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

Time: 10/4/2006 from 9:00 to 4:00

Location: Columbia River Inter-Tribal Fisheries Commission, 729 NE Oregon St, Ste 200, Portland Oregon

NEXT MEETING: November 1st at NPCC Portland 9:00 to 4:00.

NED AGENDA ITEMS for 10/4/2005

1) Introductions:

John Piccininni (BPA), Jim Geiselman (BPA), Bruce Schmidt (StreamNet), Tom Pansky (BPA), Peter Paquet (NPCC), Tom O'Neil (NWHI), Phil Roger (CRITFC), Roberto Morganti (USFS), Russell Scranton (NOAAF-NWRO), April Wollstencroft (NOAAF-NWRO), Joy Paulus (WA IAC), Bobbi Riggers (OWEB), and Stewart Toshach (NED). On Phone: Jen Bayer (USGS), David Tetta (EPA), Cedric Cooney (StreamNet/ODF&W).

Afternoon Session: Karl Jacobs (CDF&G), Brian Hale (CDF&G), John Piccininni (BPA), Jim Bruce Schmidt (StreamNet), Tom Pansky (BPA), Peter Paquet (NPCC), Tom O'Neil (NWHI), Phil Roger (CRITFC), Roberto Morganti (USFS), Hilary Forest (CRITFC), Denise Kelsey (CRITFC) Russell Scranton (NOAAF-NWRO), Joy Paulus (WA IAC), Bobbi Riggers, (OWEB), and Stewart Toshach (NED). On Phone: David Tetta (EPA), Cedric Cooney (StreamNet/ODF&W).

2) Short Updates from Workgroup Leads.

Work group structure is expected to change in 2007 to support prototype/pilot efforts currently under development. Until then work is ongoing in the following work groups:

- Riparian, Upland and Aquatic Habitat Work Group (Tom O'Neil). Tom is waiting for comments on the Inventory database before moving forward with further work. The inventory has been distributed to the PNAMP members for comment. A copy can also be obtained from Tom at habitat@nwhi.org

- Water Quality (David Tetta) The Water Quality Work group is looking to the Steering Committee for guidance on next steps. The outstanding issue is for the group to consider the extent that the National Water Quality Attributes should be included. An important task for the SC as a whole is to work to publish the meta-data from the Water Quality data systems to the NED Portal.

- Subbasin Data (Phil Roger and Peter Paquet)

Some new data is still trickling in for the subbasin collection effort. The next 'round' of sub-basin data collection is expected to being in 2008 and it will be necessary to plan for that if the problems with the previous data collection are to be avoided. This would mean deciding what level of prescriptive data management practices, standards and etc would be needed to allow the data to be systematically collected at a regional level without a follow-on effort. There is interest in modifying the EDT model for data collection to review the adequacy variables that are currently used. There is also interest in looking at whether the EDT and AHA data models can be combined.

- Technology for Data Discovery and Sharing (Tom Pansky) The Portal is now running outside of the BPA firewall and data delivery software has been installed by ESRI. This software allows Portal users to extract data from the Portal rather than just view it. The major current effort is to finalize a pilot index. The core index is based on International Standards Organization but will be customized to meet regional needs using an index first developed by the State of Oregon.

- Project Description & Performance/Temporal & Spatial (Joy Paulus) – Comments have been received based on PNAMP distribution of the document. All comments received have been constructive and supportive of the idea for improved standardization. The most challenging comment came from John Arterburn from the Colville Confederated Tribes. John had pointed out that the proposed time standard mm/dd/yyyy (eg 08/06/2006) (August 6 2006) was inconsistent with other nations who use dd/mm/yyyy (eg 08/06/2006) (8th of June 2006) and offered a compromise dd/mmm/yyyy (eg 06/jun/2006) to remove the ambiguity. Stewart said that he agreed with John that the dd/mmm/yyyy proposal was superior technically but the most common convention in the Pacific Northwest is mm/dd/yyyy and getting a change would be very difficult. In recommending the use of mm/dd/yyyy the work group was aware of other options and of the potential confusion that mm/dd/yyyy provided. In the end the work group had decided that the potential to achieve consistent reporting of mm/dd/yyyy was greater than the potential to achieve consistent reporting using dd/mmm/yyyy and this was why the work group selected it. Stewart agreed to call John to discuss. Joy and Stewart would join the PNAMP SC meeting on 10/11/2006 to discuss with the PNAMP SC.

