

**Final White Papers and
Recommendations from
Beyond Ad-Hoc:
Organizing, Administering, and Funding a
Northwest Environmental Data Network**

**Generated from Discussions at the
May 25th and 26th, 2005 Intergovernmental Workshop
Portland, Oregon**

TABLE OF CONTENTS

1.0 Background and Introduction	3
2.0 Executive Summary and Key Actions (Nancy Tosta - Ross and Associates)	5
3.0 Business Needs Case: (Burney Hill - USEPA)	10
4.0 Communications & Marketing: (Peter Paquet - NPCC & Lenora Oftedahl - CRITFC)	13
5.0 Governance: (Nancy Tubbs - USGS & Stewart Toshach - NED)	18
6.0 Financing: (Joy Paulus -WA IAC & Jennifer Bayer - PNAMP)	23
7.0 Data Content: (Mike Babcock -Yakama Tribe & Stewart Toshach - NED)	27
8.0 Technical Architecture: (David Skea – BC Ministry of Agriculture and Lands & Mike Beaty – BRec)	33

“Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network”

1.0 Background and Introduction

In the Pacific Northwest we need to improve our ability to discover and share data for fishery, habitat, aquatic habitat and upland environments. Many different groups collect and manage data of interest to a wider audience, however they mostly do it using different formats and methods. This has made data discovery and subsequent data sharing difficult, inefficient and expensive.

To respond to this challenge a workshop, ***Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network*** was convened on May 25th and 26th 2005 in Portland, OR to identify possible solutions. For the workshop the Northwest Environmental Data Network (NED) partnered with the Pacific Northwest Regional Geographic Information Council (PNW-RGIC), the Pacific Northwest Water Quality Data Exchange (PNW-WQX), the Pacific North West Aquatic Monitoring Partnership (PNAMP), the Northwest Habitat Institute (NWHI), StreamNet, State, Provincial and Tribal entities and others.

Over 75 people participated: data managers, GIS specialists and some data collection specialists. Representatives from non-profit, US Federal, State and Tribal, Canadian Provincial and other organizations attended. Participants included senior data managers/developers and program managers with lead organizational roles in federal, state, tribal and regional data efforts.

The goals of the workshop were to:

- Learn about information solutions currently being used for regional data discovery and sharing.
- Identify components and processes needed for regional information discovery and sharing e.g. metadata, best practices.
- Draft recommendations for needed administrative, organizational and funding arrangements that could improve regional data discovery and sharing.

There was consensus at the workshop that improved methods of data sharing and discovery across geographic and jurisdictional boundaries will improve our ability to answer management and scientific questions.

Day one of the workshop was mostly structured towards learning about options and solutions. In addition the participants completed a network vision exercise and identified 6 sub-topics (see below). On day two, the participants shared their ideas about solutions for each of the sub-topics. Nancy Tosta from Ross and Associates facilitated workshop sessions on both days.

The following participants then volunteered to write up ‘White Papers’ together with short and long term recommendations for each of the sub-topics - incorporating ideas from the workshop and other relevant materials:

Business Needs Case: (Burney Hill - USEPA)

Communications & Marketing: (Peter Paquet - NPCC & Lenora Oftedahl - CRITFC)

Governance: (Nancy Tubbs - USGS & Stewart Toshach - NED)

Financing: (Joy Paulus -WA IAC & Jennifer Bayer - PNAMP)

Data Content: (Mike Babcock -Yakama Tribe & Stewart Toshach - NED)

Technical Architecture: (David Skea – BC Ministry of Agriculture and Lands & Mike Beaty – BRec)

After reviewing the White-Papers, Nancy Tosta from Ross and Associates wrote up the Executive Summary and Key Actions

The authors met four times via teleconference. After the first drafts were completed the authors then had an opportunity to review and make comments on the other author’s papers.

2.0 Executive Summary and Key Actions

Resource managers, policy makers, scientists, and many others in the Pacific Northwest would like to improve their ability to use data for making decisions about salmon recovery, aquatic resources and watershed management. This requires improved means to find, access, and integrate high quality data on fisheries, aquatic and upland habitat, and water quality. Currently, these data are collected and managed across a diverse array of agencies and organizations, with different objectives, funding, and data management and formatting approaches. Data discovery, sharing, and use are frustrating, inefficient, and expensive. Recognizing this, a group of policy makers and data managers gathered at a workshop in May 2005 to begin to chart a course to address these challenges. The workshop, *Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network*, brought together the Northwest Environmental Data Network (NED), the Pacific Northwest Regional Geographic Information Council (PNW-RGIC), the Pacific Northwest Water Quality Data Exchange (PNW-WQX), the Pacific North West Aquatic Monitoring Partnership (PNAMP), the Northwest Habitat Institute (NWHI), StreamNet, State, Provincial and Tribal entities and others.

From the two days of meetings, several observations were made and challenges identified including:

- There are many existing systems - all wanting to be THE SYSTEM (some considering themselves to be “the one-stop”)
- Building shared databases is worthwhile, but requires incentives, migration from existing systems, and management of redundant systems
- Metadata registries are helpful to know who has what (e.g., Geospatial OneStop example)
- Regional systems may not be needed when national systems are being developed (e.g., Geospatial OneStop)
- Clearinghouses have been developed for data sharing regionally (e.g., hydrography data layer), but not all states or possible participants are using them
- Developing standards and protocols can take many years
- Many data dictionaries exist, but they are not “cross-walked” to optimize use as a “controlled vocabulary”
- The technology is not an issue (whether it is centralized or distributed), but who makes decisions such as how to establish and comply with standards is a challenge
- The “value” of data sharing needs to be demonstrated
- Data management and sharing requires data stewards, custodians, and business users to establish Memoranda of understanding to share data
- Governance of a regional system is complicated and requires careful thought to determine needs

The breadth of these observations and challenges led workshop participants to decide at the end of the discussions to develop a series of white papers on these topics (business need, governance, outreach and marketing, financing, data content, and technical architecture). Those papers are included herein. The summary of recommendations from the papers is shown in Table 1. Highlighted recommendations are suggested as important steps in the next year. Those that represent more immediate steps are discussed in detail following the table.

Table 1: SUMMARY OF RECOMMENDATIONS FROM WHITE PAPERS	
Business Case	Revise and advance existing “qualitative” business case presentation
	Conduct functional assessment of major databases and database technologies in the PNW and supporting standards and metadata
	Convene principals of major resource databases to formalize consensus supporting value of data network (expand NED signatories)
	Evaluate potential and requirements for “quantitative” business case (estimate costs to develop such a case)
Outreach and Marketing – short term	Develop NED brand and logo
	Develop pamphlets and handouts to describe NED capabilities
	Develop relationships with various data exchange groups
	Develop test site to show power of data linkages
Outreach and Marketing – longer term	Develop detailed plan to enlist support of agencies who generate data
	Develop glossaries and other tools to provide info to public and decision-makers
	Develop granting program to “sell” NED project
	Develop website that addresses data questions for wide range of stakeholders
	Develop speakers bureau to educate about data sharing
Governance	Present regional workshop findings to regional executives
	Identify essential additional partners and champions to support participation
	Identify current lines of authority and/or obligations – use business case to encourage participation
	Finalize formal governance and cost-share arrangement
	Develop, fund, complete, and review actions for regional data sharing/discovery
	Conduct annual executive-level review for regional data sharing signatories
Financing – short term	Establish information portal that links existing resources (with Columbia Basin as focal point)
	Explore pilot project

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

	Explore other network options to leverage (e.g., Exchange Network)
	Leverage grant opportunities and partnerships
	Organize executive funding symposium
Financing – longer term	Build long-term funding strategy
	Develop business plan that includes – financial, operational, marketing, and communications components
	Compile regionally relevant cost/benefit information
	Show costs if data sharing is not achieved
Data Content – short term	Identify costs of developing, deploying, and maintaining needed standards and protocols
	Focus on establishing minimum data exchange and sharing protocols for future data collection and management efforts
	Maintain and update regional data exchange and sharing protocols in a regional data dictionary accessible via the www.
	Develop education and outreach program to promote use of standards
	Link project and study approvals to use of regional standards
Data Content – longer term	Manage regional data legacy (inventory, prioritize, develop metadata)
	Maintain and update regional data sharing and exchange protocols via dynamic web-based dictionary
Technical Architecture - general	Adhere to policy of open standards
	Develop security requirements early and adhere to them
	Follow REST view of SOA – intelligent documents, simpler interfaces
	Actively conduct tech transfer program to support collaboration among stakeholders
Technical Infrastructure – short term	Develop and publish reference architecture for the system
	Deploy registry/repository that conforms to OGC extensions to ebRIM and provide access via simple web portal
	Deploy light-weight user configurable map browser
	Deploy access mechanism to download registered data sets
	Select small number of representative datasets and WMS services and register them in the catalog
Technical infrastructure – mid and long term	Deploy internet mapping framework that enables development of more complex applications
	Start adding rich content to catalog
	Add search engine capability to the catalog (Google)
	Deploy WFS based data replication as data access mechanism
	Continue to add data, services, rich content
	Add support for WCS (web coverage service) (for access to GEOSS data holdings)
	Develop ontology
	Interact with other data sharing networks

From this list the following are suggested as starting points to be addressed in the near term (Oct-Dec 2005).

