Earth Science Web Service Advertisement and Discovery

Liping Di[a], Peisheng Zhao[a,b], Christopher Lynnes[b]

[a]*Center for Spatial Information Science and Systems, George Mason University, 6301 Ivy Lane, Suite 620, Greenbelt, MD 20770, USA*

[b]*NASA Goddard Space Flight Center, Code610.2, Greenbelt, MD 20771, USA*

***Abstract*- As Web service technologies mature, geospatial contents and capabilities are increasingly available online as interoperable Web services. But how to publish and locate a service is becoming a big challenge. Currently, government agencies and international organizations are making great efforts to develop and implement the standards and systems for service catalogue. This paper focuses on how to advertise and discover an Earth science Web service in the NASA Global Change Master Directory (GCMD), the Global Earth Observation System of Systems (GEOSS) Registry and the Federation of Earth Science Information Partners (ESIP) Service Casting (scast). The basic registry standards, mechanisms, and information models are discussed, and an example is provided to illustrate the challenges involved and the implementation of Earth science Web service advertisement and discovery.**

I. Introduction

A new scalable Service-Oriented Architecture (SOA) is emerging as the basis for distributed computing and large networks of collaborating applications. Web services provide users interoperable methods to access and use Earth Observation data. An increasing amount of geospatial content and capabilities for Earth science are available online as Web services, significantly enhancing the interoperability of systems for geospatial information and processing. As the number and variety of geospatial Web services increases rapidly, how to publish and locate a service is becoming a major problem. Aspects of this problem are domain knowledge, service metadata, and service categories. Currently, government agencies and international organizations are making great efforts to develop and implement standards and systems for a service catalog. The NASA Global Change Master Directory (GCMD) offers a high quality directory system to enable users to advertise and to discover Earth science data sets and services relevant to global change and Earth science research. The registry system of the Global Earth Observation System of Systems (GEOSS) allows members of the intergovernmental Group on Earth Observations (GEO) to advertise those of their services that follow industrial standards and provides a formal and searchable listing and description of service registry entries. Service Casting (scast) led by the Federation of Earth Science Information Partners (ESIP) provides self-describing scalable service advertisements and allows search engines and feed readers to automatically aggregate and discover services. This paper uses Open Geospatial Consortium (OGC) Web Map Service (WMS) for AIRS Near Real Time (NRT) data as an example illustrating advertisement and discovery of Earth Science Web services in NASA GCMD, GEOSS Registry and ESIP scast.

The remainder of this paper is organized as follows. Section 2 introduces the service taxonomy of the GEOSS Registry and describes how to apply the service taxonomy to service registration and discovery. Section 3 discusses the Service Entry Resource Format (SERF) introduced by NASA GCMD and how to use the SERF to register a service. In section 4, the ESIP scast is presented to discuss a decentralized catalogue mechanism for service advertisement and discovery. Finally, section 5 presents the conclusions and plans for future work.

II. GEOSS Registry

GEOSS is a global and flexible network of content providers that provides comprehensive, coordinated, and sustained observations of the Earth system. The goal is to improve monitoring, forecasting, and understanding of the behavior of the Earth system [1]. One key aspect of the GEOSS technical architecture is that all components are standards-based services and systems are service-oriented [2]. The GEOSS Registry System (http://georegistries.info) is a key infrastructural component that enables GEO members to advertise their services based on industrial standards and provides a formal and searchable listing and description of registry entries. This registry system supports multiple standardized interfaces, including Universal Description Discovery and Integration (UDDI), OASIS ebXML-ebRS, OGC Catalogue Service for Web (CSW), to make it accessible by a wide variety of applications. Its

