

**Draft Proposal for a Data Center for the
Northwest Environmental Data Network (NED)
May 2006**

Background: the Northwest Environmental Data Network Effort.

In May 2000 the Independent Science Review Panel wrote a report on the inadequacies of the data system for Columbia Basin fish and wildlife. They noted that there were significant data gaps and significant inconsistencies in the way that data were collected and reported. They recommended a systematic approach to address a wide variety of tasks including an inventory of existing data, a survey of unmet data needs, proposals for filling data gaps, and development of standardized data collection and reporting protocols.

Following this report the NPCC and NOAAF signed an MOU to Co-manage an independent data management needs assessment that was completed in 2003 by Science Applications International Corporation. Recommendations from the assessment were adopted by the Council.

To provide a practical way for regional partners to collaborate, in December 2004, a new MOU, for the Northwest Environmental Data Network, was drafted and signed by regional partners. With the new MOU in place the partners developed a phased work plan and began work approved by the Council and Partners.

Under the MOU and through the voluntary, collaborative efforts of NED we have verified a broad-based interest in resolving these problems. NED has also coordinated a broad variety of parties to work on these issues.

The following proposal builds on those initiatives and moves us forward in resolving them.

Background to This Proposal

In December 2005 Regional Executive of the Northwest Environmental Data Network met to discuss proposed actions for regional data collection and management. They reviewed the executive considered recommendations of the 2003 SAIC study, a May 2005 Beyond Ad-Hoc Regional Data Workshop Meeting and the NED Steering Committee recommendations.

The Executive then asked the NED Steering Committee to develop a work plan and cost estimates to complete the remaining tasks and made suggestions on how to move arrangements for regional data management beyond ad-hoc.

The NED effort has deployed a phased approach, endorsed by the NED Steering Committee, the NED Co-Chairs and the Council.

Phase I involved studying the issues and problems, developing solutions, getting comments on the comments and, developing an MOU between participants. The NED strategy has been to work towards improving regional data management at minimal cost, by relying on coordinated activities and for the most part, in kind support.

Phase II involves the establishment of an organized arrangement for providing regional data management program oversight, planning and coordination; the development of protocols to allow more effective sharing of data; and, the deployment of a regional data portal where all needed regional data can be published. Once published the data is available via the www. Effective sharing and use of shared data also requires the development and deployment of data sharing standards and protocols. While significant work has already been completed, including Pilot Portal development, and Best Practices for Reporting Locational and Time Related Data, many other efforts are still needed. An increased level of resources will be needed to complete these tasks.

Phase III of the overall NED plan involves the introduction of and transition to a Distributed Database Management System (DDBMS). A DDBMS system would connect major regional environmental data systems using a common information architecture. The technology for this approach has been proven with the Pacific Northwest Water Quality data exchange, which integrates NW state water quality data systems. The next steps are to add fishery, habitat and other needed data sets. The inherent advantage of this approach is that the data is retained by the

owner/data collector who retains responsibility for the data integrity but posts it for use by others. DDBMS approaches will supplement existing warehouse approaches and allow the sharing of warehouse data.

Most significantly, ongoing success in comprehensive data management will require an independent entity, that would be charged with planning, organizing and deploying a coordinated data management program to ensure that important data necessary to understand the status of fish and wildlife in the Council's program are adequate and available and to ensure effective coordination of this program within the wider regional context.

Vision for a Northwest Environmental Data Center

The development of a Northwest Environmental Data Center will provide extensive benefits for the entire region. It will provide a much needed outlet for comprehensive, high quality data that can be used by policy makers, technical experts and the general public. It will also work with the Council and Bonneville to ensure that data gaps are filled, standard data protocols are followed, and data quality is maintained.

It is recommended that the existing Northwest Environmental Data Network be developed into the NED Center to complete this function.