1. Communicate critical needs with regional executives and solicit their agreement on the following activities. Provide quarterly updates for regional executives.
2. Convene a small working group of leads and technical staff of 3-4 representative regional data collection entities. Invite a state (WA or OR) with EPA exchange network expertise to participate. Within this group consider/agree to the following:
 - a. Development of a “Challenge Grant” proposal to U.S. EPA [Fiscal Year (FY) 2006 Environmental Information Exchange Network Grant Program] to secure funding and additional expertise in development of a regional data-sharing network¹ (states or tribal entities are eligible – can contract with others). See:
<http://www.epa.gov/exchangenetwork/grants/FY2006.pdf> Deadlines are as follows:
 - i. September 30, 2005 – Deadline for submitting questions about this notice to EPA
 - ii. **October 17, 2005 – Question-and-answer teleconference for applicants**
 - iii. **November 21, 2005 – Deadline for submitting applications to EPA**
 - iv. June 2006 – Issuance of FY 2006 Exchange Network Grant Program awards
 - b. Commitment to work together to accomplish regional goals and to leverage existing resources to the extent practical to undertake at least one joint pilot that begins to demonstrate the benefits of regional data sharing. This pilot will be based on open-standards and web services to show what might be accomplished using current data. The pilot will likely identify short-comings of data not designed to be integrated – and can thus help establish needs for minimal data standards.
 - c. Jointly develop web-services/portal that begins to expose knowledge of all existing data integration, sharing efforts within the region via searchable metadata descriptions.

¹ Possible topics to consider in grant proposal: cross-agency planning, and pilot implementation project to establish web-service based cross-agency secured web portal, using exchange network infrastructure; prototype of geo-services outlined in this paper; serious assessment of the cross-federal opportunities. Output of the grant could be operational prototype and list of specific proposals on how to take next steps forward.

- d. Discuss incentives needed to engage technical representatives in regional data sharing discussions.
3. Conduct a workshop for funding agencies to identify current expenditures for data infrastructure that might contribute to achieving regional goals for data sharing. The NED signatories could comprise this group. The group should also consider at least 3 things to do differently in the next fiscal year of funding to better contribute to accomplishing the goals of regional data sharing. Discuss incentives that might be available to promote regional goals.
4. Review with NED signatories successful models of governance that exist in other data sharing and network endeavors. Develop better understanding of why many data sharing efforts do not succeed.

3.0 Business Needs Case: for a Distributed Environmental Information Network Among Federal, State, Municipal and Tribal Governments In the Pacific Northwest

Burney Hill (EPA)

I. INTRODUCTION

The Pacific Northwest's natural resource and environmental management agencies and inter-agency organizations recognize the value of information technology (IT) and information exchange. Environment and natural resources do not align with political and jurisdictional boundaries. Cross-boundary collaboration in studying, communicating and addressing the issues and management of our environmental and natural resources requires ready access to information addressing the scope of such issues. State and tribal departments, inter-governmental organizations, non-governmental organizations and federal agencies broadly agree that the use of common data standards, data dictionaries and cross-walk tables and read/copy access to databases... information paid for almost entirely by public funds... can, would and do improve our control and protection of the environment and our management of natural resources.

Many of these departments, agencies and organizations have committed to advance the exchange of environmental information across the region. Specific reference and commitment to the value of and need for environmental information exchange is found in the Pacific Northwest's natural resource and environmental management program reviews and its interagency charters². We have a consensus that sharing information improves the resolution of issues in land, air and water management, the protection of threatened and endangered species, and the control of invasive species which challenge us at federal, regional, state, municipal and tribal levels. Responding to the challenges facing the Pacific Northwest's environmental management agencies and the Northwest Environmental Database network, assessments and plans have been completed that address the need for and value of an enterprise-level environmental information network of distributed, interconnected database networks³.

² *See:* <http://www.nwcouncil.org/dropbox/Data%20Sharing/NED%20MOU.pdf>, http://www.exchangenetwork.net/exchanges/water/pnwwqx_tpa.pdf, <http://www.nwcouncil.org/library/isab/isab2004-1.pdf>, <http://www.nwcouncil.org/library/isrp/isrp2000-3.pdf>, <http://www.reo.gov/PNAMP/Products/PNAMP%20Draft%20Plan%20Jan%202004.pdf>, and <http://www.streamnet.org/about-sn/pub-docs/fy99quick2.htm>.

³ *See:* http://www.epa.gov/oei/imwg/pdf/business_plan.pdf, http://www.epa.gov/OEI/imwg/pdf/final_blueprint.pdf, and <http://www.exchangenetwork.net/>, and

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

II. DESIRED STATE

Agencies across the Pacific Northwest, and the Nation as well, seek efficient and economical access to comprehensive data that is of assured quality and that is relevant to their missions. At the level of information technology we desire to be able to turn on our computer, to quickly and easily find a complete list of the locations and characteristics of data and other information pertaining to the environmental and natural resource challenges that we face, and to be able to acquire such data and its metadata promptly. We further desire that the database management and transfer systems function so as to promote the integration, analysis and presentation of data obtained from numerous sources. Within the sphere of information development, management and exchange we desire to have an effective and efficient organization and support structure for achieving data exchange.

III. CURRENT ISSUES and CHALLENGES

The primary issues and challenges to expanding the information or data exchange network are institutional in nature rather than technological. Current information technology has proven its ability to locate, access, transfer and present data. Our issues consist of advancing our use of common data and IT vocabularies, languages, standards and procedures for the location, access, and transfer of data. Our challenge is to move past our insular views and investments and to embrace changes in IT, IT management and local work flows that promote a greater good at system and enterprise levels which will quickly deliver profitable dividends at the project and programmatic levels within which we work. Our opportunity is to appreciate the huge successes of information exchange evident in search engines and websites on the internet and to accept these and countless other examples of the success, benefits and cost-savings of information technology referenced herein as footnotes.

IV. BUSINESS IMPACTS and BENEFITS⁴

1. Save Money/ Avoid Costs
2. Save Time
3. Increase Efficiency
4. Increase Accuracy
5. Increase Productivity
6. Increase Communication and Collaboration
7. Support Decision-making
8. Automate and Improve Work Flow
9. Expand and Enhance Information Bases

<http://www.aero.org/publications/crosslink/winter2005/06.html> .

⁴Measuring Up, the Business Case for GIS, by Christopher Thomas and Milton Ospina, 2004, ESRI Press, Redlands.

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

10. Manage Resources
11. Improve Public Access to Government
12. Establish Enterprise-level Leverage

V. OPTIONS FOR FUTURE ACTIONS

1. Revise and advance the present *qualitative* business case presentation without falling into the trap of expensive, time-consuming quantitative cost-benefit analysis and comparisons.
2. Conduct a functional assessment of the major databases and database technologies in the Pacific Northwest and their supporting data standards and metadata. Identify and clarify the common ground and important differences in these data, databases, standards and metadata. Reference national and/or international standards in current use or approaching endorsement. Prepare a report and distribute this report on-line.
3. Convene a meeting of the principals of the major environmental and natural resources databases to extend and formalize the consensus supporting the value of a data exchange network. Expand the signatories to the memorandum of agreement of the Pacific Northwest Environmental Data-exchange network. Secure these agencies' and organizations' commitment to advance the technologies, the governance and the operational support program office that will be necessary to establish such a data exchange network. Focus and capitalize on agencies and organizations that have current financial and/or programmatic commitments to operating or developing information a exchange network (e.g., ODEQ, WDOE, USGS, NOAA Fisheries, EPA, CRITFC, NPPC, BPA, PNAMP, State of the Salmon Consortium, Northwest Habitat Institute et al.).
4. Evaluate the potential and requirements for a *quantitative* business case for an environmental information network in the Pacific Northwest. Estimate the costs and time requirements for the development of such a quantitative business case.

4.0 Communications Outreach & Marketing:

Lenora Oftedahl, StreamNet Library, Columbia River Inter-Tribal Fish Commission & Peter Paquet, Northwest Power & Conservation Council

The Northwest Environmental Data Network is developing a model in which data sharing between various agencies and other organizations becomes the norm rather than the exception. In order to truly recognize the picture of the environment, all aspects of environmental data should be shared between agencies. SAIC outlined regional needs for the communication, education and outreach portion for regional data sharing (see Appendix A) in their report to the Northwest Power & Conservation Council (NPCC).

