Evolution of Web Services in EOSDIS - Search and Order Metadata Registry (ECHO)

Andrew Mitchell  
NASA Goddard Space Flight Center  
Greenbelt, MD 20771, USA  
Andrew.E.Mitchell@nasa.gov

Hampapuram Ramapriyan  
NASA Goddard Space Flight Center  
Greenbelt, MD 20771, USA  
Hampapuram.K.Ramapriyan@nasa.gov

Dawn Lowe  
NASA Goddard Space Flight Center  
Greenbelt, MD 20771, USA  
Dawn.R.Lowe@nasa.gov

Abstract —

During 2005 through 2008, NASA defined and implemented a major evolutionary change in the Earth Observing system Data and Information System (EOSDIS) to modernize its capabilities. This implementation was based on a vision for 2015 developed during 2005. The “EOSDIS 2015 Vision” emphasizes increased end-to-end data system efficiency and operability; increased data usability; improved support for end users; and decreased operations costs. One key feature of the Evolution plan was achieving higher operational maturity (ingest, reconciliation, search and order, performance, error handling) for the NASA’s Earth Observing System Clearinghouse (ECHO). The ECHO system is an operational metadata registry through which the scientific community can easily discover and exchange NASA’s Earth science data and services. ECHO contains metadata for 2,726 data collections comprising over 87 million individual data granules and 34 million browse images, consisting of NASA’s EOSDIS Data Centers’ and the United States Geological Survey’s Landsat Project holdings.

ECHO stores metadata from a variety of science disciplines and domains, including Climate Variability and Change, Carbon Cycle and Ecosystems, Earth Surface and Interior, Atmospheric Composition, Weather, and Water and Energy Cycle. ECHO provides a platform for the publication, discovery, understanding and access to NASA’s Earth Observation resources (data, service and clients). In their native state, these data, service and client resources are not necessarily targeted for use beyond their original mission. However, with the proper interoperability mechanisms, users of these resources can expand their value, by accessing, combining and applying them in unforeseen ways. ECHO provides access to its capabilities through a set of services. These ECHO Applications Program Interfaces (APIs) are based on industry standards for performing web-based computing, specifically web services profile.

Keywords - Earth Science, ECHO, EOS, EOSDIS, Web Services, metadata, WSDL, SOA, SOAP

I. INTRODUCTION

EOSDIS provides access to Earth Observation data and services through multiple systems that support the management, storage, and discovery of, and access to the petabytes of information on those data. The ECHO framework has become the primary access point for cross-

Data Center search-and-order of EOSDIS and other Earth Science data holdings archived at the EOSDIS data centers.

ECHO is a middleware component based on a Service Oriented Architecture (SOA). The system is comprised of a set of infrastructure services that enable the fundamental SOA functions: publish, discover, and access Earth science resources. It also provides additional services such as user management, data access control, and order management. The ECHO system has a data registry and a services registry. The data registry enables organizations to publish EOS and other Earth-science related data holdings to a common metadata model. These holdings are described through metadata in terms of datasets (types of data) and granules (specific data items of those types). ECHO also supports browse images, which provide a visual representation of the data. The published metadata can be mapped to and from existing standards (e.g., FGDC, ISO 19115). With ECHO, users can find the metadata stored in the data registry and then access the data either directly online or through a brokered order to the data archive organization.

The ECHO service registry allows the Earth Observing community to publish information about, and access to their functional offerings. Users can discover these services that may have applicability to their goals thru the ECHO framework. Examples of services within the Earth Observation domain include subsetting functions, translation algorithms, models and visualization services, e.g. OGC WMS services.

The Web Services (WS) based interface was introduced to ECHO using WS APIs with the release of version 8 in October 2006. In determining this path forward, the impact of this change was initially assessed with our partner representatives and was also brought forward for comments and discussion at the technical community meetings. While there was some external impact, it was assessed that this was the best path forward for ECHO and EOSDIS Evolution activities. While the WS interface was easier to work with than the legacy XML, the legacy XML interface existed in parallel with the new WS API until the release of ECHO 9 in April 2007 after which only the WS API was retained.