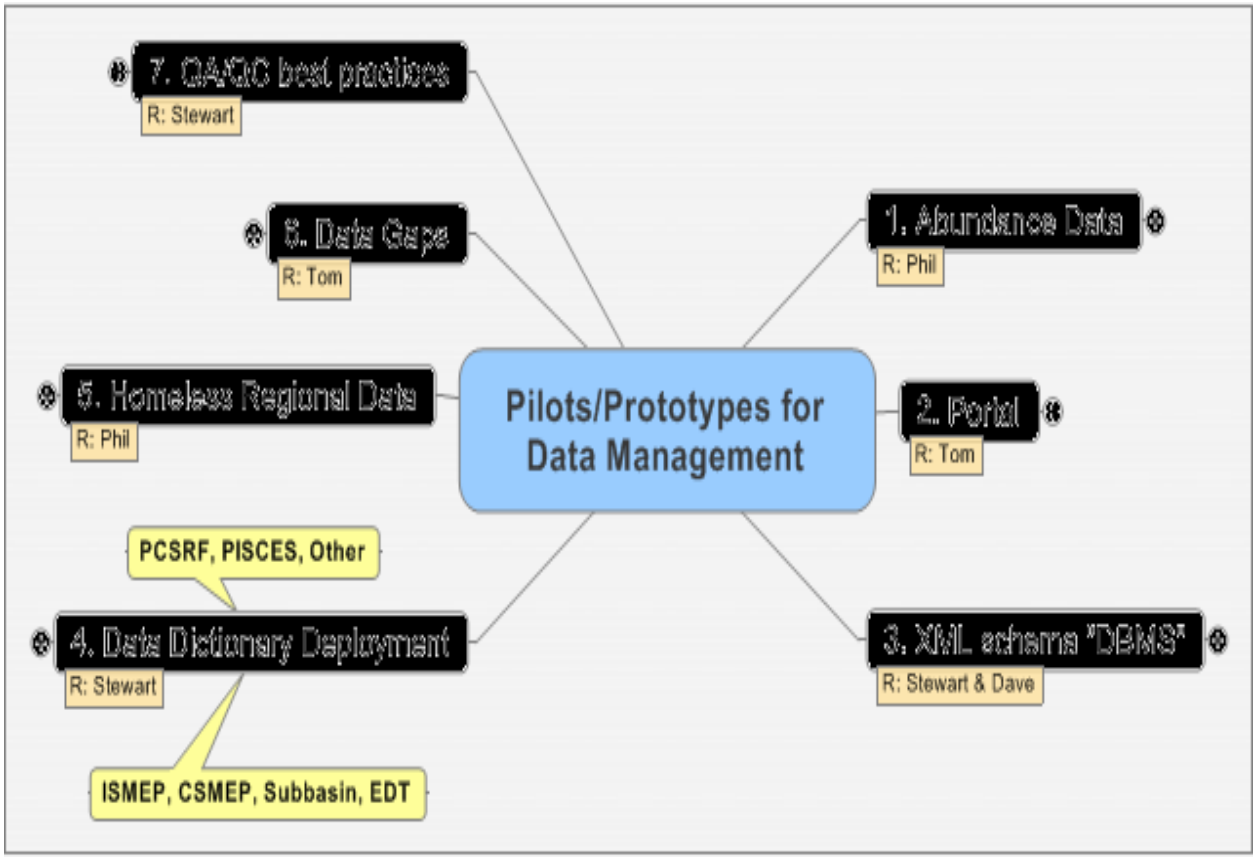
- Salmonid Data (Stewart Toshach). Work is continuing on the Regional Salmonid Status and Trend Data Collection Methods and Data Management Study – Meetings have now been held with most of the collaborating entities to get feedback on the statement of work, most recently with the State of Oregon.