What the Northwest Environmental Data Exchange seeks is a marketing plan to sell this data sharing initiative to the state, federal and tribal agencies as well as non-governmental organizations involved in natural resources in the Columbia Basin. We also need to sell NED to the general public in order to help them understand where their money is going. This effort may also help the general public to better understand the environmental situation in the Pacific Northwest. Currently the biggest message they get is the price tag involved in saving Pacific salmon with very little to show for what is spent. Creating products that educate the public and communicate what exactly scientists are doing with the money being funneled into the river should be another aim of the NED project.

We need to answer the following questions in order to come up with a plan that will help us reach customers, potential partners and funding agencies.

1. What are our products and services?
2. Who are our customers?
3. What would our customers like to buy from us?
4. Why should our customers buy from us?
5. What sets us apart from competitors?⁵ Do we have any competitors, or should these be considered potential partners?

Marketing involves two parties agreeing to a mutually beneficial exchange.⁶ In this case, participating agencies exchange access to their data for access to other agencies' data. There will be a need to develop other incentives to data sharing. Is NED going to simply be a clearinghouse for information, or will there be some value-added by NED participants/analysts? Another relationship to explore is what potential funders will get from NED in exchange for their money.

⁵ Bangs, David H. *The Market Planning Guide: creating a plan to successfully market your business, product, or service.* 6th ed. Chicago, IL: Dearborn Trade Pub., 2002.

⁶ Stern, Gary J. *Marketing workbook for nonprofit organizations*, v.1, *Develop the plan.* 2nd ed. St. Paul, MN: Amherst H. Wilder Foundation, 2001.

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. *Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.*

DESIRED OUTCOMES:

A regional data exchange to improve natural resources management in the Pacific Northwest. All agencies and tribes will contribute data to the exchange and use the exchange for data management efforts.

In addition to the cooperation of agencies and tribes, the data exchange would also be well served to educate the public about why these types of information are important and how they impact management decisions.

What other values do we wish to communicate? Of primary importance is the value of data sharing to increase the agencies return on investment. As members of NED, the agencies and tribes need to be convinced that they are gaining more than what they are expending.

Other values include the benefits of a healthy environment to the livability of the Pacific Northwest, and the economy of combining efforts to preserve and manage ecosystems.

CURRENT ISSUES/CHALLENGES:

According to SAIC's CBCIS summary, the numerous agencies and tribes working throughout the Basin are difficult for users to navigate to find the information and/or data they seek. "No single entity or site stood out as being 'the site' for information about the Columbia River Basin." In addition, while data centers may be found, they are often not 'public-friendly'. Scientists and researchers are served by data centers, but even they have to hop from one to another in order to put together a complete picture of current conditions in the Basin.

Even the data centers do not relate themselves to other data efforts in the Basin. The puzzle pieces are all laid out on the table, how do we put them together? That is left up to each individual.

OPTIONS FOR FUTURE ACTION⁷:

Step 1: What will people do? How much of what do we want by when?

- We need to attract people, attention and funds to the project.
- We want agencies and tribes to contribute data
- We want the public to understand how basic statistics drive management decisions.
- We want a test/demonstration site up by the end of this year (2005).

⁷ Stern, Gary J. Marketing workbook for nonprofit organizations, v.1, Develop the plan. 2nd ed. St. Paul, MN: Amherst H. Wilder Foundation, 2001.

Step 2: Finding your niche.

- Key audiences need to recognize the value of what NED is offering.
 - Illustrated by the website: <http://www.chesapeakebay.net/> on the lower left hand site is a box labeled View Site By Interest.

Step 3: Identify the Six P's

- Product—what do we offer?
- Publics—target audiences (managers, data collectors, data developers, general public)
- Price—what do we want in exchange?
- Place—where the product is available (or geographic coverage)
- Production—ability to meet demand and serve customers
- Promotion—what we do to convey image and motivate people to respond ; what incentives are needed to guarantee participation?

Step 4: What steps are necessary to implement the plan? Who's going to do what, by when and with what resources and support?

- Business plan is already being developed along with a governance scheme.
- Develop a common message that can be translated for different target audiences.
- Travel to agency and tribal offices to deliver personal message about NED and why the agency/tribe should participate in the effort.

Step 5: Build the image and motivate people to respond. Reinforces desired image and conveys a specific message that tells people what you want them to do.

- We want people to find the NED site easily and be able to navigate and use information found on the site.
- Building from the ground up through agencies and tribes means working with the data developers to ensure they have tools that will enable them to more easily participate in data sharing. They also need to have permission from managers and supervisors to work with NED.
- Bribery with gift certificates to Starbucks (or 'insert favorite beverage vendor here') would probably buy cooperation from a few data developers.

RECOMMENDATIONS FOR SPECIFIC ACTION TO ACHIEVE DESIRED OUTCOMES:

Short term:

1. Develop the NED brand and logo.
2. Develop pamphlets and handouts to guide the various audiences to what they need through NED.
3. Develop relationships with existing data exchange groups to ensure access to those sources of information. This effort may include financial support to strengthen and

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

support their ability to continue to provide data. In developing these relationships and showing how the data centers/exchanges can be linked may also help to sell NED.

4. Develop a test site to show the power of linkages between data originators and current data exchange groups.

Long term:

1. Develop a detailed plan to enlist the support of agencies from which data originates. This plan might best be written by a marketing group.
2. Develop glossaries and other educational tools to provide information to the general public and decision-makers.
3. Possible develop a granting program to sell the NED project. “We will give you \$\$ to support what you need to do (personnel, hardware, software, other tools); in return, you need to provide your data to the exchange.”
4. Develop a website that answers question for a wide variety of stakeholders.
5. Develop a Speakers’ Bureau to send knowledgeable speakers out to general population groups and other organizations to speak about the power of data sharing.

Table 1 Appendix A: SAIC Regional Outreach and Education Needs

Expand outreach efforts to seek buy-in from other key decision-makers and stakeholders in the region. Develop targeted outreach and education materials for key regional data network participants and supporters that clearly outline the need for a regional data network and describe the benefits and costs for such an endeavor. Ensure this outreach approach addresses the need for long-term support for a regional data network to	I.1.3	The first emphasis here is on executive level consultation with stakeholders on plans, obligations and expectations. At the same time, further public outreach is essential to raise awareness.
Develop communication and coordination hub of regional data network	I.3	Web site and collaboration site
Develop and post a regional data network guidance manual that documents everything needed to become a regional data network participant	II.5	
Expand regional outreach and investigation to other segments of the regional community not included in the original requirements analysis		
Further evaluate regional information needs against available information resources to develop acquisition strategy		
Develop management and public information/communications work groups as part of the regional administrative structure		
Develop a regional public data outreach strategy		
Conduct regionwide public workshops to advertise and seek feedback on recommendations		

5.0 Governance for a Pacific Northwest Data Network for Fishery, Water and Upland Habitat Data

Nancy Tubbs - USGS & Stewart Toshach - NED

I. Introduction:

This paper is concerned with establishing a governance structure for a **subset** of regional environmental data: fishery, water and upland habitat.

Studies recommending regional-scale data management in the Pacific Northwest point to the need for a governance structure^{1,2}. Governance is, however, a sensitive topic. It immediately plays to participants' concerns about control and the need for sensitive management of proprietary data and data security. Clearly a workable governance structure will need to overcome these concerns.

At the workshop no one disputed that governance is needed in order to effect lasting improvements in data sharing and exchange mechanisms. After all, data is a commodity and commodities are typically shared and exchanged based on rules and conventions, and usually through nominated exchanges. All of this requires governance. There must also be benefits from participation and these must be identifiable to individual participants.

Section II below identifies the desired state of governance for regional data sharing and discovery.

II. Desired State:

A formal and accountable governance (or administrative) framework would be in place and funded. It 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

The Committee 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, White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

initiatives, goals, objectives and agreements. The Committee would not replace existing efforts but would provide education, guidance, resource tools, and technical support. In summary – it would lead policy development efforts including data (not scientific) standards for data reporting and collection, provide liaison between users and producers of data, and issue tools and guidance documents.

A desired state provides for the use of common protocols, exchange formats, and standards for sharing data. Users, including the public, would be able to discover and access data from multiple data sources via the world-wide-web with single queries. Data would be made available and maintained for use by the individual entities that have responsibility for the data. The quality of available data would be described as a part of the metadata. Relevant FGDC-compliant metadata would be associated with the available data. An active regional program would recover, archive and maintain legacy data and make them available. An on-line regional data dictionary would be developed and maintained, which would allow data users to translate the data and understand what it means. Regional executives would have committed their organizations to the use of common data collection, sharing, and exchange protocols. User groups and technology specialists will be jointly involved in the development of protocols, guidance and other regional data products.

