixth Avenue, Suite 1100
Portland, Oregon 97204
503-222-5161 or 800-452-5161

Participants: Bruce Schmidt (StreamNet), Tom Pansky (BPA), Peter Paquet (NPCC), Tom O'Neil (NWHI), Phil Roger (CRITFC), Jen Bayer (USGS), Michael Newsome (USBR), Van Hare (StreamNet), Ben Zelinsky (BPA), Bart Butterfield (IDF&G), Sam Sharr (IDF&G), John Piccininni (BPA), Cy Smith (OGIC), Joy Paulus (WA IAC), Jen Bayer (USGS) and Stewart Toshach (NED). On Phone: David Tetta (EPA), Cedric Cooney (StreamNet-ODFW), Russell Scranton (NMFS-NWRO) Nancy Tubbs (USGS), Jim Gieselman (BPA). Apologies: Bobbi Riggers (OWEB), Jill Leary (LCREP), Kathy Kellon (SOS) and Chris Van Holmes (DART).

A. NED AGENDA ITEMS

1) Introductions (see above).

2) Demo of NED Portal. Tom Pansky summarized the present situation. The installation of the ESRI Toolbox version 3 is now completed with testing being completed on the development server. Transfer to the production server with external access for users is expected within the next few weeks. BPA is also undertaking some rebuild of its web servers and updating ARC IMS to provide faster service. A demonstration for Dr. Karier, the Federal Caucus and other interested clients will be developed. A data delivery extension has also been added.

Questions included:

How would the system differentiate between different versions of documents? Response: the metadata records would provide version information about each document.

Michael wanted to know if the Portal to be able to identify data that met certain data quality or assurance standards, or was collected using certain approved templates. Response: All data meeting certain standards could be placed within a particular portal folder. Portal staff would not be responsible for determining what standard the data met - it would be the responsibility of the data providers to define the quality of the data they publish.

Van wanted to know more about the ability to publish data? Tom showed how the system allowed users to log on as a 'Publisher'. Van noted that StreamNet was ready to participate and provide data and that he would provide a data steward role for the StreamNet data.

Jen asked how the system could support particular users, for example to provide data for the CBFWA State of the Resource Report. Answer: The system would provide access to metadata records for Portal published data. The challenge is to identify the data that CBFWA wants to access, ensure that it is published to the Portal, and that web or other access to this data is available – via links or other mechanisms. StreamNet expects to have discussions with CBFWA next week to discuss what data StreamNet can provide to CBFWA as summarized views of the StreamNet data.

Jen wanted to know what the plan is for Data Stewardship and Channel Management. Response: NED has completed a document *Network Portal Channels and Data Stewards Roles and Responsibilities* that identifies what is involved in managing a portal and the tasks that are needed. The first part of the plan is therefore complete- to define the task. The second part of the task is to identify how to provide this service, either through in-kind contributions of staff time or with dedicated staffing. Decisions about how many channel stewards are needed and how they will be provided have not yet been made.

Cy noted that the Channel Steward role was a difficult task if it is to be well done and thought that there was a better chance of doing it well at a regional level than at the National Level.

Russell asked if the *NED Portal Data Sharing Agreement* was intended to provide a basis for intra or inter-entity data sharing. For example some agencies have difficulty getting access to needed data. Response: The *NED Portal Data Sharing Agreement* was developed only to serve the publication of data to the Portal or the use of data from the Portal. It does not obligate any entity to share data or make data available. It is understood that there may be a need for separate data sharing agreements to support needed data sharing and that agreements could be developed depending on the specific needs of the entities- what do they want?

Cy noted that the *Agreement* currently provided only for the provision of data in view only while some users might want the ability to download data. Response: The Portal software has the ability to provide download capacity and the *Agreement* would be amended to reflect this.

Joy thanked the BPA staff for all the work they have done to get the Portal running and on the appearance of the Portal site.

3) Update on Columbia River Data Center Proposal. (Peter Paquet)

Peter summarized the 12 or so responses that the Council has received on the proposal. He noted that nearly all the responses supported the principles or goals of the proposal but differed in how they should be achieved. Many of the responses noted that existing efforts have the potential to meet most of the objectives if there was sufficient coordination and the development of a plan to assign responsibility for the needed actions. There was interest in addressing these issues in time to take proactive steps to improve data management for the 2007-2009 funding cycle. There is also an important issue of scope. It is not just a Fish and Wildlife program issue. A program solution is important, but not in isolation from other partners since there is much cross dependency on the data. Peter said that a summary of the findings would be presented to the

Council on August 15th as an information item at the Council August meeting. Dr Karier is also looking at the responses.

Joy emphasized the importance of focusing this effort on the development of products that moved data management forward.

