Data Management Services from the Digital Asset Services Hub (DASH)

Sophie Hou
hou@ucar.edu
Data Curation & Stewardship Coordinator

HAO Staff Meeting
June 27th, 2017
How do you feel about DMPs?
Agenda

• Introduction to the Digital Asset Services Hub (DASH)
• Background to DASH: the Data Stewardship Engineering Team (DSET)
• Current DASH Data Management Services
• Roadmap to DASH’s Additional Services
• Questions and Feedback?
Introduction
Introduction to the DASH

• DASH: Digital Asset Services Hub
  • https://www2.cisl.ucar.edu/dash
  • Examples of digital assets: datasets, publications, models and software.

• A centralized system that is:
  • Dedicated to provide support, engagement, and training for UCAR/NCAR's research outputs and assets.
  • Focused on sharing and allowing access and use by the broader scientific community.
    • Guidance and oversight provided by the Data Stewardship Engineering Team (DSET).
Digital Asset Services Hub (DASH)

Discovery . Access . Use . Data Services

The Digital Asset Services Hub (DASH) is dedicated to provide support, engagement, and training for digital assets from NCAR and UCAR Community Programs (UCP), including datasets, publications, software, and models.

The services and resources made available through DASH focus on supporting the digital assets in order to make them available to the broader scientific community. DASH is created and maintained by the Data Stewardship Engineering Team (DSET).

DASH Services & Resources

There are currently six DASH services & resources areas that are under development.

Please choose from the following options:

- Learn More about Data Management Plan and Related Services
- Get In-Person, One-on-One Help
- Find Answers to Questions Shared by NCAR/UCP Community
- Explore and Locate Available UCAR/NCAR Digital Resources
- Provide Public Access to Your Digital Assets
- Review Community-Recommended Solutions to Manage, Process, Analyze, and Visualize Digital Assets

Accessible from: https://www2.cisl.ucar.edu/dash
A Video Introduction to DASH

Introduction - Digital Asset Services Hub (DASH)

Digital Asset Services Hub (DASH)
Discovery, Access, Use, Data Services

The Digital Asset Services Hub (DASH) is dedicated to provide support, engagement, and training for digital assets from NCAR and UCAR Community Programs (UCP), including datasets, publications, software, and models.

The services and resources made available through DASH focus on supporting the digital assets in order to make them available to the broader scientific community. DASH is created and maintained by the Data Stewardship Engineering Team (DSET).

DASH Services & Resources
There are currently six DASH services and resources areas that are under development.

Please choose from the following options:

- Learn More about Data Management Plan and Related Services
- Get In-Person, One-on-One Help
- Find Answers to Questions Shared by NCAR/UCP Community
Background to DASH
Background to DASH: The Data Stewardship Engineering Team (DSET)

• Created by the NCAR Directorate and relies on cross-organization participation to lead efforts for providing user-focused services.
  • Current members represent each NCAR Lab, the NCAR Library, and UCAR Community Programs.

• Support efficient and effective access and discovery to all community-digital assets by a wide range of users.
  • Initial effort is the development and operation of an integrated, web-based system.
Left to Right:
Tor Mohling (Research Applications Laboratory, RAL), Linda Cully (Earth Observing Laboratory, EOL), Don Stott (Earth Observing Laboratory, EOL), Sophie Hou (Computational & Information Systems Lab, CISL), Abby Jaye (Mesoscale & Microscale Meteorology Laboratory, MMM), Daniel Ziskin (Atmospheric Chemistry Observations & Modeling, ACOM), Nathan Hook (Computational & Information Systems Lab, CISL), Rebecca Centeno Elliot (High Altitude Observatory, HAO), Matthew Mayernik (NCAR Library), Louisa Emmons (Atmospheric Chemistry Observations & Modeling, ACOM), David Schneider (Climate & Global Dynamics, CGD), Ryan May (UCAR Community Programs, Unidata), Steven Worley (Computational & Information Systems Lab, CISL).

Not shown: Brian Bonnlander (Computational & Information Systems Lab, CISL), Christy Grant (Computational & Information Systems Lab, CISL), Don Kolinski (High Altitude Observatory, HAO), Eric Nienhouse (Computational & Information Systems Lab, CISL).

