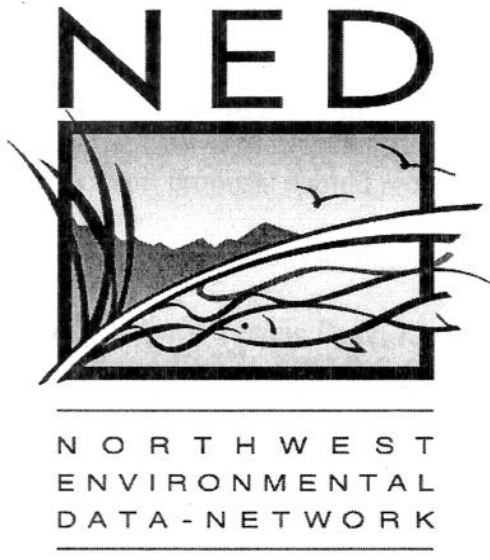


Encl mbrs  
SW  
PP  
DM  
SE  
JS



2000000026

**NED Steering Committee Response to the Council "Columbia Basin Data Center 4/24/06" Proposal**

JUN 22 2006

## **4 Existing Initiatives**

There are many initiatives already underway, which, if supported and well coordinated, could meet all the primary responsibilities identified in the CBDC proposal (in italics below).

Dealing with these in turn:

*4.1 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”*

This is primarily an IT issue which was addressed in detail by SAIC and can best be deployed by small group with appropriate expertise.

NED has already researched, developed and is testing:

- A “*high level web site that would serve as a portal*” for existing data to allow users to discover and locate the data on the web (note that the NED portal is not the same as the NED web site). It will require support from data stewards to support ongoing operation.
- The site is “*user friendly*” and uses a developed technology being used nationally by Geodata.gov and the National Spatial Data Infrastructure.
- The site relies on standard protocols and can ensure that “*data from different regions and” from different sources are compatible*”. Note that “*compatible*” does not mean that the data can be automatically compared. Comparability relies on Metadata: information about how when and why data was collected must be available before data can be compared, regardless of standards used. Tools for data integration are also needed and NED proposes to develop a distributed approach. This is consistent with the ISRP and the SAIC recommendations.

Implementing and maintaining portal functionality over time is more difficult than building an Internet portal. Participants have different abilities to publish information on the web. Much important data is maintained at a local or district level where local biologists typically do not have the training, equipment or time to undertake this task. It will take an organized effort to maintain information sharing capability as the variety of participants respond and individually adapt to new technology.

*4.2. “Data Gaps. It will be responsible for conducting inventories of existing data and determining the existence of data gaps. It would be their responsibility, in consultation*

*with various entities in the region, to facilitate approaches that would resolve gaps. Unresolved data gaps and proposals to resolve them would be reported to the Council and BPA.”*

Scientists, resource managers, decision makers and other experts on data content (not data developers or data managers) are the people who must be involved in making decisions about what data are needed, to compare needed data with existing data and therefore to identify gaps. There are a number of existing efforts working to identify needed data content in the Basin and the Region: PNAMP, CSMEP, Columbia Pilot Monitoring efforts, State IMW's, Tribal programs, the Puget Sound Partnership and cross-boundary efforts are examples.

- Once these groups have completed their own data inventories there is a further task of integrating data needs across these inventories and coordinating data collection across programs. It is important to understand that this is a two step task.
- NED is recommending the management of data content semantic information across multiple entities through the use of tools such as a regional data dictionary. This is a tool that can identify all the names, definitions and units of measure for data – to support systematic analysis and make progress towards standardization.
- The NED is testing the on-line portal to maintain a virtual inventory of data - as it is posted to the Internet. Once data providers publish metadata to the web, the inventory will be able to be queried to determine what data is collected, where and how, and what the data quality is.

**4.3 “Data Integrity.** *It will provide oversight over data quality, ensuring the integrity of the data. It will do this by periodically reviewing the procedures used by different entities to assure data quality.”*

Assuring data integrity is the responsibility of both content specialists (eg biologists) and Information Technology experts. Development and education about best data collection practices can improve the quality of the data collected. Development of software applications to screen and test data before passing it to higher levels can also improve data quality.

- Oversight of data quality is insufficient by itself. At a minimum - a set of regional Best Practices for Data Quality are necessary. Once developed Best Practices would need to be tested, deployed and maintained and the agencies doing the sampling, reporting and data analysis would need to adopt them. Education and training are key needed elements for data quality improvements.
- The NED approach and work plan has work elements for improving data quality- that can and should be addressed at all parts of the data collection cycle. Data

collectors, reporters, managers and analysts must all contribute to improved data quality.

