ESIP Discovery Cluster Testbed: Validate and Relate Data & Services - Draft

ESIP Discovery Cluster
Christine White, Esri, cwhite@esri.com
Hook Hua, Jet Propulsion Lab, NASA, hook.hua@jpl.nasa.gov

Tags/Keywords
Stakeholders: ESIP Discovery Cluster, Data Providers, Service Providers
Themes: Discovery, Data Cast, Service Cast, OpenSearch, ATOM
Technologies: Esri Geoportal Server, OpenSearch, ATOM, REST, ESIP Discovery Profile

Overview

The ESIP Discovery Cluster identified pain points in data collaboration that included discovery of services available and the data for support. The initial vision of the Discovery Testbed addresses these pain points by supporting the following activities:

- Setup validation for registration of ESIP services
- Provide some form of a service cast of registered services
- Encourage the ESIP Community to actively contribute to the repository of data and services
- Chain together data and services - e.g., exploring data and services mapping, and brokering

The Cluster will leverage the Esri Geoportal Server software to act as the repository, validation mechanism, and service casting platform for these activities. Through developing strong use cases and requirements for these activities, the Esri Geoportal Server will be configured and customized. A prototype will be installed on the Discovery testbed, where ESIP members can upload their data and services to test and provide feedback. Additional suggested enhancements will be captured and can be assessed by the Cluster; because the Esri Geoportal Server is open source, additional enhancements can be taken on by Cluster member's post-project.

Problem Statement

The grand problem is described in the 'The Grand Challenge' section of the 'ESIP Discovery Whitepaper'[1]. A small piece of that larger challenge is that there is no central ESIP-accessible repository where ESIPI members can validate their services, and/or find out what is available. Other agencies have done some work on this front – e.g., The National Snow and Ice Data Center's Libre system[2], and JPL's SciFlo project[3]. This solution would not supercede these, but instead seeks to investigate ways to interface with them as appropriate.

Use Cases

The Discovery Cluster has already identified two use cases from which requirements will be developed. These two use cases are:

1. Service Validation[4] use case that describes the process of proposing a service to the geoportal and receiving feedback on if it is valid, if not why, and if so shall it be shared through the geoportal.
2. Query Aggregated Data and Services[5] use case that focuses on a user's ability to retrieve metadata about data and services advertised by user-registered casts and/or OpenSearch services that matches the user's search criteria.

Solution Options for Addressing the Problem

Although there are several catalog/service sharing technologies[6] freely available, the Discovery Cluster has selected the Esri Geoportal Server for this project. The Esri Geoportal Server provides out-of-the-box capabilities to register and/or harvest data and services for discovery, and access them through GeoRSS, JSON, ATOM, HTML, OpenSearch, and CSW protocols[7]. It also provides the capability to display relationships between datasets and services that have relationships defined in their metadata[8].

---

[1] http://wiki.esipfed.org/index.php/Discovery_White_Paper#The_Grand_Challenge
[2] http://nsidc.org/libre/share/collectioncaster.html
[3] http://sciflo.jpl.nasa.gov/scast/
[4] http://wiki.esipfed.org/index.php/ServiceValidation
[5] http://wiki.esipfed.org/index.php/Query_aggregated_data_and_services
[6] (TO DO – list of available options)
[7] http://sourceforge.net/apps/mediawiki/geoportal/index.php?title=How_to_Search_for_Resources
[8] http://sourceforge.net/apps/mediawiki/geoportal/index.php?title=Create_Relationships_between_Resources

The Esri Geoportal Server is configurable to support validation of services/datasets; the only stipulation is that the specifications are for use with XML-based protocols. For validation of services, the ESIP Discovery Cluster has identified two initial ESIP specifications – 'ESIP Discovery Cast Atom Response Format v1.1'[9] and 'Federated OpenSearch Conventions for ESIP'[10] – with additional interoperable specifications to come.

Benefits to ESIP

- Improved interoperability of ESIP Discovery services and clients leading to broader adoption.
- Increased awareness and usage of Discovery services.
- Improved brokering of data collection and granules to compatible services that can use them.

Remaining Challenges

The culminating activity to this project - chaining together data and services, and brokering – is the most challenging aspect. The Discovery Cluster has identified that although this is the end goal, it will likely not be attainable until there is a repository with valid services and data in existence, and from that repository the use cases for user interface and technical interface can be better defined. So not only are there prerequisite things that must happen first, but the changes to the geoportal – or other component technology - to facilitate an efficient chaining interface must still be designed. This is an area for future work.

Resources (TODO)

- ESIP Federation Member Organizations (with contact information)
- Websites / Documents / publications
- External organizations (standards, professional, etc.)
- Standards
- Technologies

[9] http://wiki.esipfed.org/index.php/Discovery_Change_Proposal-1
[10] http://wiki.esipfed.org/index.php/FederatedSearchConvention