This work was performed by the authors as part of their duties as U.S. Government employees. It was supported by the NASA’s Science Mission Directorate. The opinions expressed are those of the authors and do not necessarily reflect the official position of NASA.
The service interfaces are defined using Web Services Definition Language (WSDL). Access to the services is through the Simple Object Access Protocol (SOAP). The ECHO service registry conforms to the Universal Description and Discovery and Integration (UDDI) and leverages the ISO 19119[^3] standard for service categorization. Users can search, manage, and access the contents of ECHO’s data and services registries through user-developed and community-tailored interfaces or clients. Introducing this WS API based on industry standards has helped support broad, external client development in the ECHO system. Scientists and engineers can build interfaces using ECHO’s publicly available APIs or existing ECHO client components. Using these standards, clients written in most contemporary programming languages are isolated from the underlying technologies that support the distributed communication and functionality. These clients may call the ECHO Web Services much like a local function call. Most current developer tools support these standard technologies (e.g., WSDL, SOAP, WS-I Basic Profile 1.0) natively.

The remainder of this paper focuses on how EOSDIS via the ECHO framework is being used to foster interoperability in the EO community.

II. TECHNOLOGY CONCEPT

ECHO addresses the need of Earth Observation interoperability by providing an SOA platform infrastructure utilizing web service technology for the exchange of information and services, supporting the varied needs of a rapidly evolving global community.

One system driver in designing ECHO was to increase access to Earth science data and services by providing a system with a machine-to-machine interface. As such, an Application Programming Interface (API) interface is provided. There is typically a user focused client application interacting with ECHO’s API on behalf of an end-user. This client may be a generic, query and order based client or may be specific to an end user’s research, mission, or general area of interest. ECHO’s framework of components and a spatially enabled database function together as a clearinghouse and order broker for Earth Science metadata. ECHO also incorporates a Universal Description, Discovery, and Integration (UDDI) registry that facilitates registration, discovery, and invocation of services related to the ECHO holdings. Internally, ECHO specifies application programming interfaces (APIs) and provides middleware components (including data and service search and access functions) in a layered architecture.

ECHO allows data providers to cache copies of their metadata within the ECHO database. Data providers have complete control over what metadata are represented in ECHO on their behalf. They can insert new data, modify existing data and remove old data. Clients representing various entities can communicate with ECHO via its APIs in order to perform functions such as querying or ordering on ECHO’s holdings.

![Figure 1. ECHO Flow Diagram](image)

ECHO uses an open system approach and ensures that user interfaces fully address user/scientist needs by specifying and publishing domain APIs that accommodate independent ECHO clients. These APIs are independent of the underlying transport protocols used. All metadata are held in a database with spatial extensions. The metadata model is derived directly from that used by the Earth Observing System Data and Information System (EOSDIS) Core System (ECS).

ECHO’s service capabilities have been categorized into 6 basic types. Table 1 represents those types and provides an allocation of ECHO services to the categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Services</th>
</tr>
</thead>
</table>
| Participation   | Support user and organization participation.     | * Authentication  
                     |                                                 | * Provider  
                     |                                                 | * Group Management  
                     |                                                 | * User  |
| Publication     | Support publication of data and service resources. | * Ingest  
                     |                                                 | * Extended  
                     |                                                 | * Data Management  |
| Discovery       | Support finding of resources that have been published. | * Catalog  
                     |                                                 | * Extended  
                     |                                                 | * Taxonomy  |
| Ordering        | Support ordering and delivery of data.           | * Order Management  
                     |                                                 | * Order Processing  |
| Brokering       | Support facilitation of data and services integration. | * Invocation  
                     |                                                 | * Invocation Utility  |
| Eventing        | Support notification of dynamic events.           | * Subscription  
                     |                                                 | * Event Notification  |

Since ECHO uses platform independent web service definitions for its API there are no requirements for a client programming language. However, some languages lend themselves to web service based development better than others.