significant issue is to define a lightweight service taxonomy based on service interfaces to capture domain knowledge and classify service instances. This taxonomy, as shown in table 1, uses the service's running behaviors, including category, standard, version, binding and profile, to represents service entities and relationships in formal structures [3]. Figure 1 shows the detailed information that is necessary to register and discover the example OGC WMS. In addition to service name and service description, information URL that refers to a human-readable service description, e.g. HTML documentation or a metadata file for the service interfaces, and an interface URL that refers to the service interface used by software, e.g. the OGC capability file (or service WSDL file), are essential as basic service information for service registration. Contact information, including name and email, provides the person or position name responsible for answering questions about this service instance. The service geographic extent and time period of the information provide the spatio-temporal attributes of the data this service hosts, which is very useful to evaluate the fitness of a particular service for a specific decision or assessment. The example OGC WMS is compliant with OGC WMS version 1.1 and is categorized as a Portrayal and Display Service. This registered classification information is used mostly to facilitate the discovery of available services.

TABLE I
GEOSS SERVICE TAXONOMY

Name	Description	Examples
Service Category	A class or group of geospatial services with common roles in the information processing of data-information-knowledge.	catalog/registry service, data access service, portrayal and display service, data transformation service
Service Standard	A specification of how a service fulfills its functionality through its standardized interfaces.	OGC Web Coverage Service (WCS), OGC Web Feature Service (WFS), OGC Web Map Service (WMS), OGC Catalogue Service for Web (CS/W)
Standard Version	A series of revisions of a service standard.	OGC WMS 1.1, OGC WMS 1.3, OGC WCS 1.0, OGC WCS 1.1
Service Binding	A concrete accessing protocol and data format specification.	Simple Object Access Protocol (SOAP), Hypertext Transfer Protocol (HTTP) POST, HTTP GET
Service Profile	An agreed-upon subset and interpretation of a specification	ISO19115:2003/ISO19119 Application Profile for CSW, OWL Application Profile of CSW

