Introduction

Many NASA Earth Observing System (EOS) have either already reached the end of their active life or are nearing it. Preservation of data products is a fairly well defined task for the NASA EOS Data Centers or DAACs. However, supporting documentation and other artifacts from these missions are also critical to the long-term studies of our planet’s climate, and to aid future generation’s ability to understand climatic changes. The challenge is how to preserve these items along with the traditional data products.

The Goddard Earth Sciences Data and Information Services Center (GES-DISC) has implemented a repository system, which is capable of long-term archive of documentation artifacts and other associated digital content. The existing GES-DISC Repository System is based on Fedora Commons, an open-source repository management software, for asset savings and flexibility.

The first mission to utilize the GES-DISC Repository System was the High Resolution Dynamics Limb Sounder (HIRDLS) on the Aura spacecraft. Since then, the GES DISC has gathered documentation from the UARS and TOMS into the Repository. The Microwave Limb Sounder (MLS) team has begun delivering some early pre-launch documents to the GES-DISC Repository System as well. Other missions in planning or progress include AIRS, OMI, SORCE, SNPP, Sounder, and TRMM.

NASA Earth Science Data Preservation Content Specification 432-SPEC-001

Being able to understand and interpret data from these older missions after the experts familiar with them have moved on or are no longer available is a concern to NASA. The importance of preserving the Earth science documentation and data resulted in the issuance of the “NASA Earth Science Data Preservation Content Specification” (423-SPEC-001) for the Earth Science Data and Information System (ESDIS) supported data centers. Documents and data in the GES-DISC repository are archived and classified according to 423-SPEC-001 into 9 categories listed in the table below for each mission.

Available at https://earthdata.nasa.gov/files/NASA_ESD_Preservation_SPEC.pdf

Fedora Commons Interface

The GES-DISC uses Fedora Commons, an open-source repository management system that is used in many universities, research centers, and libraries. It comes with a simple web-based GUI interface which provides for easy administration of the system. The GUI also allows one to enter objects or datastreams (these can be of any type document, image, source code, binary data, etc.) into the system. The system uses XML to manage the objects. The GES-DISC has also developed a command line script to allow batch ingest of objects into the Fedora Repository.

Ges DISC Data Preservation Implementation

Overview of the physical objects storing, tagging, storage in archives and distribution system

Restricted documents are currently only available by contacting the GES-DISC user services.

1) Identify documentation
   GES DISC Science Support staff identified specific information needed permission in Data Preservation Mission List by working closely with the original mission teams to sort out documents for preservation

2) Specify and implement preservation environment
   a) Local archive based on open-source system Fedora Commons
   b) Implementation is complete for HIRDLS, TOMS and UARS missions
   c) Other missions MLS, OMI, AIRS, SORCE, TRMM, etc. are in progress

3) Retrieve documentation
   Public documents accessed by users from GES DISC web site

4) Implement retrieval and distribution services
   a) Access for internal GES-DISC users
   b) External Access via WWW for unrestricted (public) documents
   c) External Access for restricted documents (ITAR) via User Services contact
   d) Iterate with other DAACs/community

Public Access of Preservation Documents

External users access the publicly available documents by visiting the mission specific documentation page for that instrument. The Fedora repository system is at the backend and makes access to the linked documents possible. Note that restricted objects (ITAR, proprietary, or software) are not accessible through the public interface. Four missions are now public:


Lessons Learned, Challenges, Future

- Heritage missions require extensive work to identify and classify documents
- Restricted (ITAR or Proprietary) requires special handling vs. Unrestricted
- Incorporate DOI metadata into repository (if available)
- Replace or upgrade Fedora Commons which is currently 5 years old
- Move Repository to a NASA ESDIS-wide Repository