

[Progress in Data Management Planning](#) [1]

Submitted by NancyRitchey on Mon, 2014-10-06 16:22 Tuesday, January 6, 2015 - 13:30 to 15:00

Event: [Winter Meeting 2015](#) [2]

Session Type: [Breakout](#) [3]

Expertise Level: [Beginner](#) [4]

Collaboration Area: [Data Management Training](#) [5]

Abstract/Agenda:

Recent focus on data initiatives and policies from the government, funding agencies and publishers has highlighted the importance of Data Management Planning. How have we used this increased attention to improve our Data Management Planning policies, approaches, and tools? The goal of this session is sharing leading practices, approaches and tools that will further improve Data Management Planning across the Earth Science Data Partners. Speakers from USGS, USDA, NASA and NOAA will present their organization's perspective on this topic followed by a panel discussion addressing challenges, gaps and potential solutions.

Notes:

Session on data management planning.

Erin's idea - saw a presentation at a NOAA meeting and ask them to give a similar session at the ESIP meeting.

The ESIP short course pamphlets were highlighted - to teach data management planning. Ruth said it was 35 peer reviewed videos with associated presentations. Available for reuse and remix.

The meeting is to talk about progress. Where do we want to go in the future. Representatives will talk about their progress in data management planning.

First talk

Jeff de La Beaujardiere from NOAA

Talked about traditional planning activities. Mission or project specific.

In 2011, they issued a data management planning directive. All data in NOAA should have a data management plan.

Several groups (OAR and OER) used this template extensively.

A revision/complete re-write was created and is effective 2015-1-1. The committee just voted on it, and it is available online.

Reviewed some of the procedural directives. A new one is on data citations for example. Assigning DOIs to archival data sets. And a data access directive - must make data accessible or get a waiver to not make it available.

Clear requirements and responsibilities - they are now clear and crisp and easier to understand.

If you are the manager - you are responsible for the data your program made. you can delegate it, but someone senior has to be in charge of it.

Improved data management plan template. Publicly available and made general so others may use it.

They also fixed redundant or multipart questions, so that it can be an online form.

Categories in revised template include asking if you have provided funding for data management. If you are sending your data to an archive, and when. How do you maintain it before you send it out. And other questions as needed.

Challenges - awareness, willingness, metrics.

Question - Bob Downs. Asked about accommodating collections in a plan, instead of just a data set? Jeff - you could do that if you wanted to. Or you could have a more specific plan. It is one approach you are now allowed to take.

Second presentation

NOAA Environmental Data Management: Improving Satellite Data Archive Planning

Helen Wood

Overview - discussed nature of the data being collected. AND it is well described. But you also need storage systems, archives are more than a storage system though.

Starting two new satellite programs, and need to think about the future, next generation.

Data management plans are thick. Data centers are preparing to hold these data, and need to know for how long. And what it means to have a complete plan. Not just because you have one means it is complete. But readiness to hold data is being done differently/different time than readiness to collect data.

Principles for effective environmental data management (recommendations from 2007). There are 9 principles that are rock solid

1. archived/accessible
2. end-to-end management
3. involve users
4. partnerships
5. metadata
6. expert stewards
7. what data to archive
8. discovery, access, and integration
9. planning process

Set of rules based on these 9 recommendations. And are written for the top level, which can provide

support to the lower levels.

Data Lifecycle - Data Management Activities, Usage Activities, Planning and Production Activities etc. Not just delivering data, but the entire life of the data.

Review of the template Jeff showed - which includes resources for managing data for the long term.

Approach - need a policy, has a long term archive, and make extensible, and make sense where we live and work. And then apply that to the systems we are building today.

Discussed examples that did not include ways of pulling out primitive data and ensuring it was still accessible. Missed opportunities.

Questions? Someone asked about the 9 principles - wanted to write them down. The report is available on the NRC website, for free. www.nap.edu/catalog.php?record_id=12017 [6]

Third speaker

Ted Payne

USDA Geodata Management Planning (GDMP)

Geospatial data.

Instead of accessibility - usability. So you don't have to build a new system to use existing data.

Discussed where they fall in the hierarchy. This includes working with the OSTP, OMB, GAO, and FGDC.

Program manager for 17 agencies, to get them to develop a working plan.

Overarching drivers within department of Ag.

Want to align with previous efforts, data structures. FGDC and NSDI. OMB has a number of drivers as well.

Be transparent.

GDMP USDA drivers

oversight and reporting

Stronger portfolio management understanding - what missions, what functions in those missions, and what supports those (alignment exercise). Now, during sequestration, they can create these narratives. How to work with ESRI, positioning one's self to negotiate with these organizations. How many people have accessed the information. And the amount of licenses that are out there etc.

Three principles: Discoverability, Accessibility, Usability

Challenges

1. Discoverability

1. GAO, OMB study to use CKAN

2. Confusing set of messages to your community

3. Pushback, and say should have better guidelines

2. Accessibility and usability

1. Strong language telling where to go

2. stronger communication pattern internal to USGA

Enterprise geodata maturity

Data management policy, have not started operations, but moving in the right direction. Important slide to show higher ups.