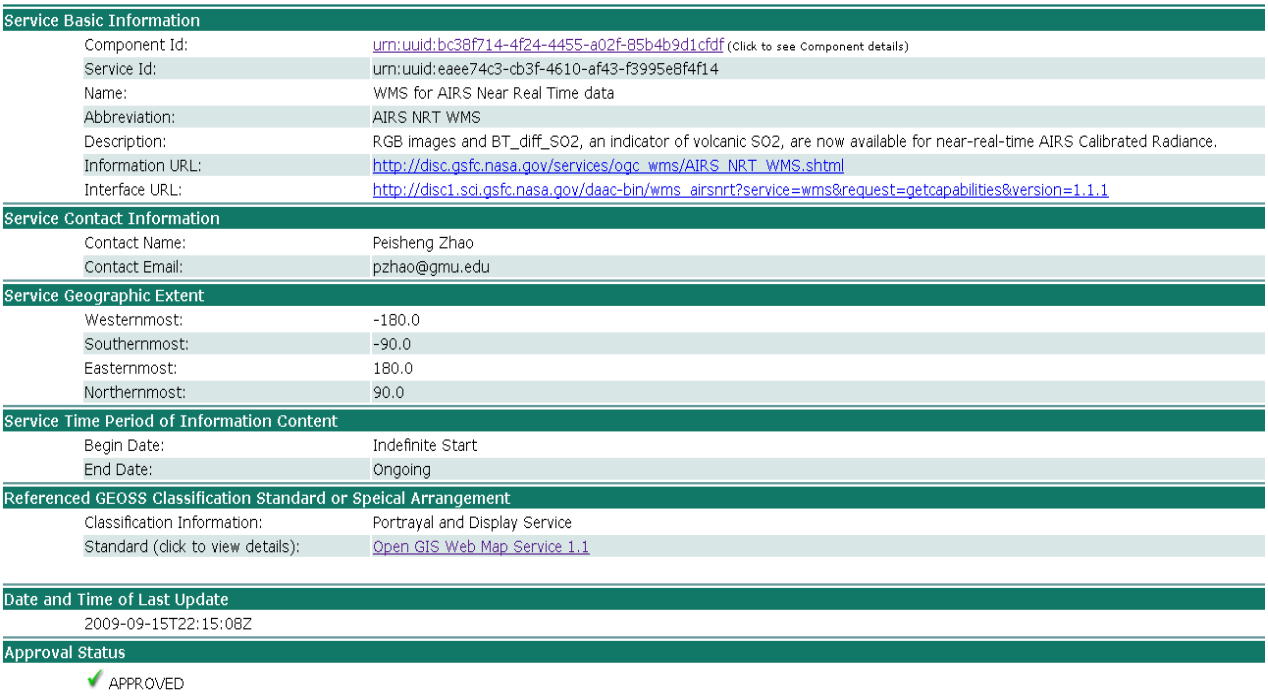

Fig. 1. GEOSS Service Instance

III. NASA GCMD

NASA GCMD (http://gcmd.nasa.gov) is a promising catalogue system that enables users to discover access, use and publish Earth science data and data-related services relevant to global change and Earth science research. This catalogue system holds more than 30,000 records of Earth science data, services, and ancillary information covering all aspects of Earth and environmental sciences. The GCMD uses Service Entry Resource Format (SERF) to record service directory entries related to the acquisition, processing, retrieval, viewing, analysis, archival, and interpretation of Earth science data services. The SERF is a XML document that includes mainly seven mandatory fields for service registration. Listing 1 shows a SERF for the OGC WMS example.

```
<Entry_ID>NASA_AIRS_NRT</Entry_ID>
<Entry_Title>OGC WMS for AIRS NRT Data Products</Entry_Title>
<Service_Parameters>
  <Service_Category>Earth Science Services</Service_Category>
  <Service_Topic>Web Services</Service_Topic>
  <Service_Term>Information Management Services  </Service_Term>
   <Service_Specific_Name>Web Map Service</Service_Specific_Name>
 </Service_Parameters>
<Science_Parameters>
   <Science_Category>Earth Science</Science_Category>
   <Science_Topic>Atmosphere</Science_Topic>
   <Science_Term>Atmospheric Chemistry</Science_Term>
   <Science_Variable>Carbon and Hydrocarbon Compounds
</Science_Variable>
  </Science_Parameters>
<ISO_Topic_Category>Climatology/Meteorology/Atmosphere</ISO_Topic
_Category>
<Service_Provider>
   <Service_Organization>
   <Short_Name>NASA/GSFC/SED/ESD/GCDC/GESDISC</Short_Name
>
    <Long_Name>Goddard Earth Sciences Data and Information Services
Center (formerly Goddard DAAC), Global Change Data Center, Earth
Sciences Division, Science and Exploration Directorate, Goddard Space
Flight Center, NASA</Long_Name>
   </Service_Organization>
   <Service_Organization_URL>http://disc.gsfc.nasa.gov/</Service_Organi
zation_URL>
  <Personnel>
    <Role>Service Provider Contact</Role>
    <First_Name>Peisheng</First_Name>
    <Last_Name>Zhao</Last_Name>
    <Email>peisheng.zhao-1@nasa.gov</Email>
    <Phone>301-614-5312</Phone>
    <Fax>301-614-5268</Fax>
    <Contact_Address>
     <Address>NASA GSFC, Code 610.2</Address>
     <City>Greenbelt</City>
     <Province_or_State>MD</Province_or_State>
     <Postal_Code>20771</Postal_Code>
     <Country>U.S.A.</Country>
    </Contact_Address>
  </Personnel>
 </Service_Provider>
 <Summary>This is one of the GES DISC's OGC Web Map Service
(WMS)instances for AIRS NRT data products which are currently available
for Level-1B and Level-2, and are specially useful for users whose primary
interest is the low latency for data availability..  </Summary>
<Metadata_Name>CEOS IDN SERF</Metadata_Name>
 <Metadata_Version>VERSION 9.7.1</Metadata_Version>
```