(https://ncar.ucar.edu/data-stewardship-engineering-team-dset)
Current DASH
Data Management Services
Data Management Guidance

Training and Education Materials & Best Practices

Categories:

- Data Management Plans (DMPs)
- Digital Object Identifiers (DOIs)
- Data Management Training and Education Resources

Data Management Plans (DMPs)

In order to assist and support the NCAR community in meeting the Data Management Plan (DMP) requirements when submitting proposals to potential funders/sponsors and fulfilling the DMP requirements after the proposals are awarded, the following resources are currently available:

- DMP Preparation Guidance and Template Document
- Samples:
  - Reanalysis Dataset
  - Community Earth System Model (CESM)
  - Simulations
  - Observations and Simulations
  - Measurements and Instruments
- DMP Checklist for Awarded Proposals
DMP Preparation Guide and Template

• **Nine major sections for the DMP:**
  1. Products of the Research (Type of Data Produced)
  2. Data Format
  3. Metadata
  4. Access to Data and Data Sharing Practices and Policies
  5. Policies for Re-Use, Re-Distribution, and Production of Derivatives
  6. Archiving of Data (Data Storage and Preservation of Access)
  7. Cost of Implementing the DMP
  8. Roles and Responsibilities
  9. Additional Comments

• Each section includes a description of the section and questions to consider for the section.
3. **Metadata**

- **Description:** Describe the format and standard in which the data/products will be described(documented).

- **Questions to Consider:**
  - What form of description/documentation, or metadata, will be used to describe the data/products produced (i.e. what details will someone else need in order to be able to understand, interpret, and use the data/products)?
    - If you are using metadata standard(s) (e.g. ISO19115, Dublin Core, EML, etc.), which metadata standard(s) are you using and why have you chosen it(them)?
  - What is the essential information needed to understand, interpret, and use the data/products from this research/project?
    - Metadata may include descriptions of research details such as: experiments, apparatuses, computational codes, etc.
  - What contextual details are needed to make the data/products you create or collect meaningful?
  - How will you create or capture these metadata?
  - When existing standards are absent or deemed inadequate, please document the rationale(s) and provide any proposed solutions or remedies.
DMP Preparation Guide and Template - Continued

• Designed to be an educational/reference resource.

• Reviewing the document and the actual DMP with the Data Curation & Stewardship Coordinator could help in understanding the document and applying it to meet specific proposal requirements.

• It is an evolving document, and we welcome comments and feedback to help in improving the content.
DMP Checklist for Awarded Proposals

• **Five** areas to consider:
  1. Maintaining the DMP during the Project Lifecycle
  2. Acquiring the Supporting Resources for Data Management
  3. Documenting Project Process/Workflow and Products
  4. Establishing a Relationship with a Long Term Archive/Repository for the Final Project Products
  5. Working with the Archive/Repository to Share and Provide Access to the Final Project Products

• Each section includes a description of the section and a list of recommended activities to complete.

• Questions or requests for additional information can be sent directly to the Data Curation & Stewardship Coordinator or via [datahelp@ucar.edu](mailto:datahelp@ucar.edu)
Current DASH
Data Management Services

• **Five** components are currently available to help UCAR/NCAR project teams in meeting data management plan (DMP) requirements:
  • Preparation Guidance.
  • Template.
  • Samples.
  • Frequently Asked Questions.
  • DMP Checklist for Awarded Proposals.

• In person consultation and workshop/training seminar can also be requested.
  • With the Data Curation & Stewardship Coordinator (me!).
Other Data Management Services

- Reviewing data management plans before submissions.
- Preparing descriptions and documentations (metadata) for the project and its results.
- Setting up a data management schedule and the responsible roles.
- Assigning Digital Object Identifiers (DOI).
- Selecting an archive/or a repository for depositing project outputs.
- Sharing news and progress regarding curation and stewardship efforts for digital assets.
Roadmap to DASH’s Additional Services
Roadmap to DASH’s Additional Services

• Other services that are currently being developed:
  • Additional training DMP training and education materials.
  • DASH Search and Discovery.
  • Getting Assets into DASH.
  • Software and Tools resources.

• Provide complementary assistance during Proposal and Award process from the DMP’s perspective:
  • Setting up reminders and milestones in preparation for project reports.
  • Updating status and progress.
  • Documenting accomplishments.
Search & Discovery

308 datasets found for "atmospheric pressure"

ERA-40 Pressure Level Forecasts
The pressure level forecasts from the 6-hourly runs of ECMWF ERA-40 reanalysis project are available here.

Extended Reconstructed Sea Level Pressure (ERSLP)
The extended reconstructed sea level pressure (ERSLP) analysis was constructed using the most recently available Comprehensive Ocean-Atmosphere Data Set (COADS) SLP data.