4) Federal Caucus Data Needs. (Michael Newsom)

Michael Newsom outlined a presentation that he had given recently to the Federal Caucus as a part of consideration of the Data Center Proposal. He summarized the development of the John Day data dictionary (originating from BPA funded work by Tracy Hillman in the Upper Columbia) and its deployment into the Integrated Status and Effectiveness Monitoring Program (a cooperative effort by the NWFSC and the BORec that is funded by the BPA). The NWFSC web site describes the project as follows: "This project seeks to design, implement, test, and evaluate Status and Trends and Effectiveness Monitoring programs for salmon and steelhead populations and habitat. ISEMP pilot projects are being developed in the John Day, Salmon and Wenatchee basins of the Interior Columbia. In addition to monitoring salmon and steelhead populations and habitat within pilot basins, ISEMP is designed to test the robustness of monitoring protocols, indicator metrics, and sampling designs currently used in monitoring programs. ISEMP plans to develop tools to facilitate effective data analysis, management and communication."

Michael has recommended more widespread use of the ISEMP database and template driven collection methodology within the region. He drew a distinction between entities that develop data base capability such as the NWFSC and those, such as StreamNet that implement them, and thought this broad distinction was a useful model for moving forward. He plans to bring this proposal to Dr Karier and to include it, if possible into collaboration efforts under the remanded BIOP. Peter noted that this sounded like a warehouse solution which has not been supported by either the ISAB or the SAIC reports to the region as long-term solutions to data integration, and that the trend was towards distributed approaches. In addition it did not provide for collection of a wide diversity of non-salmonid data.

Joy offered that, on the development side, the development of tools to manage standards and protocols, such as the pilot protocol manager was important since it had the potential ability to support multiple data collection issues and provide templates for collection of data that ultimately supported more effective sharing.

Cedric wanted to know what sort of data the pilot was focused on, was it just Salmonid data or was it a broader range? Response: It was going to be broader and the template was now being revised to capture more data. There was a June 07 deadline for completion of the template. Cedric offered that there was a problem with the methodology of having only a restrictive template. What happens to data that does not fit the template, we lose it.

There was discussion on QA/QC, with recognition that it was a substantive problem that affected the end-to-end data system, from collection design to collection to reporting to database management and eventually to analysis. More work was needed in this area and we have only

touched the tip of the iceberg. There was a need to take the QA/QC issues to a broader audience: CSMEP, CBFWA, PNAMP and others.

5) Use of NED Data *Network Portal Channels and Data Stewards Roles and Responsibilities* and *NED Portal Data Sharing Agreement* (Stewart Toshach). These documents were discussed in relation to progress on the Portal – see item 1 above. Any new comments should be forwarded to Stewart.

6) Update on International Columbia River Basin Center of Knowledge proposal. (Peter Paquet) Peter reported that a meeting had been held with representatives from the U.S. and Canada to discuss the data and information sharing among various U.S. and Canadian entities. There was consensus that all information regarding the Columbia River in its international dimensions should be considered for sharing via the proposed Center of Knowledge website, through hypertext links to other websites and through a separate data portal or portals that would link to specific data. The group discussed the goals and objectives proposed in the draft issue paper developed by the Columbia Basin Trust and the Power Council and posted on their websites for public comment. There was consensus that the goals should emphasize that the website and portal(s) would be managed adaptively so that changes could be made over time as needed, and that "governance" should be added to the types of information that would be included. There was also consensus that name "International Columbia Basin Center of Knowledge" isn't quite right and that the group will work to come up with a better title.

7) Report on Subbasin Planning Data Management paper. (Phil Roger)

Phil provided a power point summary of the lessons learnt from the effort to pull together the data that was collected as a part of the Subbasin planning efforts. A particular product is the addition of GIS capability to the EDT data sets. The challenge was to integrate some 50 disparate data sets - a daunting task - for future use and broader landscape analysis. Phil has been asked to provide a briefing to the Council on this topic in September and wanted to know if the presentation should be broadened to address more generic data management issues in the region and to begin to respond to how some of these issues can be addressed through pilot efforts.

Comments:

How can data integration be achieved, for example for the State of the Resource report?

Phil's example, while important, in not the first - Roberto Morganti expressed similar views of the difficulty of working with disparate data for Northwest Forest Plan data reporting.

While it is great to have the data available, it will still not be possible to analyze across the subbasins because the data collection integrity does not exist. Different collection methods were used and most data was provided without metadata, so the level of confidence in rolled up data will be low.

Despite limitations it is still a major step forward.

The ISEMP could be put forward as a pilot example to work towards

A policy 'pilot' could also be promoted, for example to focus on data quality improvement, how to include standards and how to document data quality.

Support for the portal can be viewed as a pilot effort.

Peter noted that the ISAB review of the Council Columbia Basin Data Center proposal emphasized the need to deal with data access and data quality control, identifying "an urgent need for coordinated entities responsible for providing access to and quality control of the diverse range of environmental data accumulating in the northwest region."

Stewart noted that in the event that the presentation is broadened as Phil is suggesting, there are a number of resources to work from: The 2000 ISRP report, the 2003 SAIC report, the 2005 NED Workshop White Papers, the Columbia Basin Data Center Proposal and responses.

It was agreed to try to broaden the presentation. Phil asked for feedback on the latest version of his ppt. presentation (posted on the NED web site) with ideas and suggestions for pilot efforts by August 14th. It was agreed that Stewart will set up a conference call on Monday August 14th at 1:00 pm to discuss the comments.