III. Current Issues/Challenges:

There are still questions about how the right people can be brought together but there is no argument that the “right” people include data generators and as well as senior administrators, technologists, and policy representatives.

Obtaining broad participation. Increased and broader participation is needed. Past participation in regional-level data management has been difficult to sustain. Data sharing and discovery based on a single issue limits both the breadth of participation and the duration of participation. As single issues are solved, or re-prioritized, the data-sharing capacity and data-integration potential are greatly diminished. A desirable alternative is to develop a data sharing and exchange governance that survives individual problems or issues. Apart from a political “ah-ha” that this realization needs, such a governance task requires a commitment to a ‘trickle’ of funding rather than the “feast or famine’ strategy that now characterizes most regional data management efforts.

There was some difference of opinion on whether governance should be a “light touch,” mandatory, or voluntary.

It was even suggested, that to be successful, the process should make itself invisible or redundant. In practice, this question has one answer - governance would need to be a light touch and voluntary. There is no mandatory region-wide authority requiring data sharing and exchange (except for certain subsets of data, e.g. some treaty data) and there is not likely to be region-wide mandatory authority. By default then, the approach is light and voluntary. Some entities may decide to make data sharing mandatory, and existing legal authorities already make data sharing mandatory, e.g. FOIA requests.

There were varying opinions on whether NED is the right umbrella for a governance structure.

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

There is no correct answer to this question. A move to a more regionally-governed approach rather than the current ad-hoc approach is likely to happen when the pain and cost of existing arrangements are more than the pain and cost of making the necessary changes. Previous efforts have produced some positive results but have not resulted in a regionally-governed approach. If the NED umbrella is not the right governance umbrella for data sharing and discovery, then management choices are to: 1) modify NED to make it 'right', 2) develop alternatives outside of NED or 3) do nothing. The only alternative suggestion offered at the workshop was to create the governance entity as an NGO.

What would a model of governance look like?

Governance requires an explicit organizational structure and functions. Needed governance structure and functions are described in Fig. 1 below: Governance Structure for a Regional Environmental Data Network for Fish, Water and Upland Habitat Data:

Review Function. Complete independent peer review of regional protocols, standards, plans and funding proposals based on a separate charter. This function could be optional; however, independent peer review of data network planning, protocols or funding proposals is desirable.

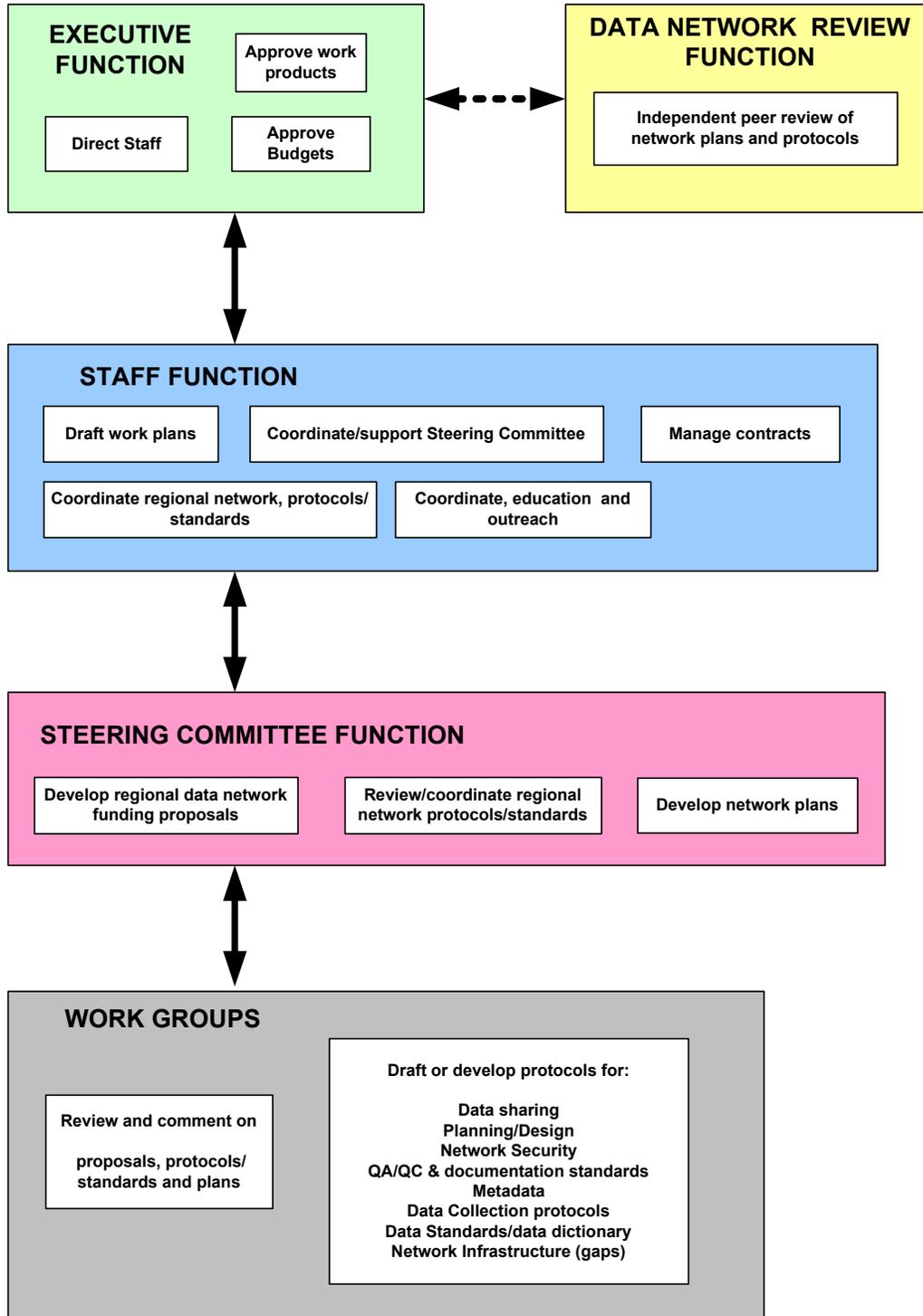
Executive Function. Follow obligations under Memoranda of Agreements (or other binding agreements), hire and direct staff; approve protocols/standards, plans and funding proposals; cooperate with regional data entities. Chair Steering Committee.

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.

Steering Committee Function. Lead policy development efforts, including needed data (not scientific) standards for data reporting and collection, provide liaison between users and producers of data, and issue tools and guidance documents.

Work Group Function. Key individuals from the public, stakeholder groups, agencies and others provide input to network proposals and plans and develop protocols and other needed network components and policies: for example, for data-sharing agreements.

Governance Structure for Regional Environmental Data Network: for Fish, Water and Upland Habitat



White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

IV. Future Governance Action needed to move towards the desired state:

- 1) Present findings from the regional data workshop to regional executives.
- 2) Identify essential additional partner entities and a champion in each entity who would work to support their agencies' participation.
- 3) Identify current lines of authority and/or obligations for data sharing and discovery and use the business case to support participation from new partners.
- 4) Finalize a more formal governance and cost-share arrangement and present to regional executives.
- 5) Develop, fund, complete and review short- and long-term actions for regional data sharing and discovery.
- 6) Conduct an annual executive-level review meeting for regional data network signatories.

6.0 Financing:

Joy Paulus -WA IAC & Jennifer Bayer - PNAMP

I. Introduction

Financial coordination across regional governmental and organizational boundaries is undoubtedly one of the biggest challenges in aligning access and delivery of environmental data needed for salmon recovery and watershed health. Access to regional fisheries, water quality, water quantity and uplands habitat data is inhibited because we haven't found a solution to this problem.

The NED Intergovernmental Data Sharing Workshop demonstrated regional interest in integrating existing information holdings and coordinating new data collection efforts. Workshop participants recognized that it is more cost effective to coordinate and share a common approach than to function as independent "stovepipes" of information. This approach clearly makes sense from a long-term perspective. However, commitment to finance modernization of current information sharing practices is critical for success.

To meet the difficult challenge of financing this collaboration, we must have executive level commitment from regional entities. Executive support will be necessary to leverage existing programmatic investments and procure new funding to sustain a collaborative effort. This will ultimately yield a regional-scale return on this investment.

