The IRIS Data Management Center: Enabling Access to Observational Time Series Spanning Decades

T. Ahern (1), R. Benson (2), and C. Trabant (3)

(1) IRIS, Data Management Center, Seattle, United States (tim@iris.washington.edu), (2) IRIS, Data Management Center, Seattle, United States (rick@iris.washington.edu), (3) IRIS, Data Management Center, Seattle, United States (chad@iris.washington.edu)

The Incorporated Research Institutions for Seismology (IRIS) is funded by the National Science Foundation (NSF) to operate the facilities to generate, archive, and distribute seismological data to research communities in the United States and internationally. The IRIS Data Management System (DMS) is responsible for the ingestion, archiving, curation and distribution of these data. The IRIS Data Management Center (DMC) manages data from more than 100 permanent seismic networks, hundreds of temporary seismic deployments as well as data from other geophysical observing networks such as magnetotelluric sensors, ocean bottom sensors, superconducting gravimeters, strainmeters, surface meteorological measurements, and in-situ atmospheric pressure measurements. The IRIS DMC has data from more than 20 different types of sensors.

The IRIS DMC manages approximately 100 terabytes of primary observational data. These data are archived in multiple distributed storage systems that insure data availability independent of any single catastrophic failure. Storage systems include both RAID systems of greater than 100 terabytes as well as robotic tape robots of petabyte capacity. IRIS performs routine transcription of the data to new media and storage systems to insure the long-term viability of the scientific data. IRIS adheres to the OAIS Data Preservation Model in most cases.

The IRIS data model requires the availability of metadata describing the characteristics and geographic location of sensors before data can be fully archived. IRIS works with the International Federation of Digital Seismographic Networks (FDSN) in the definition and evolution of the metadata. The metadata insures that the data remain useful to both current and future generations of earth scientists. Curation of the metadata and time series is one of the most important activities at the IRIS DMC. Data analysts and an automated quality assurance system monitor the quality of the incoming data. This insures data are of acceptably high quality.

The formats and data structures used by the seismological community are esoteric. IRIS and its FDSN partners are developing web services that can transform the data holdings to structures that are more easily used by broader scientific communities. For instance, atmospheric scientists are interested in using global observations of microbarograph data but that community does not understand the methods of applying instrument corrections to the observations. Web processing services under development at IRIS will transform these data in a manner that allows direct use within such analysis tools as MATLAB® already in use by that community.

By continuing to develop web-service based methods of data discovery and access, IRIS is enabling broader access to its data holdings. We currently support data discovery using many of the Open Geospatial Consortium (OGC) web mapping services. We are involved in portal technologies to support data discovery and distribution for all data from the EarthScope project. We are working with computer scientists at several universities including the University of Washington as part of a DataNet proposal and we intend to enhance metadata, further develop ontologies, develop a Registry Service to aid in the discovery of data sets and services, and in general improve the semantic interoperability of the data managed at the IRIS DMC. Finally IRIS has been identified as one of four scientific organizations that the External Research Division of Microsoft wants to work with in the development of web services and specifically with the development of a scientific workflow engine.

More specific details of current and future developments at the IRIS DMC will be included in this presentation.