Listing 1. Core elements of SERF

SERF uses *Service_ID* and *Service_Provider* to record the basic service information, such as organization, contact person and contact address. Its key property is to make use of a four level hierarchy of service keywords [4], Category/Topic/Term/Service Specific Name, to identify a service in general, and a four level hierarchy of science keywords [5], /Category/Topic/Term/Variable, to associate a service with a science topic, term and variable. Thus, users can discover the services that best fits a scientific problem. SERF also includes some optional elements to provide specific service information, as shown in Listing 2. For example, *Sensor_Name* and *Source_Name* are the names of the instrument and platform used to acquire data related to the service, and *Related_URL* specifies some URL links that point to a service description and the service endpoint.

```
<Sensor_Name>
   <Short_Name>AIRS</Short_Name>
   <Long_Name>Atmospheric Infrared Sounder</Long_Name>
 </Sensor_Name>
 <Source_Name>
   <Short_Name>AQUA</Short_Name>
   <Long_Name>Earth Observing System, AQUA</Long_Name>
 </Source_Name>
<Related_URL>
   <URL_Content_Type>
    <Type>GET SERVICE</Type>
    <Subtype>GET WEB MAP SERVICE (WMS)</Subtype>
   </URL_Content_Type>
   <URL>http://disc1.sci.gsfc.nasa.gov/daac-bin/wms_airsnrt</URL>
 </Related_URL>
```

Listing 2. Optional elements of SERF

Overall, SERF focuses more on the scientific description of the data related to the registered service, than on how to access this service programmatically. Although in general a service keyword is very useful for identifying services, missing service implementation information, such as the

algorithm, will still result in service discovery mismatching. The GCMD provides a Web portal for Web service advertisement and discovery. This portal provides an online tool to aid users in building a SERF and a wizard to guide users to find a service step by step based on the hierarchical SERF elements as shown in figure 2.

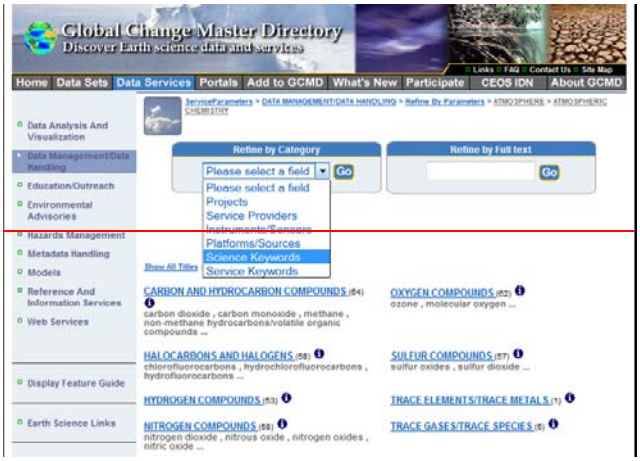

Fig. 2. Service discovery in NASA GCMD

IV ESIP SERVICE CASTING

The GEOSS Registry and NASA GCMD are centralized catalogue systems which require service providers to register services in and users to discover services from a major central repository. The ESIP scast is a mechanism for a decentralized service catalogue that does not require a centralized system and a certain central repository. By using Atom syndication feeds, a simple HTTP-based protocol for creating and updating web resources, ESIP scast enables the following features [6, 7]:

- Dispersing the governance of service advertisement closer to providers.
- Using machine-readable metadata to categorize services
- Providing interface description links for automatic service invocation.
- Importing Outline Processor Markup Language (OPML) for simple aggregation.
- Providing an open search interface or using a feed reader for scalable discovery.