The EOSDIS Core System (ECS) data model[^1] is used as a basis for ECHO. ECHO incorporates the ECS concept of granules and collections. A granule is the smallest aggregation

[^1]: http://observer.gsfc.nasa.gov

This work was performed by the authors as part of their duties as U.S. Government employees. It was supported by the NASA’s Science Mission Directorate. The opinions expressed are those of the authors and do not necessarily reflect the official position of NASA.
of data that is independently managed (described, inventoried, and retrieved). A collection is a grouping of 0 or more granules that all come from the same source, such as a modeling group or institution. A collection could also be a logically related set of all granules such as all instances of a given version of a geophysical product derived from an EOS instrument’s data. Collections contain information common across all the granules they contain.

ECHO extensibility is ensured by its component architecture, which allows new capabilities and functions to be plugged in, modeling relationships between services/APIs/UIs, and continued prototyping. The ECHO system is developed incrementally to allow for insight and feedback during the development cycle.

III. DATA PROVIDERS & DATA HOLDINGS

ECHO data providers ingest metadata that represent their Earth Science data holdings to ECHO to make them available for search-and-order. These providers register their holdings in ECHO before publishing them for public viewing. Providers can declare a set of translation and validation rules which ECHO uses during ingest processing. The translation rules convert the XML metadata representation from the provider’s metadata model into the ECHO common metadata model. This process is used for data quality and integrity control.

All the EOSDIS Data Centers\(^6\) except the Crustal Dynamics Data and Information System (CDDIS) and the Ocean Biology Processing Group (OBPG) are currently ECHO Data providers.

IV. EXTENDED SERVICES

An ECHO Extended Service provider is an individual or organization that participates with ECHO by providing access to functions that create Earth Science Data or information. The services are "Extended" because they are not native to ECHO. ECHO advertises these services as well as brokers these services, as the service provider requests. ECHO provides interfaces and mechanisms that allow organizations to publish their services and correlate their service to the data types in the clearinghouse. These services allow providers to introduce and maintain web service interfaces, implementations, GUIs, and advertisements.

A web service interface defines a web service API including the messages and parameters necessary for implementing it represented in a Web Services Definition Language (WSDL) document. Service interface documents (.wsdl documents) are maintained by the ECHO system and managed by the ECHO operations team and the ECHO partner community. A web service implementation is a physical instantiation of a web service interface hosted by an ECHO partner. Like a web service interface, the service definition is also described by a WSDL document. An advertisement is used to provide broadcasting or publications about data. The advertisement services are associated with the data items that they publicized; they also provide descriptions of the broadcasted services. A web service GUI is a user interface for web services that may be registered in ECHO. Web Service GUIs are important because they provide a human with a way to interact with a web service.

V. CLIENT

An ECHO Client Partner works within the ECHO Community by developing a software application that communicates with web services available through the ECHO API. The ECHO API allows for various client types including user-interactive, metadata harvesting and other batch processing for scientists and other clients. Most ECHO clients provide access to ECHO’s Earth Science metadata and browse catalog and order broker, while other clients facilitate Data Partner data management and metrics collection. ECHO currently has 3 operational clients as well as multiple successful past clients and clients under development.

Culminating many years of service to the Earth Science community, the EOS Data Gateway (EDG) was officially retired on February 28th, 2009 and replaced by ECHO’s Warehouse Inventory Search Tool (WIST)\(^7\). WIST is the primary web-based client for discovering and ordering cross-discipline data from all of ECHO’s metadata holdings. WIST allows users, including those without specific knowledge of the data, to search science data holdings, retrieve high-level descriptions of data sets and detailed descriptions of the data inventory, view browse images, and submit orders via ECHO to the appropriate data providers.