  Stewart and Tom Pansky discussed, with Vivian Hutchison the availability of metadata training from the NBII program at the USGS. NED will take a lead in soliciting needed training across the NED partners. If there is sufficient interest we will work with Vivian to coordinate training which USGS will provide. The first available time for training (for 1.5 day sessions) will be in May 2007. The most likely location is the BPA computer lab in Portland. Stewart will work with PNWRGIC, PNAMP and NED to help coordinate metadata training needs.

3) Development of NED proposal, to the Council, for regional-level pilot or prototype efforts. We discussed the findings and next steps from the 9/19 ad hoc work group meeting attended Peter Paquet, David Tetta, Phil Roger, Tom Pansky, Tom Iverson and Stewart. Seven possible pilot/prototype efforts were identified at the meeting (see graphic below) – with volunteer leads and a goal of writing up 2 page summary papers for each prototype or pilot effort.

- Abundance Data – Phil
- Portal – Tom Pansky
- XML Schema (DDBMS) Stewart
- Data Dictionary Deployment Stewart
- Homeless Data (Phil)
- Data Gaps (Tom Iverson)
- QA/QC –Stewart

Phil and Stewart had prepared draft 2 pagers for discussion. Many comments were made on the drafts and the list of topics. The next step is for the all the authors to complete their draft two-pagers. It was agreed that they would be circulated on 10/11/2006 and the ad-hoc work group would meet again on 10/13/2006, at CRITFC from 9:00 through 12:00, to continue work on the prototype proposal development.



The schedule for completing the task is as follows

10/13 at CRITFC - 9:00 to noon
10/19 if needed - time and place TBD
11/1 at NPCC - during regular NED meeting

4) Consideration of comments on: *Best Practices for Data Dictionary Definitions and Usage* and *Check List for Organizing Field Collection and Management of Data*. The Best Practices document is posted on a Wiki site that David Tetta has established at http://nedworkgroup.pbwiki.com/FrontPage Comments posted are all positive. Joy has arranged for an edit of both documents. Final changes will now be made to complete the documents.

4) Other work items/updates (30 min)

- CBFWA – StreamNet Review. Tom Iverson has advised that the Summary Findings from the Workshop will be circulated on 10/5/2006 to the CBFWA Steering Committee for comments. It was agreed to discuss these further at the next NED meeting – November 1st

- Using the List Serve to comment on NED document/s. Stewart encouraged NED SC members to try the Wiki to comment on documents – see details above.

- International Columbia River Basin Center of Knowledge. Deferred to next meeting.

- OFWIM Conference. Bruce Schmidt and Janet Hess-Herbert are planning on attending.

- Fed Caucus Data Needs. Stewart reported that the Caucus wanted an update on what NED's work plan is – but that the work plan is currently under development (at the Council's request) as a set of proposals. Jim suggested that if NED wanted the Caucus to consider contributing to NED funding it would need to provide proposals very soon. There was further discussion about the need to get more formal participation in NED from the Caucus. Jim said he had raised the issue more than once with the Caucus and some members were reluctant to sign – the concern he has heard is that regional needs could end up conflicting with national needs. It was agreed that Jim would discuss the need internally with BPA Executive, with a view to encouraging formal Caucus participation, and Peter Paquet would discuss with Tom Karier, with a view to encouraging formal State participation from Idaho and Montana.

.

B. PNAMP DATA MANAGEMENT AGENDA ITEMS

1) PNAMP Inventory Task update. Bruce said the inventory was not complete and a time extension has been requested. There were a number of lessons learned with the effort: the difficulty of getting good locational data and trying to complete the effort in the field

season when many field staff are out of the office. The prototype work is expected to be completed by the end of November

2) PNAMP Scientific Data Analyst. Jennifer Bayer, with technical support from Joy and Stewart, has completed a description of work for a PNAMP Scientific Data Analyst. The main task is to evaluate and test Protocol Builder. It is not clear that the Federal Caucus will fund the work. There should be discussion about the possibility of requesting funding for protocol builder work through NED if the Federal Caucus does not fund. Jennifer and Stewart will discuss possibilities.

3) Information item. PNAMP has taken a lead on working to finalize the "Comparison of Data/Monitoring Groups" table that Bruce Schmidt had first drafted in 2005. PNAMP had received a number of comments and is providing a revised draft for comments. Tom Pansky asked for an introductory paragraph or two to set the context for the document.