Basically, each entry in the scast (Atom feed) represents the advertised service by including a set of links, the service type, and content information. The links point to a machine-readable interface description (e.g. WSDL, WADL, or OGC WXS GetCapabilities call), a service executable endpoint URL, and human-readable documentation. The service type information defines the service semantics and an invocation protocol to interpret the entry links. The content information can contain any XHTML or XML including additional information and metadata. Listing 2 shows a scast (Atom feed) for advertising the example service. This document explicitly presents the example service as an OGC WMS so that a machine/program can do an OGC GetCapabilities HTTP call to parse the service interface and invoke the service automatically, and also to provide an example of the WMS GetMap call to show how to get a map. A scast document can be deployed at any convenient Web-accessible location and updated to re-advertise at any time. Thus, users can subscribe to this document to get up-to-date information about a service in a Web browser, as shown in Figure 3.

```
<feed xmlns=" http://www.w3.org/2005/Atom"  xml:lang="en"
xmlns:scast="http://sciflo.jpl.nasa.gov/serviceCasting/2009v1">
<title>AIRS Near Real Time WMS</title>
 <link rel="self" href="http://g0hep12u.ecs.nasa.gov/airs_nrt_wms.xml"/>
  <updated>2010-03-04T19:33:59.368863</updated>
  <author>
    <name>Christopher Lynnes</name>
    <email>Chris.Lynnes@nasa.gov</email>
  </author>
  <id>uri:http://disc1.sci.gsfc.nasa.gov/daac-bin/airsnrt_wms</id>
   <entry> <title>AIRS Near Real Time WMS</title>
    <id>uri:http://disc1.sci.gsfc.nasa.gov/mapserv-bin/airsnrt_wms</id>
    <updated>2010-03-04T19:33:59.368863</updated>
    <summary>AIRS Near Real Time WMS views of BT_diff_SO2 (an
indicator of volcanogenic SO2) and RGB images
    </summary>
    <scast:serviceSemantics>OGC.WMS</scast:serviceSemantics>
    <scast:serviceProtocol>HTTP</scast:serviceProtocol>
    <category schema="scast" term="AIRS Near-real-time SO2" />
    <link type="application/xml"  title="Server endpoint"
        href="http://disc1.sci.gsfc.nasa.gov/daac-bin/airsnrt_wms" />
    <link rel="scast:serviceInterface" type="application/xml"
        title="Service interface description"
        href="http://disc1.sci.gsfc.nasa.gov/daac-
bin/airsnrt_wms?SERVICE=WMS&amp;VERSION=1.1.1&amp;REQUES
T=GETCAPABILITIES" />
    <link rel="scast:serviceEndpoint" type="application/xml"
        title="Server endpoint"
        href="http://disc1.sci.gsfc.nasa.gov/mapserv-bin/airsnrt_wms" />
    <link rel="scast:serviceDocumentation" type="text/html"
      xml:lang="en-us"  title="Service documentation"
href="http://disc.gsfc.nasa.gov/services/wxs_ogc.shtml" />
    <link rel="alternate" type="text/html" xml:lang="en-us"
        title="Service documentation"
        href="http://disc.gsfc.nasa.gov/services/wxs_ogc.shtml" />
    <content type="xhtml">
      <div xmlns="http://www.w3.org/1999/xhtml">
     <p><b>Example Call (South America, BT_diff_SO2, Descending
node)</b> </p>
      <p> <![CDATA[
http://disc1.sci.gsfc.nasa.gov/daac-
bin/wms_airsnrt?SERVICE=WMS&VERSION=1.0.0&REQUEST=GETM
AP&CRS=EPSG:4326&FORMAT=image/png&BBOX=-100,-60,-
20,20&WIDTH=512&HEIGHT=512&LAYERS=AIRS_SO2_D,coastline
```

```
]]>   </p> </div>
   </content>
  </entry>
</feed>
```