Responsibilities of the Center

The Center would not be responsible for collecting and compiling data, except for the collection and maintenance of metadata associated with the management of the North West Environmental Data-Network's Portal. The collection and compilation function would remain the responsibility of project sponsors and fish and wildlife managers. No analysis would take place at the NED Center. The NED Center would, however, be responsible for ensuring that the programs necessary to collect and manage the full array of important data--for hatcheries, harvest, hydro passage, and habitat--are available, reliable and adequately documented. The NED Center would work with the Council, Bonneville and other Partners staff to identify and remedy shortcomings in the current system.

Organizational Arrangements of the Center

The Organizational arrangements of the NED Center are shown in the attached diagram and would include the following:

A Review Function. Complete independent peer review of regional protocols, standards, plans and funding proposals based on a separate charter. Independent review would be established through a Regional Data Management Panel comprised of now fewer than 3 and no more than 5 members set up as joint effort

by the ISRP/ISAB. The membership of this panel would be chosen to ensure that it represents the interests of both the Columbia Basin and the wider region.

Executive Function.

The NED Center Executive would meet obligations under the Memoranda of Agreement (or other binding agreements), hire and direct staff; approve protocols/standards, plans and funding proposals; Coordinate with regional data entities to develop data sharing agreements; and Chair the Steering Committee.

Staff Function.

Draft plans and funding proposals; organize and coordinate work groups and work group products; work on education, and outreach; and manage projects and coordinate education and outreach efforts. A particular staff role would be to provide for data stewardship for the NEDC Portal Staff would work at the direction of the Executive.

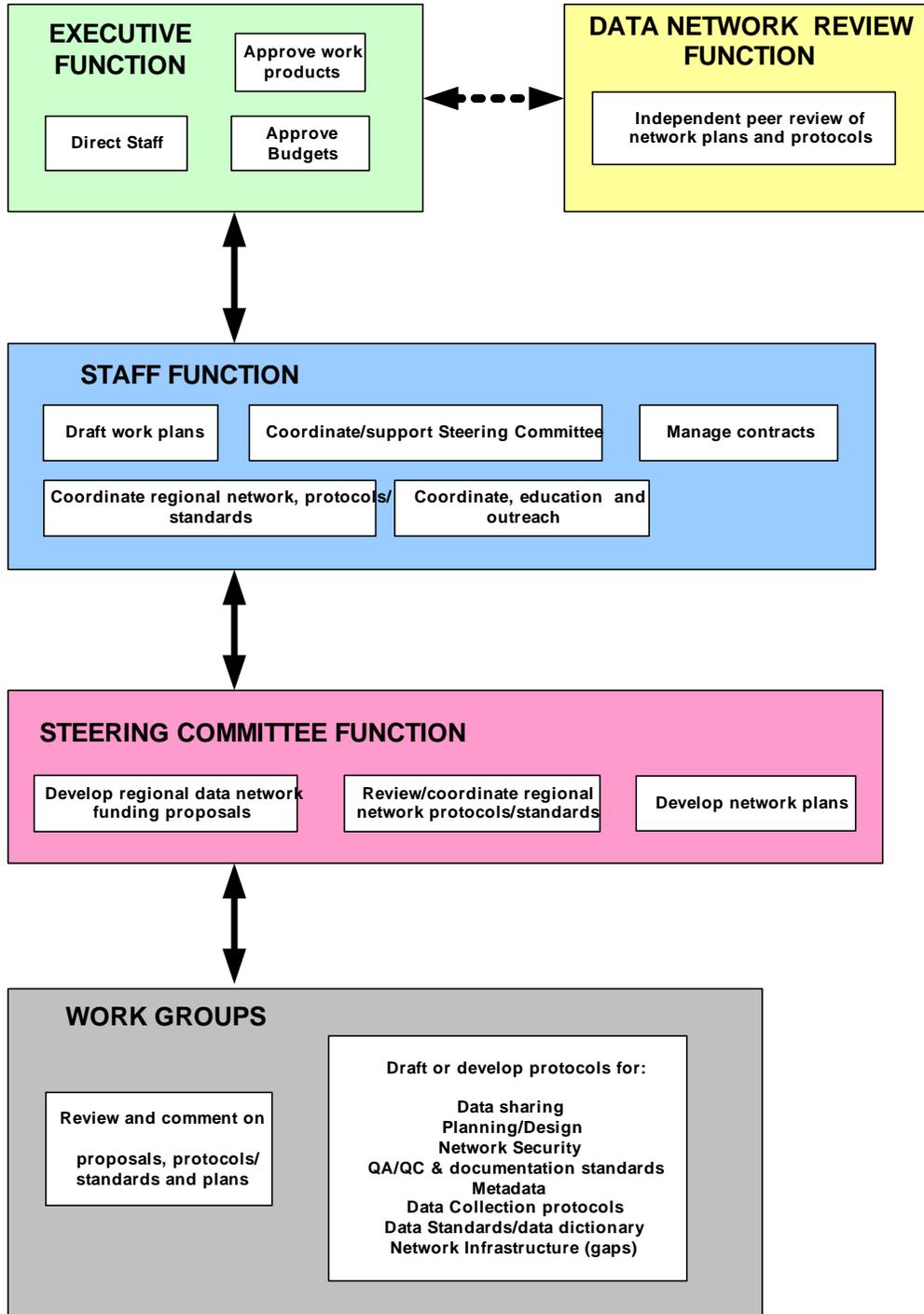
Steering Committee Function.