II. Desired State:

The desired funding state would include a governing body to provide regional leadership, including the ability to procure and direct funding to carry out the agreed upon goals of the partners. It would ensure that:

- Fiscal needs are clearly identified and aligned with products (results), including separate needs for establishment and maintenance of the partnership
- Policy and executive level support is obtained and sustained over time
- Funding partnerships (mechanisms) are established and function through time
- Funding to share and exchange data is available so all organizations are able to participate equally
- Communication and marketing plan is established and implemented for regional coordination

In order to achieve this desired state, a series of smaller steps to prove this model's viability would be required. First steps may include:

- Compile a catalog of functioning data exchange networks and distribution systems with specific identification of the participants, description of the

information, and technologies used for evaluation for applicability and adoption in the Pacific Northwest

- Compile a list a potential pilot projects that could be used as a “proof of concept” for achieving regional goals and justification for the re-alignment of funding over the longer term
- Procure interim funding through grants and/or partnerships to cover an appropriate pilot project
- Build on the information sharing model that EPA developed (Exchange Network) and capitalize on work and experience gained in the Pacific Northwest
- Establish an information sharing and integration funding strategy (institutional, administrative and communications) to support collaborative information management
- Develop a funding partnership outreach plan
- Link goals defined in other white papers (from NED Workshop) to the funding plan
- Link key regional resource management goals with information management goals to ensure their alignment

A governing oversight body should be established to work with regional executives to ensure that short-term goals and achievements are communicated and that longer-term funding strategies are pursued.

III. Current Issues and Challenges

There are many issues facing data managers in the region. Complex reporting requirements (i.e. the need to report information at programmatic, organizational, state or federal jurisdictional levels) is just one example. However, it is generally not technical issues that plague the integration and coordination of environmental information but rather large scale institutional change that comes slowly and not without trepidation. The difficulty in financing and/or diverting earmarked funds to new cross-organizational efforts is formidable. The ability to align individual organizational business needs with needs at the regional scale is the crux of this problem.

Specifically, challenges to data coordination include:

- Difficulty aligning programmatic activities across organizations and jurisdictions
- Difficulty redirecting existing funding sources
- Difficulty securing new funding sources
- Difficulty expressing the financial benefit to each partner from cross-agency coordination, therefore making funding difficult to procure
- Capitalizing on the need to meet internal data requirement first so that work benefits collaborative effort.
- Difficult to demonstrate cost of missed opportunities due to lack of coordination
 - Difficult to demonstrate cost to others when requirements for data and information don't line up with others' needs and agencies go their own way
- Difficult to identify extraneous costs where duplication of effort occurs

- "Start-up costs" are high since some disciplines (fish, habitat) are new to standardization efforts

Addressing these issues will make information more accessible and useful to government policy makers, decision makers and citizens. It will also ensure more effective use of existing water, fish and habitat data, ultimately resulting in better environmental decisions.

IV. Options for Future Actions:

To achieve the desired state of financial support for collaborative data sharing in the region, we must work through these issues and seek opportunities to collaborate on shared solutions.

Progress may be made by a series of small steps while still keeping our eyes on the larger goal. The broad objectives would be to:

- Promote cooperation, collaboration, and technology transfer among state, federal, and local governments and other entities that collect, manage, and/or use fish, water and habitat data.
- Explore alternative funding strategies and incentive programs for public and private partnerships
- Incorporate data sharing requirement into local and regional recovery funding plans and grants for state and federal financing assistance

Specific short and long-term goals to focus on in order to achieve success over time would be:

Short Term Options:

- Begin with an information portal that links existing resources at one access point with the Columbia Basin being the focal point
- Focus on small steps that show progress and success – explore a pilot project that tests the premise
- Find convergence between existing activities and this one (for example, leverage work that's been done by the Exchange Network)
- Leverage grant opportunities and partnerships in identified priority areas
- Organize an Executive Funding symposium to begin the discussion about budgets and funding to help defray participation costs. Identify the key player for this executive discussion (for example: ESA (NOAA); Columbia Basin Fish and Wildlife Program (Federal Caucus and Northwest Power Conservation Council); Clean Water (EPA); other needs such as Harvest Management Programs, Federal/state/tribal land management programs, County Lands Use and Permitting)

Long Term Options:

- Build a long-term funding strategy
- Develop a business plan that includes financial, operational, marketing, and communications components that would meet this shared vision
- Compile a regionally relevant cost/benefit information to meet the long-term vision of data access, standardization and sharing
- Demonstrate what the additional costs would be if coordination not achieved (and the savings if coordination successful)

7.0 Data Content: Developing, Adopting and Deploying Regional Content Standards and Protocols

Authors: Mike Babcock –Yakama Tribe & Stewart Toshach - NED

I. Introduction:

Determining the data or information “content” that would be desirable and possible for a Regional data network is a continuing discovery process. In 2003 Science Applications International Corporation (SAIC) identified the need for standards and protocols to help us discover, understand or share data⁸. There are at least 3 different aspects of data standards – content standards (standards for what the data is), exchange standards (standards for how data is exchanged), and data management standards (standards for how data is stored and archived).

SAIC recommended that the following standards/protocols concerning content be developed to support more efficient and effective sharing and exchange of data:

Metadata tools. Metadata is essential for exchanging, sharing and using data. For distributed architectures they provide the basis for searchable indices of information

Geographic data: latitude and longitude, map coordinate datum and map coordinate projection.

Data collection storage and analysis: for example common calendar/data policy, methods codes, regional data dictionary, common monitoring methods, codes and station names

Quality Assurance and Quality Control: Procedures and consistent approaches to complete quality assurance and quality checking. Users of data must be able to understand the quality of the data

Documentation standards for data processing and analysis: Written material that explains how the product was generated and what assumptions were used. Much is maintained in a metadata record however it is important that detailed descriptions of data derivation be maintained.

System security protocols: It is necessary to define security protocols and chain of custody, for certain shared data sets, for example: who has ability to create,

⁸ Science Applications International Corporation, May 2003. Recommendations for a Comprehensive and Cooperative Columbia River Information System. Report to the North West Power and Conservation Council.

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

update, delete or edit data files. Users also need to know that the data is backed up and/or mirrored data sets are available.

The SAIC report completed a preliminary inventory of information resources. Further work on the inventory is a necessary step to define some data content for a regional data network.

An important data content question concerns the **type** of data needs to be shared and exchanged. SAIC reported that 90% of their survey respondents were concerned data collection efforts concerning fish information, 67% with water quality, 63% with water quantity, 60% with habitat, 57% with physical environment, 57% with hydrology, 53% with land use and ownership, 37% with wildlife and 43% with other. Clearly, while survey respondents have multiple overlapping interests, the information collection efforts are weighted heavily towards fishery data and related water resource issues.

II. Desired State:

The desired state of regional network data/information for fishery, water and habitat data content would be a network of data networks that would provide decision-makers, researchers and the public with access to comprehensive data/information they trust. Standardized regional data collection, quality assurance and storage protocols would be implemented and used by all data gathering and processing entities and all legacy data would be brought into conformance with standard storage schemas. Data gathered within any basin by any project on any topic (such as juvenile Spring Chinook out migration), could be included in queries, summary statistics or trend analyses encompassing other basins or projects in a timely and meaningful way. Processed data (information), in the form of interpretive reports would be indexed and easily accessible through search engine functionality. All data and information would be geo-referenced with common parameters to allow spatial analysis and presentation. Metadata and data dictionaries would be complete, concise, available via the web, and inclusive of the regional information spectrum. Most of the forgoing depends on a clear understanding of data content and the adoption of data standards/protocols for network participants.

III. Current Issues/Challenges and Options for Future Action:

There is professional agreement on what data content standards and protocols are needed to support information management systems that deliver understandable and high quality data. As mentioned above, the SAIC study identified a list of needed

standards and protocols. This list has broad acceptance. This is not too surprising since the list is relatively generic.

Standards must be developed (or adopted), deployed and maintained

The current challenge is develop, deploy and maintain the needed standards. This requires a considerable level of effort and a particularly acute focus on the details of the data. Semantic differences and terminologies must be understood at a detailed level. A close working relationships between information system data developers and content specialists is essential: for example with scientists who have a detailed knowledge about what the data is and what it means. A cross-disciplinary effort is needed and it is relatively time intensive. Substituting data developers for content specialists or vice versa does not work. Work with PNAMP and other content-related groups will help data managers understand what content specialists need.

Approach needs to be systematic

Standards and protocols are parts of a systematic approach. While all parts are needed and must be interrelated they do not all need to be developed at the same time. Therefore there is some flexibility in planning and completing these tasks.

Nevertheless, a systematic planned approach is needed, with executive level buy-in. The need for standards and protocols must, eventually be reflected in program priorities. This means that regional executives need to set clear deadlines for delivery of data protocols and standards and fund the efforts to deliver products in a timely fashion. While some (mostly larger) organizations have developed internal standard data management practices this is not the case externally, between organizations. Since current efforts are mostly ad-hoc this is particularly important. Without timely products and commitment to deployment we will never get “ahead of the curve”.