\(^6\) http://nasadaacs.eos.nasa.gov/about.html
\(^7\) https://wist.echo.nasa.gov/api/

This work was performed by the authors as part of their duties as U.S. Government employees. It was supported by the NASA’s Science Mission Directorate. The opinions expressed are those of the authors and do not necessarily reflect the official position of NASA.
ECHO holdings. The EOLI Web Client\(^8\) provides access to ESA catalogues of EO products via a standard web browser, as well as to the catalogues of other data providers (for example DLR and NASA ECHO) and is part of ESA’s eoPortal.

The Search 'N Order Web Interface (SNOWI)\(^9\) tool is a web-based ECHO client used to search a select set of NSIDC data holdings. It is intended as a lightweight tool for quick access to data for its user community. SNOWI allows searching for and ordering data from NSIDC and other Distributed Active Archive Centers (DAACs). Unlike WIST, SNOWI only searches DAAC EOS data and Near Real-Time Ice and Snow Extent (NISE) data (plus MOD02, MOD09, and MOD35), and it does not offer subsetting.

There are several developmental projects underway to implement ECHO clients that target specific Earth science data disciplines. One such client is focusing on NASA Global Water and Energy Cycle and another is focusing on multi-sensor measurement-based model for long-duration analysis of climate variables.

Currently under development is an iGoogle\(^10\) gadget named NESI, NASA Earth Science Imagery (NESI) that will allow users to browse EO imagery generated from NASA data obtained via the ECHO architecture. NESI will allow users to submit requests for imagery based on spatial and temporal location and provide information associated with the event that image will portray. It is hoped that the user community will provide events that will showcase the ability of NASA earth science data to produce striking imagery.

In an attempt to showcase the cadre of capabilities available using ECHO web services, developers have been demonstrating new uses of ECHO. One such demo used ECHO Services to provide content for a 3rd party geospatial tool. Using the Google Earth\(^11\) application, the spatial extent of ECHO collections were displayed on a 3D representation of Earth.

In another widely applicable demonstration, ECHO in collaboration with the Open-source Project for a Network Data Access Protocol (OPeNDAP)\(^12\) team, developed a prototype for how to take advantage of the capabilities of ECHO and OPeNDAP data access services. The demonstration explained that a scientist needing satellite instrument data from two different data centers can retrieve it via an ECHO query through the ECHO web services and have a Matlab\(^13\) computing environment client invoke the search; access the result set URLs as well as access the satellite data. By leveraging these independent sets of services, clients can perform granule- or inventory-level queries and get direct references to support direct data access to those results.

VI. SUMMARY

ECHO has evolved over a number of years from a prototype to the current flexible and extensible operational system. The multi-organization, inter-disciplinary content of ECHO provides a valuable service to a growing number of Earth science applications and interdisciplinary research efforts. ECHO streamlines access to digital data and materials and brokers orders and other services from Earth Science applications to data providers. When data and/or service providers publish metadata representing their data holdings or services to the ECHO registry, they make their data and services resources available to a wide range of potential users. By building a client on ECHO’s service API and publishing the client as a service, client partners expose the availability of their clients to new users.

EOSDIS’s move to incorporate Web Services standards has increased client development in the ECHO system by reducing the complexity of integration and delivering platform and technology independence. Web Services’ broad applicability reduces the number of tools needed while leveraging Internet protocols and infrastructure. EOSDIS’s multi-organizational content in ECHO, the use of Web Service technology and a Service Oriented Architecture (SOA) provides a valuable service to a growing number of Earth science applications and interdisciplinary research efforts.

More information on ECHO can be found on the ECHO web site\(^14\).

---

\(^1\) http://catalogues.eoportal.org/eoli.html
\(^2\) http://nsidc.org/data/snowi/index.html
\(^3\) http://www.google.com/ig
\(^4\) http://earth.google.com/
\(^5\) http://opendap.org/
\(^6\) http://www.echo.nasa.gov/

This work was performed by the authors as part of their duties as U.S. Government employees. It was supported by the NASA’s Science Mission Directorate. The opinions expressed are those of the authors and do not necessarily reflect the official position of NASA.