

Repository Cross-Linking at the National Center for Atmospheric Research
Jennifer Phillips
Matthew Mayernik
NCAR Library

Abstract: The National Center for Atmospheric Research (NCAR) builds and maintains a number of repositories for data and scholarship. Using these as a development test bed, our project demonstrates how multiple repositories of diverse resources can exchange and connect related information via complementary workflows and metadata sharing. Our poster maps out how we are building cross-link between our data and scholarship repositories, on the one hand establishing relationships between resources upon submission by researchers and on the other establishing technical connections between repositories on which to build out future interoperability.

The National Center for Atmospheric Research (NCAR), a Federally-Funded Research and Development Center operated by the University Corporation for Atmospheric Research (UCAR) for the National Science Foundation (NSF), provides and manages a number of data repositories and well as an institutional repository for scholarship. The NCAR Computational and Information Systems Lab (CISL) and the NCAR Library received an NSF EAGER grant to produce a pilot implementation of repository cross-linking, using as the development test bed data and scholarship repositories provided and managed by NCAR/UCAR.

NCAR CISL built, maintains, and operates the Earth System Grid, which has served as the primary mechanism for the distribution of climate model results for international projects, and is the primary distribution system for many UCAR climate and environmental models and model results. The ESG is a grid-based infrastructure for the distribution and access of climate data. The NCAR Library operates the OpenSky repository, which hosts and provides access to the record of scholarship produced by NCAR and UCAR staff, including peer-reviewed publications as well as conference presentations and posters, reports and archival materials. OpenSky is in the process of transitioning from a locally developed repository platform to an Islandora/Fedora repository. The objective of our joint project is to demonstrate through a proof-of-concept implementation how multiple repositories of diverse resources can exchange and connect related information via complementary workflows and metadata sharing.

Our conference poster will map out our approach to establishing cross-links between the ESG and OpenSky. This will in part show how we establish workflows that allow researchers who deposit resources in one system to initiate the deposit of related resources in the other. So for example, if a researcher is publishing a report or depositing an item of scholarship in OpenSky, they will be able to initiate the deposit of data or software into the ESG. We will also show how links between the related resources held in the two separate systems will be exchanged and made visible in the respective systems. Producing and maintaining these cross-links will require the development of common workflows for identifying and depositing related resources, and the development of technical connections between the two repositories, leveraging APIs that enable metadata and/or file exchange. By developing such common workflows and information exchange protocols, our goal is to demonstrate how repository interconnections can be built in ways that establish connections between related resources (e.g. data, software, services, etc.) thereby providing even further benefit to users. While some of workflow development, protocol specifications, software and use cases are specific to NCAR/UCAR context, through this pilot project we work toward a longer term vision of developing processes for researchers to more easily deposit resources in relevant repositories, and for links between repositories to be created more routinely.