Institutional challenges

As well as programmatic challenges there are critical institutional challenges. In the current, primarily ad-hoc data management environment the lack of standards essentially reflects a lack of awareness, and subsequent programmatic commitment of the need to improve data sharing and exchange.

It is important to understand that within scientific entities and programs, *corporate* (standards based) data management programs and habits are relatively immature. Where standard practices have been developed they are different. Within most scientific entities, corporate data management is a relatively new initiative. Managing data is also mostly viewed as a staff support task that must be completed in order to meet the entities more important core mission. The budgets of most agencies are allocated to deliver (often competing) projects and products. Therefore program managers have a large influence in budget allocation. Every resource they ‘give up’

to corporate data efforts diminishes program success. Change towards corporate data management therefore requires a strong-voice for an enterprise level data system, since the corporate interest must prevail over the individual program interest. In most entities related to environmental data management this has not yet happened and therefore it is a major challenge. It is also important to recognize that there have been past failures in corporate regional data management efforts (for multiple reasons) – to the extent that there is understandable caution about committing to large scaled efforts.

Standards and protocols development and deployment also requires extensive cooperation and collaboration within agencies and externally. Cooperation and collaboration happens between people at a personal level. Usually it does not happen by accident and must be nurtured and supported. In other, words it is a significant challenge to build collaborative and cooperative actions, and to support them. Can a minimum set of inter-organizational data sharing and exchange standards be developed and deployed?

Concerns about sharing data

Fears about data sharing and exchange are also a potential challenge to participation in regional data sharing and exchange efforts. For example, some entities may not want to share legacy data that may be “ugly”, such as data that is not validated or subject to QA/QC. Anecdotal accounts suggest, for example, that some programs have collected data for narrow purposes such as completing annual reports with few resources invested into making further use of these data. Entities may therefore be reluctant to expose data like this. To avoid time and resources being spent by entities to massage or clean up data before exposing it, a data amnesty system could be established.

There are also institutional and individual attitudes to data ownership that affects sharing. In the scientific community sharing may not be supported until the scientific community that collected the data has completed all analysis and publications that they need to make. This can lead to long delays before the data is available to others who need it. Data sharing guidelines and protocols can help with this challenge.

Adequate Participation

Developing or adopting standards or protocols requires broad participation. To earn acceptance, and to understand nuances across entities all data entities with an interest in collecting using or sharing regional data, broad and diverse participation is needed. Getting full participation is a critical challenge.

A significant input of time and focused effort is needed. Since most entities first priority is to intra-agency obligations and second priority (if it is a priority at all) is to inter agency obligations it is difficult for many agencies to make staff commitments.

On the other hand most of the standards and protocols identified as needed in this region could be completed if entities committed FTE Resource's. A six-month FTE commitment from each major entity would go a long way on these tasks.

Alternatively the work could be supplemented with contract support. For acceptance though, staff still need to be closely involved. Since most agencies have not yet developed data management standards and protocols for their own use, participating in a cooperative effort would provide direct benefits to each agency as well as benefits to the region as a whole. If contract support is the basic method to complete development and deployment entities will need to develop a joint budget and be willing to support it.

Dealing with Legacy data

There are substantial legacy data resources in the region. How can these resources be related to a system based on standards and protocols? Common data formats for the specific information or data sets that need to be shared are needed. Given the scope of existing legacy data sets, this should be completed on a priority basis. Therefore decisions are needed on these priorities. Individual data collection entities would process their data into a form that can fit this format. Meta-data would be used to describe how this was done for individual data sets. Based on this meta-data the end user would decide what if any of these data they want included in their analyses, and derive their own justifications.

IV. Options for Future Action: Short and long terms steps necessary to address the challenges and move towards the desired state?

Short Term.

- 1) Identify the costs of developing deploying and maintaining needed standards and protocols. What can be developed now and what can be done later? What are the priorities? What standards and protocols have worked well elsewhere and so could be reused? Identify incentives for participation.
- 2) Focus on establishing minimum data exchange and sharing protocols for **future** data collection and management efforts – for implementation as soon as they are available but no later than 2007. Identify resources needed to complete these tasks and get entity support for resource commitment. Develop these protocols and standards so that all regional data entities have an opportunity to participate.
- 3) Maintain and update regional data exchange and sharing protocols in a regional data dictionary. The dictionary should be accessible via the www. Subject the standards and protocols to peer-review.

4) Develop an education and outreach program to promote use standards and protocols

5) Link project and study approvals to the use of regionally acceptable standards and protocols.

Long Term.

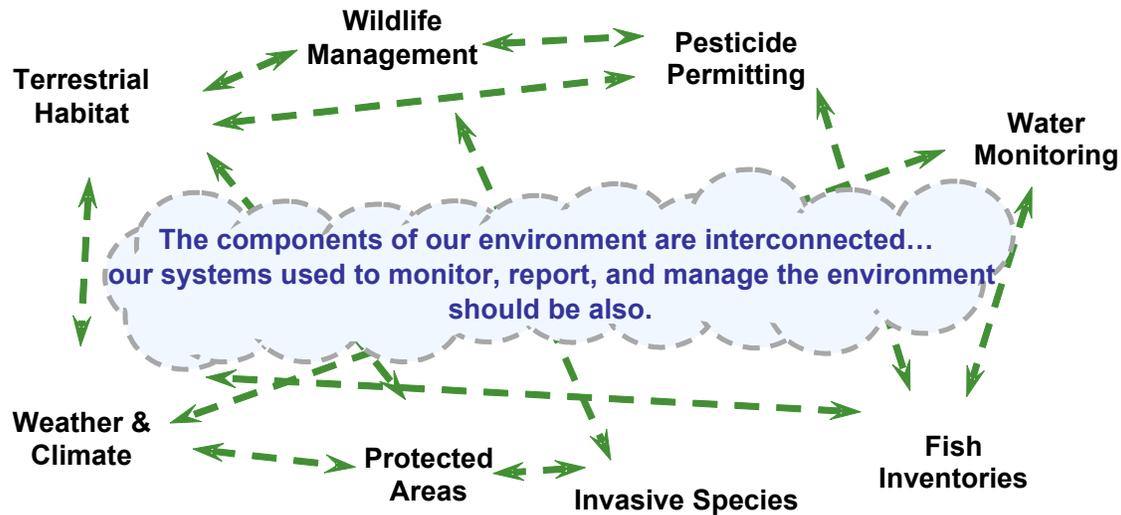
1) Managing our data legacy. Inventory legacy data sets and prioritize in terms of importance in answering scientific and management questions. Develop metadata records for these data sets where they do not currently exist and make these meta-data accessible. Support the archiving and recovery of these data.

2) Maintain and update regional data sharing and exchange protocols and standards via a dynamic regional data dictionary on the www.

8.0 Technical Architecture:

David Skea – BC Ministry of Agriculture and Lands & Michael Beaty – USBRec

I. Introduction



The interconnected nature of the world around us demands that specific environmental stewardship programs be aware of other programs that have overlapping interests and hopefully be able to interact with them. When one examines the extent of overlapping subject areas it quickly becomes apparent that there are very few silos, everything is connected to something else. This contrasts to arrangements for data management where we have multiple institutional silos or ‘databases’ that are usually poorly connected.

From a technology viewpoint the need for integration of the myriad of different systems can be addressed in two ways: 1) develop a monolithic centralized system or 2) develop a loosely coupled distributed system. Given the maturity of the internet, open standards for interoperability and the emerging service oriented architecture (SOA) frameworks, the clear technical direction for NED is to develop a loosely coupled distributed system. This system will be based on:

- A reference architecture to ensure that components will work together in the overall system,
- Adoption of open standards to guarantee interoperability, and
- Use of the SOA framework to support discovery and access to both data and services over the internet.

