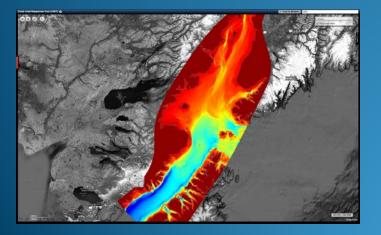






Holistic Cyber-Infrastructure for Ocean Data





Rob Bochenek Information Architect

Summary

- Background Exxon Valdez Oil Spill Restoration and Monitoring Program
- Lessons Learned
- System Data Flows
- Architecture and Infrastructure
- Demonstrate Data Management Tools
- Demonstrate Data Integration Tools

Background- Exxon Valdez (2002-2006)

- 1 Billion \$ restoration/research program started 1989
- 250 Million \$ Gulf Ecosystem Monitoring Program (GEM)
- Interact intimately with high profile data managers, administrators and scientists
 - Phil Mundy, Steve Hankin and Peter Cornillon
 - Access to Charles Falkenberg files/memos
- "Great Incubator"

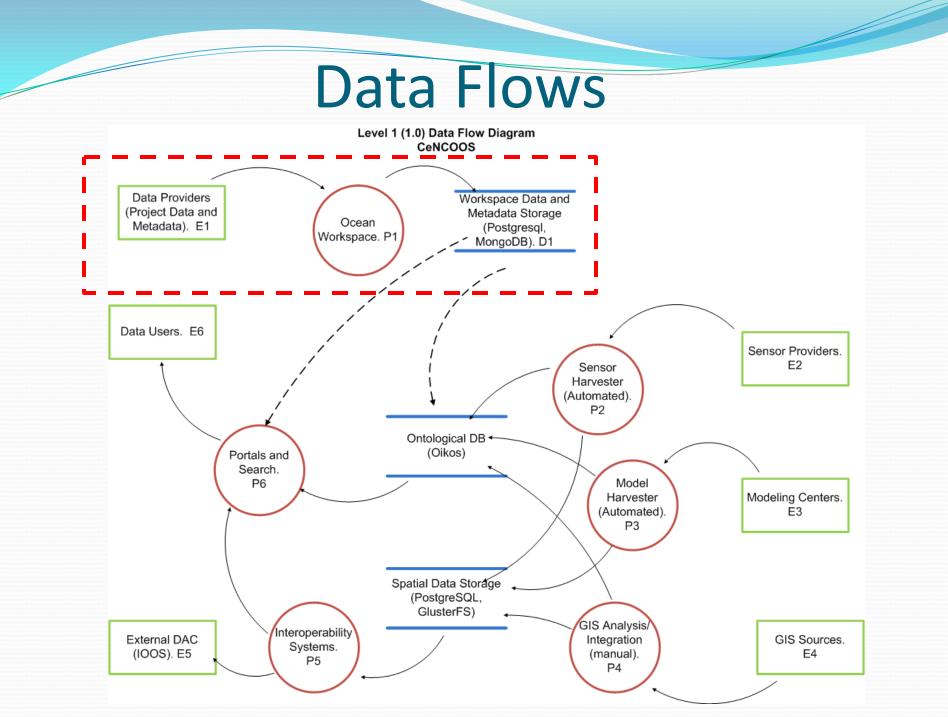


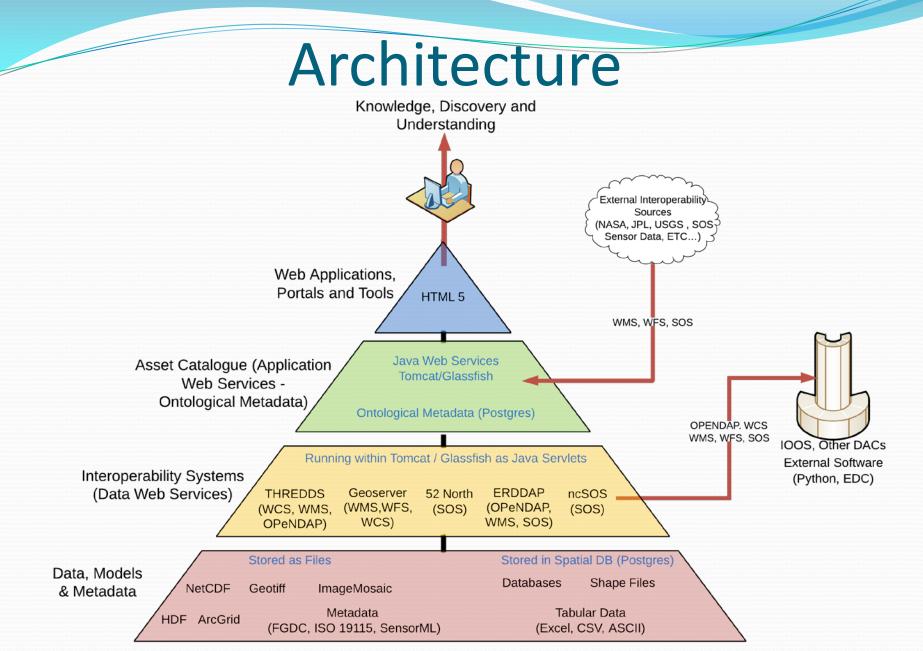
Lessons Learned

- Initiate DM activities at the beginning of projects not after
- Standards and requirements are not enough. Scientists need tools to effectively execute on their data management requirements
- Data integration and visualization are important for discoverability, assessment, re-use and management applications
- Two types of data (Structured/Machine Produced and Research Campaign Based). Each require unique approaches

Research Campaigns

- Researchers
 - Need ways to securely share data and information products between study teams
 - Need tools to generate metadata and publish data to meet DM requirements
- Data Managers
 - Develop DM protocols in vacuum using own methods and systems
 - Difficult to execute job effectively (not very satisfying) resulting in staff turnover and other problems....
- Program Managers
 - Want more transparency to entire process





Vector Spatial Data, Databases/Tabular Formats

Infrastructure

- Axiom operates two High Performance Computing Clusters (Portland, OR and Providence, RI) for data system geo-replication and high availability
- On the fly visualization and analysis of complex ocean data requires sophisticated hardware approach
- HPC provides horizontally scaling framework for compute, storage and bandwidth
- Shared infrastructure between NOAA Clients



Research Workspace: Scientific Collaboration and Data Management Platform



- Researchers organize themselves into teams for projects and larger scale research campaigns
- Data, sampling designs, contextual information, analytical workflows and results can be securely shared and transformed among team members
- Users can generate scientific metadata for information resources (ISO 19115-1/2)
- Users can then elect their project and selected data files to be published to publicly accessible portals.