The International Surface Pressure Databank version 3
This dataset contains the International Surface Pressure Databank version 3.2.0 (ISPDv3), the world's largest collection of pressure observations. It has been gathered...

International Surface Pressure Databank (ISPDv2)
The International Surface Pressure Databank (ISPD; Cram et al. 2015) is the world's largest collection of pressure observations. It has been gathered through international...
# Metadata Entry Tool

<table>
<thead>
<tr>
<th>NCAR Dialect Element Name</th>
<th>NMDE.dit Path</th>
<th>Element Definition</th>
<th>Format Constraint</th>
<th>Subelement? (Yes or No)</th>
<th>Controlled Vocabularies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Asset Type</td>
<td>Asset Type</td>
<td>Type of asset described by this metadata</td>
<td>Controlled Vocabulary</td>
<td>No</td>
<td>See the pull down menu for dataset</td>
</tr>
<tr>
<td>2 - Author</td>
<td>Cited responsible party (under Core metadata - Data identification - Citation - CI_Citation)</td>
<td>The person(s)/institution(s) receiving credit, as in a citation</td>
<td>Controlled Vocabulary</td>
<td>Yes - Individual name, Organisation name, Position name, Role</td>
<td>Choose &quot;author&quot; as the value for the subelement &quot;Role&quot;</td>
</tr>
<tr>
<td>3 - Title</td>
<td>Title (under Core metadata - Data identification - Citation - CI_Citation)</td>
<td>A name given to the asset or the name by which the cited asset is known.</td>
<td>None</td>
<td>No</td>
<td>None</td>
</tr>
<tr>
<td>4 - Resource Type</td>
<td>Descriptive keywords (under Core metadata - Data identification - Descriptive keywords)</td>
<td>Type of asset</td>
<td>Controlled Vocabulary</td>
<td>Yes - Keyword, Type, Thesaurus name (additional subelements are included)</td>
<td>See the pull down menu for Model</td>
</tr>
<tr>
<td>5 - Location (URL)</td>
<td>Location (URL)</td>
<td>Web accessible landing page</td>
<td>URL</td>
<td>No</td>
<td>None</td>
</tr>
<tr>
<td>6 - Publication Date</td>
<td>Date (under Core metadata - Data identification - Citation - CI_Citation)</td>
<td>Date when the asset was first made available</td>
<td>DateTime</td>
<td>Yes - Date, Date Type</td>
<td>Choose &quot;Publication&quot; as the value for the subelement &quot;Date Type&quot; 2014-08-01</td>
</tr>
<tr>
<td>7 - Metadata Date</td>
<td>Metadata date</td>
<td>Date stamp when metadata record was created or last updated</td>
<td>DateTime</td>
<td>No</td>
<td>None 2014-09-20</td>
</tr>
<tr>
<td>8 - Publisher</td>
<td>Cited responsible party (under Core metadata - Data identification - Citation - CI_Citation)</td>
<td>The organization (or group) that made the asset available.</td>
<td>Controlled Vocabulary</td>
<td>Yes - Individual name, Organisation name, Position name, Role</td>
<td>Choose &quot;publisher&quot; as the value for the subelement &quot;Role&quot;</td>
</tr>
<tr>
<td>9 - Resource Support</td>
<td>Resource support contact (under Core metadata - Data identification)</td>
<td>Person, group, or institution to contact for support on asset</td>
<td>Controlled Vocabulary</td>
<td>Yes - Individual name, Organisation name, Position name, Role</td>
<td>Choose &quot;pointOfContact&quot; as the value for the subelement &quot;Role&quot;</td>
</tr>
</tbody>
</table>

### Metadata Entry Tool Details

- **Dataset:** See the pull down menu for dataset
- **Author:** Choose "author" as the value for the subelement "Role"
- **Model:** NCAR Global Climate Assimilation (CFDDA)
- **Publication Date:** Choose "Publication" as the value for the subelement "Date Type" 2014-08-01
- **Metadata Date:** None 2014-09-20
- **Publisher:** UCAR/NCAR - Comp Laboratory
- **Resource Support:** Choose "pointOfContact" as the value for the subelement "Role" 2014-08-01
- **Contact:** John Plain, CISL, As...
Currently, an organization wide repository does not exist.
Benefits of DASH’s Recommendations and Services

Credit/Citation/Recognition

Transparency

Understandability

Discoverability/Accessibility

Re-use/Reproducibility/Validation

Compliance
Thank You!
Questions? Comments?

Sophie Hou
(datahelp@ucar.edu, hou@ucar.edu)