II. Desired State:

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

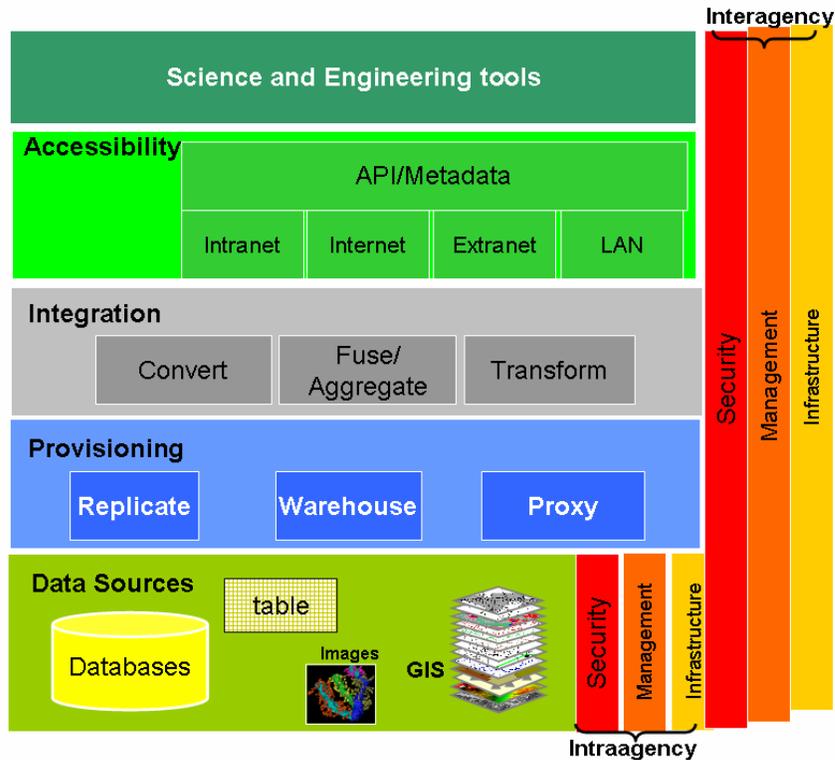


Figure: DataNet framework (from THE USACE DATANET – PROVIDING EFFICIENT DATA ACCESS THROUGH A NET-CENTRIC FRAMEWORK, C. Denise Martin, U.S. Army Corps of Engineers, May 2004)

What we would like to see is a loosely coupled distributed network of services and data, accessible over the web, and based on open standards including those from the Open Geospatial Consortium (OGC), OASIS and W3C. The following hypothetical “use case” illustrates how this concept might be applied.

A person working for a small municipality has collected detailed information on a number of contaminated sites. The plan is to re-visit these sites periodically and update the information, possibly adding new sites or retiring other sites. The objective of gathering this information is two-fold: 1) to make this data available to organizations doing planning on a larger scale, and 2) to monitor the effect of these contaminated sites on water quality and fish habitat within the municipality. Being a small municipality, the budget is small, under a thousand dollars. How can this be done?

1) First, the data must be housed in an environment that supports the OGC WMS (Web Mapping Service) specification. For example, a standard PC running Windows with a static IP connection to an IIS web server and Manifold 6.0⁹. Manifold is a desktop GIS

⁹ This is not a product endorsement, just an example of an environment that supports OGC WMS services and is reasonably priced. There are also free GPL-licensed options but they require more internal IT support.

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

that sells for \$250 USD. Next, our municipal employee goes to the NED portal and registers this WMS service with the NED catalog server. Done! Anyone with access to the NED portal can now point at this dataset and include it in an analysis or mapping application.

2). Recognizing their limitations to build a GIS application that monitors the effects of contaminated sites on water quality and that the affected watershed boundaries extend beyond their boundaries, the municipality turns to NED to find pertinent data and services and to organize a community of users that will enable them to leverage their development efforts.

This community of users then decides it wants to build a water quality and fish habitat application that runs on the Web. The application allows the user to identify an arbitrary stream reach or water body and generates a report as to habitat suitability. The application is based on an Internet Mapping Framework (such as the OGC compliant offering from Moxi Media). The framework supports the rapid development of spatial web applications and, with no internal knowledge of data, it derives its information from web services registered in the NED catalog. Services to run this application include:

- Topographic base information coming from a WMS, e.g. the USGS National Elevation Dataset (NED).
- Road network data coming from a WFS (Web Feature Service).
- Contaminated site information coming from a group of WMS services like 1) above.
- Upstream query service coming from a WPS (Web Processing Service), e.g. USGS Streamstats, that given a point on a stream, returns the upstream watershed polygon for that point.
- Other data, e.g. tax parcels, zoning districts, land use, land cover, can also be included

Notice that the hypothetical application not only accesses data distributed on several servers but also make use of services that carryout computation, e.g. returning the upstream polygon based on an input stream point. Once this application is built, it can also be registered in the NED catalog for others to discover.

From this use case the following requirements can be ascertained:

- Data should be owned and managed “at the source”, not duplicated by being pulled into a central warehouse or housed redundantly in multiple operational systems. This does not necessarily eliminate the need for NED or other organizations to compile and host some data (e.g., bandwidth issues, complex security issues, etc.), but it does minimize this requirement.
- Data will be accessed via a small number of standard interfaces. For our example three standards interfaces from the OGC are used: the WEB Mapping Service

(WMS, maps rendered as images), the Web Features Service (WFS, actual geospatial data), and the Web Processing Service (WPS, computed results).

- Data will be exchanged using self-describing XML (eXtensible Markup Language) technology. Specifically, all data exchange will be based on a XML dialect defined by an XML Schema. This will include: GML for spatial data, SensorML for sensor data, XACML for security policies (or another XML security infrastructure if XACML fails to maintain support), and other dialects for other types of scientific data.
- Deployment can be carried out by institutions of any size. The implementation described in 1) above represents the cheapest entry into this field. There are GPL-licensed free software (WMS → MapServer, WFS → GeoServer, spatial database → PostGIS) options. Although free, they take a higher level of effort to deploy and maintain, perfect for universities and small research organizations. At the high-end there are offerings from Galdos, Ionic, ESRI and CubeWerx that support enterprise level systems based on Oracle, IBM and Microsoft databases.
- Support of a registry/catalog for geographic information and geospatial services (see figure 1 below). This will enable the dynamic discovery of web-based services or data. Ideally, it should support both the UDDI interfaces and the OGC extensions to ebRIM (the Registry Information Model from the ebXML initiative). The goal is to catalog all shared resources, such as service descriptions, XML (and GML) schemas, definitions of coordinate reference systems, map styles, geographic data sets, imagery, technical report and sensor data records.

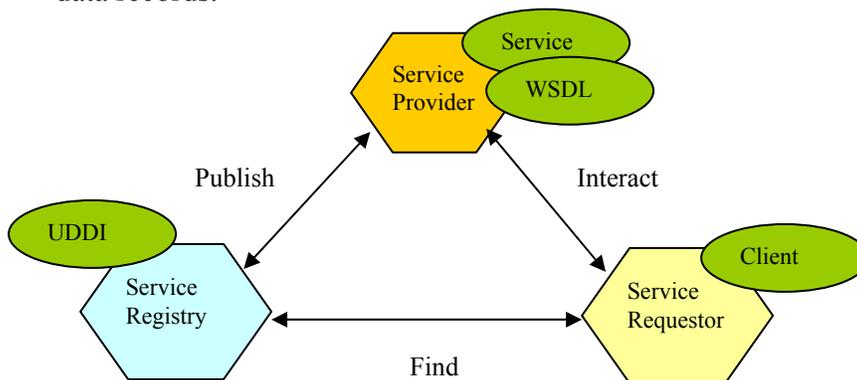


Figure: Service-Oriented Architecture (from THE USACE DATANET – PROVIDING EFFICIENT DATA ACCESS THROUGH A NET-CENTRIC FRAMEWORK, C. Denise Martin, U.S. Army Corps of Engineers, May 2004)

NED will also have to support the exchange of physical dataset for users who want to support standard desktop mapping application (GIS systems) and other types of analysis. However, the more copies of a dataset that get made, the more non-coordinated changes happen and over time the copies get mutated into different, new datasets. These new copies never get reconciles and nobody wants to become custodians of the new data.

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

Over the long term, keeping data close to the custodian, and only allowing distributed access to it, is the best architectural option.

III. Current Issues and Challenges

From the technology viewpoint there are no current issues or challenges facing the development of the NED network. One could build it today based on current open standards and widely available components. However, there are two core component that should be understood because of their long term impact and two issues related to competing standards. These will be discussed below:

Security: Initial visions for distributed data-sharing systems like NED typically ignore security issues as the purpose is to share data, not hide it. However, security issues almost always emerge once the system goes into production. Emerging standards such as XACML and SAML provide a strong foundation on which to build a security architecture. Another useful tool is the ability to hide data behind a WPS service. This allows access to aggregated or summarized views, not the raw data that may have restrictions on it. Security requirements should be collected early and worked into the architectural design from the beginning.