The Steering Committee would work on policy development efforts, including identification and development (or location) of needed data standards for data reporting and collection, provide liaison between users and producers of data, and develop commonly needed tools and guidance documents. The Steering Committee would be comprised of individuals whose Executive have signed the NED MOU.

Work Group Function.

Key individuals from the public, the Steering Committee, stakeholder groups, agencies and others would work to provide input to network proposals and plans and develop protocols and other needed network components and polices: for example, for data-sharing agreements. Work groups would be led by members of the Steering Committee with membership open to any participants, at the discretion of the Steering Committee.

Fig. I Governance Structure for Regional Environmental Data Network Center



Summary of Functions of the Center:

The Center would,

- organize around the principles of cooperation and integration
- involve communication between users and providers of information on a regular basis
- create a decision-making Steering Committee and associated work groups that are independent of the missions of any particular agency
- include federal, tribal, state, local government and private interests and user groups such as the scientific community, academic institutions, public information specialists, decision makers, program managers, and civic/environmental groups
- deliver and maintain data exchange and sharing agreements and tools, and coordinate regional data-management planning for data sharing and exchange
- develop and maintain a business case, funding and incentives
- provide evaluations of approved projects to ensure that they have met their data defined data management objectives

The NED Center would integrate and provide oversight for a regional data network and serve as a goal-setting body to ensure that the network addresses key priorities as expressed through plans, initiatives, goals, objectives and agreements. The Center would not replace existing efforts but would provide education, guidance, network resource tools, and technical support.

Primary Responsibilities of the Center

1. **Internet Access.** It will be responsible for maintaining a high level web site that would serve as a portal for existing data. This site would be user friendly for policy leaders, technical experts and the general public. It would rely on standard protocols as necessary to ensure that data from different regions and from different sources are compatible. And it would offer sophisticated web based tools for graphing, mapping, and consolidating data.

2. **Data Gaps.** It will be responsible for conducting static and dynamic inventories of existing data and determining the existence of data gaps and in consultation with various entities in the region, and facilitating solutions that would resolve gaps. Unresolved data gaps and proposals to resolve them would be reported to the Council and regional Partners.

3. Data Integrity. It will provide oversight over data quality, and have authority for setting minimum data quality standards necessary to ensure the integrity of the data. It will do this by periodically reviewing the procedures used by different entities and recommending solutions to the Council and Regional Partners.