**4.4 “Data Standards.** *It will propose standard protocols for data collection, data reporting, and data quality to be considered for adoption by the Council. These protocols would be applied to BPA funded projects as stated in the Program”*

Data standards for collection of data are best developed by data content specialists and there are (as mentioned above) existing groups in the region that are working on parts of these tasks. Data reporting standards and data quality standards can be developed for each of data collection or reporting method or protocol. Note that data quality solutions are not generic, but specific to each collection method. NED recognizes the need to develop and deploy standard protocols for multiple data sets and is working with other data content and data management groups to promote their use and development.

### **5.0 Statement of Work**

The NED SC is willing to contribute to the development of a statement of work to define needed tasks for improved data management, regardless of the scope of the effort.

### **6.0 Some Gaps in the Council Proposal:**

The NED SC believes that there are several important topics, relevant to successful regional data management, missing from the Council proposal. They are as follows:

- A need for formal agreements when adopting and using standard protocols and reporting methods.
- A need to improve data handling efficiency by automating routine tasks and shortening data pathways.
- An ability to consolidate data. Here, consolidating data means assembling disparate data into common views without manual data manipulation. A NED approach would lead to the use of a distributed approach, based on existing technologies and the adoption of agreed eXtensible Markup Language standards, to allow queries to multiple databases as an alternative to the manual collection of all data into data warehouses.
- An agreed upon structure for designing, planning and evaluating regional data programs. Currently there is no plan within the region or within the F&WP for developing and maintaining data management resources. NED can provide a collaborative forum for this function.
- The problem appears to be more complex than described in the (understandably brief) proposal:

- Many of the detailed issues can only be addressed by coordination amongst content groups (eg biologists) and data technology specialists. A long-term solution requires ongoing collaboration between these two types of expertise.
- Change at the data collection level is difficult because there is usually limited IT expertise and resources available. An expedient way to affect changes in data collection is to provide incentives, (equipment, expertise, and applications) which makes their jobs easier, as well as standards that allow data integration.
- The problem is large and complex, involving multiple programs, agencies, and different levels of information needs. Rather than a single “engineered” solution, we may make fastest progress by addressing one or a few specific needs to demonstrate success and build credibility. We also need to take the time to “grow” a regional data network.
- Management needs determine information needs and gaps. In the past, management needs have not always been clear. There is now, however, some convergence on a subset of information needs which are supported by a number of programs. These include fish population VSP parameters, and habitat conditions characterized by NOAA’s PFC report and EDT input data.
- Along with the need for access to a broad range of data is the need for tools to interpret that data. For instance, electric power managers use a variety of models to estimate future demand and balance that against expected generation capacity. Likewise, fish and wildlife interests need tools to aggregate data across a variety of spatial scales, create derived metrics of change, estimate the impacts of habitat change on fish and wildlife populations and etc. Analysis tools need to be developed collaboratively to assure a broad range of acceptance and trust in their results. Examples of needed tools are those being proposed by the Northwest Habitat Institute.

## **7.0 F&W Program Data Management Alternatives that Complement Regional Data Management Strategies**

There are some initiatives that could be taken immediately by the Council F&WP that would complement the broader regional approach.

7.1 Adopt Best Practices for “Reporting Locational and Time Related Data” which has been completed by NED and is available for adoption at any time.

7.2 Advise all BPA contractors and internal programs that they will, in the future, be required to comply with Council approved data reporting and data quality standards, beginning with interim standards for 2007-2009 funded programs. Publishing data to the NED portal is a recommended step in making the information discoverable and shareable.

7.3 Complete a more detailed and explicit review of current proposals for funding of data management initiatives from 2007-2009 for the Councils F&WP. A deliberate is warranted. The first step is to decide what data collection programs will be supported. The second step is to decide what mix of data management initiatives will best serve F&WP and regional program obligations.

7.4 Set a deadline (suggest October 2006) for completion of a first set of regional interim data collection and reporting standards from the multiple data content efforts that the BPA is funding. When the interim reporting standards are complete, test them on a subset of key FWP projects in 2007 and apply them more broadly in 2008 and 2009.

## **8.0 Conclusion**

While the NED SC understands the urgent need and the importance of making changes to improve data management for the Council's F&WP, it considers that doing so within a regional perspective and working through the programs that have already been developed, and in particular the NED effort to coordinate regional data initiatives, will provide the strongest foundation for the needed changes. The ISRB, the SAIC and the 2005 NED White Papers all identify the need for a strong regional solution and provide details of needed changes.

The NED Steering Committee and the Council have previously endorsed the findings of the ISRP and the SAIC studies that recommended a single regional entity to provide a forum for the development and completion of a distributed data design, data standards and other data resources. The Steering Committee is willing to provide this role in the region, however the current MOU anticipates funding and support to complete this task. The NED SC will bring a more detailed proposal to the Council if it is needed.

Alternatively, if there is a desire to proceed more rapidly at the F&WP level with initiatives that complement existing ongoing regional strategies - four actions have been outlined in Section 7.0 above. In the event that this is the preferred option, the CBDC should be developed collaboratively with regional partners and NED can provide an advisory committee function.