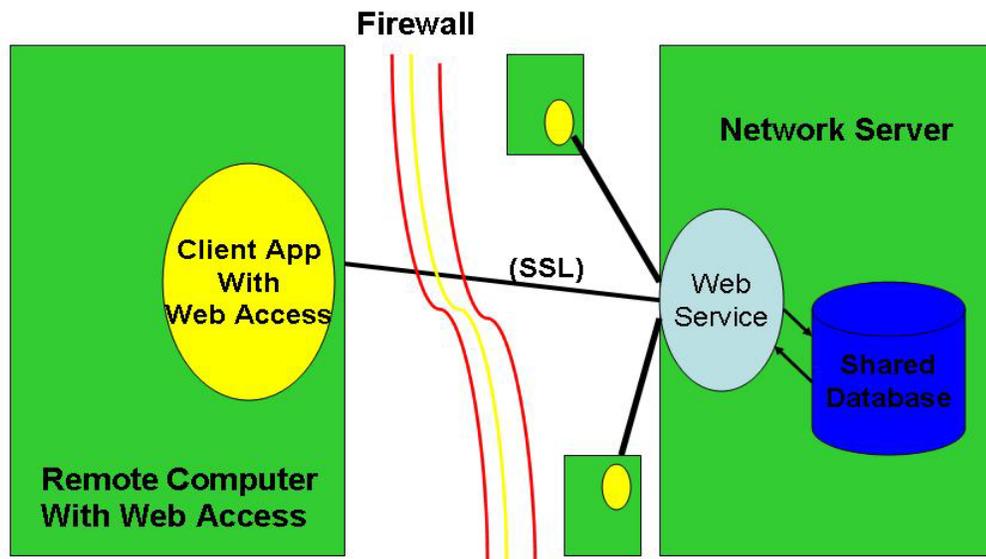


Figure: Database Service (from THE USACE DATANET – PROVIDING EFFICIENT DATA ACCESS THROUGH A NET-CENTRIC FRAMEWORK, C. Denise Martin, U.S. Army Corps of Engineers, May 2004)

Ontology development and merging: Discovery services are only effective if you know what you're looking for and metadata are published in a searchable form. This is a huge

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

problem for the Web in general and one that NED must successfully address if it is to grow and prosper. A suite of interrelated technologies are emerging aimed at addressing this issue. These include: RDF, OWL (Web Ontology Language), and SWRL (Semantic Web Rule Language). This is certainly not a production solution but significant work is being carried out in this area (e.g., <http://sweet.jpl.nasa.gov/sweet>) and progress in this area should be followed.

SOAP versus REST: Service Oriented Architecture (SOA) is an architectural paradigm based on distributed components linked by standardized open interfaces. There are two main streams of development in this area. The first is the Web Services Framework being promoted by Microsoft which is centered around the SOAP interface. This framework is highly complex and tries to emulate distributed object models (e.g., CORBA) with XML content. The second approach, known as REST, strips away much of the complexity and focuses on the documents. This is the paradigm adopted by the OGC and should be followed for the NED architecture.

Discovery, ebRIM versus UDDI: The initial publish→find→bind architecture of the web services framework was based on a directory specification (the find part) known as UDDI. This specification has a very structured hierarchy and was found to be too restrictive for general e-business. To support a more flexible discovery service the ebXML community have developed ebRIM specification. This specification was adopted as a standard by the W3C in 2005. As the NED catalog will need to hold information about a wide variety of objects, adoption of the ebRIM model is recommended.

IV. Options for Future Actions:

From a technical perspective there are clear directions for the development of such an infrastructure. Many similar networks are being developed and a common theme is emerging.

General Options:

- 1) Adhere to a strict policy of using open standards. This does not mean using only open software, it simply means that interaction with proprietary software is through open standards.
- 2) Develop security requirements early and adhere to them. There should be a fairly strict formal process for changes to security policies.
- 3) Follow the REST view of SOA; intelligent documents, simple interfaces.
- 4) Actively conduct a technology transfer program and support collaboration between stakeholders.

Short Term Options:

- 1) Develop and publish a reference architecture for the system.
- 2) Deploy a registry/repository that conforms to OGC extensions to ebRIM and provide access to it via a simple web portal.

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

- 3) Deploy a light-weight user configurable map browser that will let people view spatial layers found in the registry.
- 4) Deploy an access mechanism for downloading registered data sets.
- 5) Select a small number of representative datasets and WMS services and register them in the catalog.

Midterm Options:

- 1) Deploy an Internet Mapping Framework that enables rapid development of more complex applications consuming WMS, WFS and WPS services.
- 2) Start adding more rich content to the catalog: more datasets, WMS, WFS and WPS services as well as links to other objects, reports, spreadsheets, etc.
- 3) Add search engine (Google) capability to the catalog.
- 4) Deploy WFS based data replication as a data access mechanism.

Long Term Options:

- 1) **Add more data, services and other rich content types (real-time sensors) to the catalog.**
- 2) **Add support for Web coverage service (WCS). This will be the main access mechanism to Geos data holdings (Earth observing satellite sensors).**
- 3) **Ontology development.**
- 4) **Interaction with other data sharing networks.**

V. Glossary

OGC → Open Geospatial Consortium, Inc.: non-profit, international, voluntary consensus standards organization that is leading the development of standards for geospatial and location based services. Currently there are 275 companies who are members of OGC including: ESRI, Intergraph, Oracle, IBM, ...

GML → Geographic Markup Language: XML encoding for the transport and storage of geographic information, including both the geometry and properties of geographic features.

WFS → Web Feature Server: describe data manipulation operations on OGC Simple Features (feature instances) such that servers and clients can “communicate” at the feature level.

WFS Filter → Syntax for forming spatial queries in a WFS transaction.

WMS → Web Map Server: creation and display of registered and superimposed map-like views of information that come simultaneously from multiple sources that are both remote and heterogeneous.

White Papers and Recommendations from the May 25th and 26th, 2005 Intergovernmental Workshop. Beyond Ad-Hoc: Organizing, Administering, and Funding a Northwest Environmental Data-Network.

SFS → Simple Feature SQL: specification application programming interfaces (APIs) provide for publishing, storage, access, and simple operations on Simple Features (point, line, polygon, multi-point, etc).

OGC-CS → Catalogue Services specification: defines a common interface that enables diverse but conformant applications to perform discovery, browse and query operations against distributed and potentially heterogeneous catalogue servers. The registry/repository can be used to hold:

- ▶ XML documents
- ▶ XML Schemas (XSD documents)
- ▶ UML models (XMI documents)
- ▶ Map Service configuration files (AXL, MapServer map files, WMS capabilities documents)
- ▶ Coordinate reference System descriptions (CRS documents)
- ▶ OGC Styled Layer Descriptor documents (SLD documents)
- ▶ Web Service interface descriptions (WSDL documents)

XML → eXtensible Markup Language: a syntax for data transport and a suite of specification for defining interoperability over the Web. Everything on these pages are part of XML or XML related.

XML Schema → One way of defining the structure of a XML document. Replaces DTD. GML is a set of XML Schema documents.

SAML → Security Assertion Markup Language: XML-based framework for exchanging security information based on the expression of security in the form of assertions about subjects. These assertions are stated using XACML.

XACML → eXtensible Access Control Markup Language specification. Access is controlled by defining Access Control Policies in XACML and applying them to registry objects.

SOAP → Simple Object Access Protocol: a lightweight protocol for exchange of information in a decentralized, distributed environment. It is an XML based protocol that consists of three parts: an envelope that defines a framework for describing what is in a message and how to process it, a set of encoding rules for expressing instances of application-defined datatypes, and a convention for representing remote procedure calls and responses.

ebXML → a modular suite of standards for conducting electronic business. An initiative sponsored by the OASIS, a global consortium that drives the development, convergence and adoption of e-business standards.

ebRIM → Registry Information Model: specifies the abstract model for the objects and metadata contained in a ebXML registry.

UDDI → Universal Description, Discovery and Integration: like ebXML, but more restricted in scope than ebXML.

WSDL → Web Services Description Language: an XML-based language for describing Web services and how to access them.

XSLT → eXtensible Stylesheet Language Transformation: a language for transforming XML documents into other XML documents. Example, you can use FME to transform a ESRI Shape file into simple GML2 and then use XSLT to transform the data into GML3.

XPath → a language for finding information in an XML document (pointers from one part of the document to another).

JTS → Topology Suite: a Java library that implements in a robust way the OGC Simple Feature Specification. Developed in BC and Vivid Solutions is used around the world. Galdos, Safe Software, and ESRI have all used JTS!

JUMP → Unified Mapping Platform: a Java GUI-based application for viewing and processing spatial data, providing a highly extensible framework for the development and execution of custom spatial data processing applications.

uDig → Like JUMP but based on IBM's Eclipse Java framework. Currently supports a GML2 WFS client.

FME → Feature Manipulation Engine: Is used for spatial data interoperability by transforming one format into another.

GeoTools → An open source Java GIS toolkit for developing standards compliant solutions. It provides an implementation of OGC specifications as they are developed.

GeoServer → An open source implementation of a WFS server. Currently limited to GML2.

MapServer → The most successful open source implementation of a WMS. Very widely used including at the Ministry.

IMF → Internet Mapping Framework: A Java framework for developing Web mapping applications. Supports both OGC open specifications and the ESRI ArcIMS map server.