C. AFTERNOON SESSION 1:00 through 4:00

NED Workshop

Use of EPA Distributed Data Base Management Technology to Collect and Distribute Habitat, Fish and Water quality Data

Karl Jacobs (Program Manager) and Brian Hale (Contract Application Developer) from the California Department of Water Resources joined the meeting to provide a briefing of the data management work completed in the Bay Delta and Tributaries Project and other efforts. A powerpoint of the briefing is on the NED web site at http://www.nwcouncil.org/ned see October 4th Agenda at Item C. You can also look at the BDAT site at see http://bdat.ca.gov. Karl and Brian provided our workshop with ideas about the needed the practical steps and their lessons learned from working on an EPA exchange grant and building their system in general.

Background and NED Interest:

Phase III of the NED work plan is to investigate what it will take to migrate towards the use of Distributed approaches and technologies for collecting and distributing data.

NED has an interest in exploring the use of DDBMS systems. In addition, the EPA exchange network has grants available to support pilot level efforts to apply the EPA exchange technology

and we want to discuss this. Information about the grants is available at http://www.epa.gov/exchangenetwork/grants/index.html An example of an Application prepared by the California Department of Water Resources *2005 Exchange Network Grant Submital* is posted on the NED web site for your information. See http://www.nwcouncil.org/ned October 4th, Agenda at Item C.

Agenda

1:00 (See above for participants)

1:15 Application of DDBMS to data collection and sharing – what steps, skills and resources are needed to develop and complete a pilot level project. An important task is developing translation type tools, or schema, that will facilitate the use of web services as a way of implementing DDBMS for biologic and habitat data – and this is one of EPA's priorities for funding network grants this year. Karl outlined how the BDAT system works. Essentially the program offers turn-key development of data entry systems for data collectors. The data is collected on MS Access templates which are customized depending on the collection and reporting needs of the users. No data analysis is offered by BDAT. Business rules for operation of the system are clear in that the providers of the data remain as owners of the data and they decide when to submit data to the warehouse and are responsible for QA/QC. The data is consolidated within BDAT to provide for regional level views of all data provided by the partners. The forgoing is what BDAT refers to as a distributed data management system. Some in the PNW, for example StreamNet, would refer to this as a centralized system.

BDAT is also a partner in an EPA Exchange system. BDAT uses this technology to exchange the data from their warehouse (above) with other data systems. The Entire system is called the BDAT Cooperative Data Management System (CDMS). In the PNW we would call the EPA system components a distributed system. CDMS provides for the exchange of data across multiple nodes within and outside BDAT. Each of the nodes represents a database with a collection of data from other providers. The data is accumulated in different ways according to the content interests and needed data business rules of the different groups.

Karl and Brian offered some practical suggestions for a pilot level undertaking in the Pacific Northwest:

- Base the proposal on a clear and strong business need – do not propose to collect all data for all uses;
- Clearly define the pilot in terms of scope and time;
- Evaluate the potential to field-test the controlled language that EPA is developing for fishery and habitat related data as a part of the pilot effort – for example the schema being developed by the Office of Water Quality;
- Develop a critical key question that will be answered by the pilot; and,

- If possible relate the need for the pilot to 303d or TMDL issues.

Next Steps:

It was agreed that the NED SC should proceed to work collaboratively with Regional Partners to develop a proposal for an exchange grant and a request for a cost share from the NPCC.

Three components of such a proposal were discussed:

1) Evaluate the use of EPA developed control language for Fishery and Habitat data and modify as necessary to meet PNW needs in a common schema.

2) Develop pilot for end to end demonstrations of data collection, using controlled language for data that is not currently collected, for example Tribal data.

3) Demonstrate an 'enterprise level' data integration – where the locally collected data (in 2 above) is integrated with other data via EPA exchange nodes.

Projects such as ISEMP, CSMEP and PNAMP and others are potential and important co-partners.

Project Host: CRITFC has offered to consider taking on the role as formal project sponsor. CBFWA is also a potential participant in coordinating fish and wildlife agency support and involvement and providing administrative support.