Open discussion?

4th talk

Viv Hutchison

Planning for data management for the USGS

Mixture of people with library degrees. - Science Data Management.

USGS - 7 different mission areas (Core Science Systems is one). Distributed all over the place.

ODI - Open data initiatives: 2013

Working on data management type information. And these were helpful to back up what they have been working on for the USGS.

Science data lifecycle model - had to build their own, as none were exactly appropriate. published if interested. Emerging new policies for DM - Mid January release. Instructional memos then policies. 4 coming out. One is the life cycle. One about metadata, one about releasing data, and one about preservation of data.

A closer look - unreleased policy for planning.

USGS data management website

www.usgs.gov/datamanagement [7]

Data management planning vs. program

Planning tools - multiple options for planning in sciencebase.

Data management planning process at the USGS

Milestones - 2011 WG developed a planning template across mission areas. One fill out before getting funding, and the other if you get funding. NCCWSC

Example of the two phase process looks like.

NCCWSC - reviews the plan with a data steward, provides comments back to a PI. If funded, then

they work with the Data Steward to complete the plan. And to make sure the data goes back into the repository/ScienceBase.

Trying to learn lessons, and promote adoption elsewhere in the USGS.

Challenges - Education, follow-through, and next steps (linking plans to metadata).

Questions - Ruth had a comment about linkages... There is an RDA interest group on this topic. Living data management plans.

Donald - within NOAA, we have some of these connections. One of the exports is the ISO metadata file. It is already mapped in to a standardized way. Have you thought about that as an output here?
Viv - Yes. We are working on that. We have build a dashboard, and we want to take that concept and expand it to cover this so it is more of a data management dashboard.

Wants to have a data steward in every single program!

Fifth talk

Ruth Duerr

NSIDC

Data management planning at NSIDC

NSIDC is not a large federal agency trying to be uniform.

Overview of what NSIDC is - co-opership with NOAA and University of Colorado Boulder.

NSIDC informatics. Technology and data management. And to help NSIDC meet its mission.

Way oversimplification diagram in slides. How they relate and getting them herded together. Smaller bits into something bigger than the parts.

Data management planning - because funded by multiple agencies, what they can do varies dramatically.

NSF - you have to put data in ACADIS but isn't paid for.

ELOKA - might manage the data, but who ones it? The knowledge holder.

DAAC - has assigned data that they are required to deal with, and NASA is pretty good at data management. But also get requests for data to be put in by investigators who think the DAAC should pay for it. Have an accession process for that type of work.

Trying to work to move the whole data management thing back in to the research process itself. Not just writing the plan, but tools for the field which can help organize data, and make ready to go in to the archive. The Data Conservancy packaging tool helps with that.

Willing to send data managers into the field during data creation to help understand. Very useful.

Data management is a low priority for domain scientists, but also want others data to be accessible. BIG PROBLEM.

Terminology problem - data management plans - researchers; and data curation plan - repositories.

Question - as a researcher, it is not sufficient for a researcher to just identify an archive. Have to look at if it will handle the data, format migration etc.

Progress in Data Management Planning

Published on Commons (<http://commons.esipfed.org>)

Bob - follow up. Thinks from an appraisal perspective, we wouldn't know if someones data is worthy until after we see the results of their work. So data management plan specifying in advance where they are going to send the data, is assuming a lot.

Joe - data management plan vs curation. Plan happens once and gets ignored. Curation....

Question - is it common for researchers to contact the repository before hand? Ruth - it depends on the funding agency, and even in NSF, the program.

Session Leads:

Name: [Nancy Ritchey](#) [8]
Organization(s): [NOAA/NCDC](#) [9]
Email: nancy.ritchey@noaa.gov [10]

Notes takers:

Name: [Sarah Ramdeen](#) [11]
Organization(s): [School of Information and Library Science UNC-CH](#) [12]
Email: ramdeen@email.unc.edu [13]

Creative Common License: Creative Commons Attribution 3.0 License

Teaser: Sharing leading practices, approaches and tools that will further improve Data Management Planning across the Earth Science Data Partners

Accepted:

Keywords: [data management plans \(DMP\)](#) [14]

Source URL: <http://commons.esipfed.org/node/2697>

Links:

[1] <http://commons.esipfed.org/node/2697>

[2] <http://commons.esipfed.org/2015WinterMeeting>

[3] <http://commons.esipfed.org/session-type/breakout>

[4] <http://commons.esipfed.org/taxonomy/term/260>

[5] <http://commons.esipfed.org/collaboration-area/data-management-training>

[6] http://www.nap.edu/catalog.php?record_id=12017

[7] <http://www.usgs.gov/datamanagement>

[8] <http://commons.esipfed.org/node/2696>

[9] <http://commons.esipfed.org/taxonomy/term/1492>

[10] <mailto:nancy.ritchey@noaa.gov>

[11] <http://commons.esipfed.org/node/557>

[12] <http://commons.esipfed.org/taxonomy/term/373>

[13] <mailto:ramdeen@email.unc.edu>

[14] <http://commons.esipfed.org/taxonomy/term/1302>