Note teleconfernence meeting has been set up from 1:00 through 3:00 on August 14th. Call in details will be sent to you in a separate e-mail.

The subsequent revised document would then be discussed again at the NED meeting on September 6th ahead of a planned presentation to the NPCC in September. The idea is to have a joint presentation with representation from different NED participants.

8) Other products to complete: a) Check List for data collectors. b) Data dictionary semantics. (Stewart Toshach)

Stewart identified two NED products that can be completed in the near term.

The check list for data collectors was written as guidance to data collectors working on the PNAMP protocol testing project. It has a potential wider audience and can be completed for this purpose.

The data dictionary product was drafted last year but has not yet been completed.

The Steering Committee supported the completion of both of these products.

9) Update on Salmonid Data Tracking Project. (Stewart Toshach). See item below.

10) Meeting with ISAB (Peter Paquet) . Information item. Peter, Stewart and John Stein met on June 30th with the ISAB, as a part of their deliberations on the Council Columbia Basin Data Center proposal, to brief them on the NED effort.

11) Short Updates from Workgroup Leads (if time). Reports from the work group leads will be taken up at the NED September meeting (September 6th)

B. PNAMP DATA MANAGEMENT AGENDA ITEMS

1) PNAMP Inventory Task**.** Update. Bruce Schmidt. Technicians have been employed and the inventory has been distributed for completion by regional monitoring participants. Completion is expected by the end of the 2006 FFY.

2) A Presentation about the Pacific NW Aquatic Monitoring Partnership and NED to the Organization of Fish and Wildlife Information Managers (OFWIM) conference in Minneapolis, MN has been requested. The conference is being held from October 16-19. Stewart has been invited to make a joint presentation to the OFWIM meeting for PNAMP and NED. Jen and Stewart would prepare a presentation. The NED Steering Committee supported this action. John Piccininni will follow up on the availability of funding to support anticipated expenses.

3) PNAMP Effectiveness Monitoring Meeting. The PNAMP effectiveness Monitoring Committee is reviewing a "Data Management Design for regional Project Tracking to Support Monitoring and Evaluation" and has asked for input from NED. Stewart is a coauthor. Tom O'Neil volunteered to review the document for NED. Copies of the document will be circulated when it is available.

4) Training opportunities for metadata. Jen has had a discussion with Vivian B Hutchison, Metadata Coordinator, National Biological Information Infrastructure (NBII) about metadata training opportunities. There are currently opportunities to join an upcoming session session and also to have more custom training for groups with special interests, in late winter. Jen will follow up on more details of what is available and let us know.

C. AFTERNOON SESSION

Peter Paquet, Joy Paulus, John Piccininni, Jen Bayer, Tom O'Neil, Bruce Schmidt, Phil Roger and TomPansky met from 1:00 to 2:00 pm to discuss the planned work for the *Study of Salmonid Status and Trend Data Collection Work Process.*

Stewart outlined that the effort was intended to document the current methods being used in the region to collect and roll up Salmonid data. A collaborative approach is being sought with Fish and Wildlife entities or agencies to help define the needed products and provide information for the effort. The main current challenges are to 1) manage the scope of the effort and 2) develop a list of essential contacts.

Discussion and Comments:

Define the scope first because that will define the essential contacts.

Make it clear that the information when collected will be available to contributors and the public– the current proposal is not clear.

Arrange to post the findings to the NED Portal
Develop the statement of work for CBFWA context because that is who will be completing the work.

Develop a list of benefits for participants. For example 1) it is necessary in any event to document the data collection process to meet internal agency needs, and 2) the results may be used for agencies to satisfy metadata reporting. Emphasize that it is important to focus on metadata for tabular data since GIS specialists are further ahead in reporting spatial metadata than most groups who work with tabular data.

Another approach to the scaling challenge would be to focus on a smaller geographic area and complete the effort as a pilot.

Work on developing a template that defines the minimum amount of data that needs to be collected to define the collection methodologies.

Look at the work completed for the CSMEP project which focused on data rich and data poor sub basins as an example of what can (and cannot) be achieved.

Sam Sharr and Bart Butterfield commented that the product is likely to be of interest to the IDF&G and they are interested in participating.

Cedric Cooney suggested that some of the agencies are feeling "over-surveyed" so it would be important to clearly identify what they will get out of their participation. Stewart will follow up with Cedric to identify best contacts and ways to work ODF&W.

Stewart will follow up with other Fish and Wildlife agency and Tribal staff to understand what is the best way of working collaboratively, will revise the draft SOW to reflect the comments offered above and will take a first effort to more narrowly define the scope.

NEXT MEETINGS

September 6, 2006. NED Steering Committee and PNAMP Data Work Group. Conference room of NPCC 9:00 through 4:30

October 4, 2006. NED Steering Committee and PNAMP Data Work Group. Conference room of NPCC 9:00 through 4:30

Phone call in details and agenda will be distributed later.