4. Data Standards. It will identify or propose standard protocols for data collection, data reporting, and data quality to be considered for adoption by the Council and regional partners. These protocols would be applied to BPA funded projects as stated in the Program and, unless partners have more stringent standards to the Partners operations.

In order to achieve these results the NED Data Center must establish and maintain a reputation for neutrality and objectivity. It must also maintain a clear set of operational principles and standards against which data management proposals and activities can be evaluated.

At the same time it must continue to work closely with the many entities in the region that are users or purveyors of data. Independent review of NED Data Center programs and other data issues would be conducted, as necessary through a Regional Data Management Panel comprised of now fewer than 3 and no more than 5 members set up as joint effort by the ISRP/ISAB. The membership of this panel would be chosen to ensure that it represents the interests of both the Columbia Basin and the wider region.

It should be emphasized that the Data Center would, by itself, have no responsibility for actually collecting data in the field, a task often associated with monitoring activities. Instead, the Data Center would either be the recipient of this data or it would access it after it has been posted by others.

Implementation of the Center

In order to achieve this vision there are several important steps that need to be taken by the Council, the NED Co-Chairs and the NED partners, will need to convert the vision outlined in this paper into a more precise statement of work for the NED Center.

The Statement of work will need to ensure that the NED Center, can:

1. Can implement the tasks on schedule.
2. Can maintain a high level of integrity and neutrality.
3. Identify where data are collected but not publicly available.
4. Will remains independent from actual data collection, with the exception of the actual collection of regional metadata tracking system, PISCES, that promises to provide considerable value to the regional fish and wildlife effort this new Data Center is more than information technology, and will require a somewhat broader set of skills.

Bonneville will also need to develop a budget proposal for this initiative. As it does this they are encouraged to develop a more comprehensive data management budget that

includes all of the existing data management efforts in the region. The comprehensive budget should ensure that there are no duplications of functions and should identify efficiencies that can be gained by concentrating the planning and coordination of these data functions into one single entity.

Examples of how data management challenges can be solved by the Center

In particular, there is no comprehensive data inventory available for the Columbia Basin that identifies data gaps and no projects have been approved or even proposed to close those gaps. The Council has yet to approve guidelines for collecting or reporting data and there is still no Internet access for broad categories of Columbia Basin Data.

Addressing these in turn.

A comprehensive data inventory. By their very nature inventories are challenging. So much data is being produced that as soon as an inventory is produced it will be out of date. The solution that NED Steering Committee has developed is to use the NED portal as a way of maintaining a dynamic inventory. The success of this approach will depend on the degree to which the Portal is populated. For new data from the Fish and Wildlife program this would happen quickly if the F&W program conditioned all projects with an obligation to publish their data through the web portal. The task of completing the inventory for legacy data, and data collected by other entities would remain. The proposed solution to that challenge is to use data stewards to work proactively and systematically to publish other data via the portal. Where this data does not already have a web-enabled home a program is needed to collect, host and web enable this data. It is a significant task. The strategy of the NED SC has been to tackle the ongoing problem first (publishing new data).

Guidelines for Collecting and Reporting Data

The SAIC report recommended these actions. The NED SC depends in part on the work of PNAMP (established in 2005) to provide many of these data collection guidelines. PNAMP was established specifically to develop and evaluate protocols for data collection and is comprised mostly of scientists. It made sense for PNAMP to complete parts of this task because scientists, rather than data managers are best positioned to develop guidelines for collecting data.

There are however, some guidelines that apply to all data collection efforts and need not be identified by a scientific group. For these NED is taking the lead on development. One set of guidelines, *Best Practices for Reporting Location and Time Related Data* is complete and is currently available for adoption by the NPCC and other groups. NOAAF executive have signaled that they plan to deploy these Best Practices. Another area of focus concerns data sharing agreements. Exchanging and sharing data is administratively more complicated than one person sending a copy to another. In many instances permissions and assurances are required. NED is working to develop these data sharing

tools for adoption and use. The first task is writing sharing agreements for the data sharing portal.