Policy Analysts: Data Policy Breakout

Approach of Breakout

- Worked from IWGDD points first, amplified from EPA points.
- General affirmation of policy elements

Amplifications

- Affirm roles and responsibilities as key
  - Needs to particularly address recognition
- Look for ways to have policy elements that can be pulled together to meet agency and laboratory needs (building blocks).
  - Harmonization across agencies and within agencies for data policies is highly desirable.
- At an agency level, policy describes outcomes, not details of procedures or implementation.

Amplifications

- Data is an asset (enterprise resource) and needs to be managed as such
- Not all data needs to be kept. Policy needs to be clear about how those decisions are made
  - Benefits may be to users in other agencies, localities, or even countries.
  - Policy governs the level of investment process decisions (data access tools, metadata, data quality objectives …)
  - Deal with “negative value” wording.

Amplifications

- Policy can balance access versus control
- Data is part of the information age and generating value for what's been done and achieving science/technical mission (domain knowledge)
  - Chief Data Officer key to getting value for data/manage data asset
    - Real cost to agencies to have "chief data officer"
    - How can CDO relate to CIO (tends to be on HW/SW, business operations).
    - Data management is integrally related to Enterprise Architecture (enables or hinders)

Researchers: Data Policy Breakout

Develop a SDM Plan for full life cycle

- The agency shall require its programs and projects to develop data management plans that cover the full data life cycle (from planning for data creation to accessible archiving, disposition/preservation). Each agency shall develop procedures for compliance. The data management plan shall include methods for compliance. The plan must be reviewed on a regular basis. The agency shall have a reasonable number of templates that encourage consistency that cover the data management plans aligned with communities of practice.

Identify Scientific Data w/ Metadata

- We need metadata following accepted standards for content and format to assure certain conditions are met. At the agency level, need to cover certain facets:
  - discovery
  - access
  - reuse
- **Administrative metadata** (includes appropriately maintained tools for each version)
- **Preservation** (includes anything that is needed to allow access; preservation of the metadata goes along with data). Preserving bits, readability, understandability (documentation to explain what bits mean what).
- Note: Someone must maintain the tools that would be able to access the metadata and data to allow access in perpetuity.

Statement of Guiding Principles for Digital Scientific Data Preservation and Access

- The agency shall have processes, compliant with all relevant laws, to determine which data shall be preserved and accessible. Includes a method to determine which data should be preserved and made accessible. Note: Preservation and access policies should be addressed separately.

Science Managers: Data Policy Breakout

Develop a Scientific Data Management Plan that Covers the Full data Life Cycle

- Scientific data created by an agency is made publicly available in a timely manner (Barrier: need to think about what to release and who decides when to release – not realistic to release all data?)
- Data management should begin at the inception of project/effort and be an integral part of project planning, budgeting and management.

Supplement ~~Identify~~ Scientific Data with Metadata to Enable ~~Needed Business Operations~~ current use and future reuse

- The metadata should be linked directly with the data
- Technologies should be identified and implemented to enable linking and discovering data
- Ontologies should be established and/or identified and implemented to enable linking and discovering data
- Tools that create/analyze/model/access data may need to be retained in addition to output data to make the data understandable for future use/reuse.

Manage Scientific Data as Enterprise Assets ~~or Liabilities~~

- Important to manage as a resource that has an appraisal process (value)
  - Consider visibility, impact, uniqueness of data
  - Value of data should be appraised at a defined frequency (what is not valuable today could be valuable tomorrow)
  - Records management (there are schedules for disposition – ensure that policy complies with the law)

Manage Scientific Data for Appropriate Control

- Need to identify (or create when necessary) appropriate long-term repositories for scientific observation data

Statement of guiding principles for digital scientific data preservation and access

- No information provided

Issues that are not addressed…

- Need for continuously-funded data curation/stewardship function and staff within the agency to ensure continuity
- There should be an interagency ontology similar to FEA that each agency can map their respective ontologies to, illustrating the types of federal data assets across the government.

Issues that are not addressed…

- Tools that create/analyze/model/access data may need to be retained in addition to output data to make data understandable for future use/reuse.
- A policy does not necessarily ensure a cultural change in behavior.
- Incentivize and encourage data management planning at the appropriate level within the agency.
- Policy should include a glossary that clarifies the meaning of specific terms in the policy.
- We believe it is also useful to have meeting about available technologies for enabling scientific data management.

Operational Users: Data Policy Breakout

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Manage Scientific Data as a National Scientific Resource		• Strategic goals may change over time. • What happens if it is a "virtual" enterprise with multiple partners? • Where should the Scientific Data Officer sit? • "Enterprise" and "asset" have connotations.	• Suggested wording change.

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Develop SDM Plan that Covers the Full Data Life Cycle	• Make sure there is a glossary and that the plan is understandable by all	• Terminologies differ	

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Identify Data with Metadata to Enable Needed Business Operations	• Begin applying metadata early in the life cycle Periodic review/appraisal process would be a point where metadata could be changed • Community accepted metadata structures with extensions • Persistent unique identifiers for data	• Metadata needs may differ by user group, use and by near-term versus long-term use. Business operations may change over time. • Need to update or add metadata at the end of the project or at other points in the life cycle It is hard to anticipate secondary and tertiary uses.	• Address this in the data management Plan • Include data dictionaries and methods catalogs.

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Manage Scientific Data for Appropriate Control	• Address embargo and sunset		• Sunset may mean different things to different communities. Here is where we need the glossary.
Maintain Version Control and Change Control on Data Sets	• Connect to model version if applicable	• Unless frozen models change constantly	• Different computer environment may need to be noted in the metadata as well

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Retain Data Commensurate with Its Value	• Periodic review/appraisal process. • Consider whether data can be reproduced cheaper than the archiving.	• We really need to talk about cost. • Consider the value to the broader scientific enterprise as well as the agency enterprise Should conduct a risk assessment, the risks of keeping, doing nothing or of deleting.	• Establish a review committee with stakeholder representation.

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Ensure Scientific Data Management Practices Integrate with KM Initiatives	• Connect to budget, HR, and other KM areas.		

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Statement of Guiding Principles for Digital Scientific Data Preservation & Access	• User requirements driven	• You need to talk about both preservation and access “in the same breath”. Access is moot if you haven’t preserved. Dark vs. Light Archives	

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Assignment of Responsibilities		• Assign role for reference service for the data • Assign across the life cycle and improve interactions between the players.	• Organization may need to be found for high value but orphaned datasets • Consider how to keep the metadata up-to-date when someone leaves or the organization changes.

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Description of Mechanisms for Access to Specialized Data Policies		• This particular element isn't clear.	• Need to reword, but didn't have time.
Statement of Intention for Cooperation, Coordination and Partnerships	• Submission guidelines & agreements • Affiliated archive status with NARA		• NOAA's agreement might be exemplar

Policy Elements and Principles	Best Practices	Issues & Challenges	Solutions
Provisions for Updating and Revisions of the Policy Document (Living Document)		• Drivers may include changes or new Administrative or Congressional policies directives. Could also be driven by OGC decisions or Administrator directives.	• IWGDD considers this a living document. This is at such a high level, do we really consider this to be changing often?
Data Rescue Policy			

END