Listing 3. ESIP scast (Atom feed) document

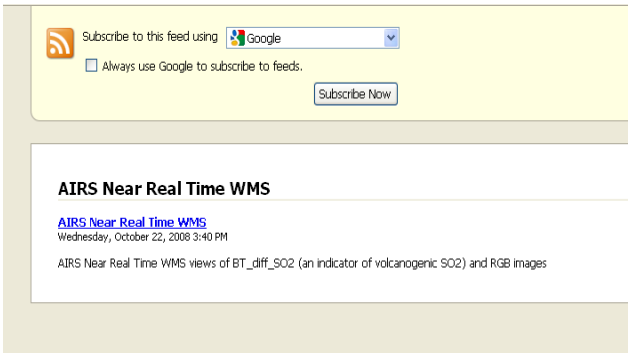

Fig. 3. Scast subscription in a Web browser

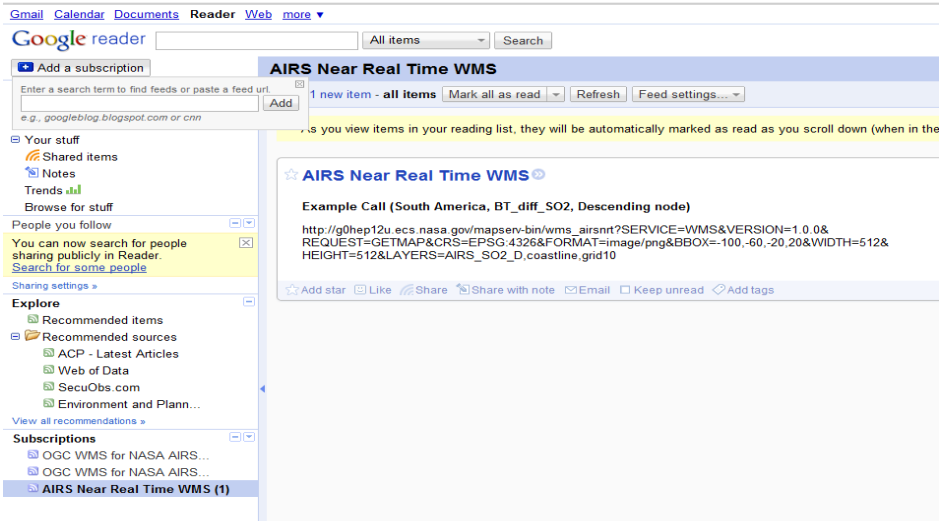

Fig.4. ESIP scast discovery using Google Reader

V. CONCLUSION

GEOSS Registry, NASA GCMD, and ESIP scast are the service catalogue systems that are being used most widely. The GEOSS Registry uses a hierarchical taxonomy to describe service running behaviors and to use a formal structure to represent service entities and relationships. NASA GCMD defines a set of service and science keywords to associate a service with a scientific problem and identify the category to which a service instance belongs ESIP scast is a decentralized self-publish/subscribe system for a service catalogue that encodes service type, It links information using Atom syndication feeds to enable service discovery in Google already supports Atom feed searches. The user already can search services advertised by ESIP scast using Google. This search capability extremely enhances the capability for discovering services. Google Reader (http://reader.google.com) is a Web-based aggregator, capable of reading Atom feeds. By using Google Reader, a user can search in ESIP scast to discover and view services. Figure 4 shows how the example service appears in Google Reader. To subscribe or discover the example service, users need to input the service's scast URL or the other keywords, into Google Reader's *Add subscription* box to ask Google to search relevant ESIP scasts. Once the example service is found, users can subscribe to it to view its detailed information and get real-time information when the service is updated.

a variety of search engines and feed readers. This paper discusses the basic registry standards, mechanisms, and information models. It examines service entities and registration entries that allow services to be advertised and discovered. This paper uses OGC WMS as an example to describe and discuss how to advertise and discover an Earth science data service in the GEOSS Registry, NASA GCMD, and ESIP scast.

To improve Earth science Web service advertisement and discovery, a full comparison of different registry information models, taxonomy semantics, and programming/user interfaces in different scenarios is being planned.

REFERENCES

[1] GEO. (2009, May 5). *GEO - Group on Earth Observation*. Available: http://www.earthobservations.org

[2] G. Percivall, "Web Services for the Global Observing System of Systems," in *Geospatial Web Services: Advances in Information Interoperability*, P. Zhao and L. Di, Eds., ed Hershey: IGI-Global, in press.

[3] Y. Bai, *et al.*, "GEOSS Component and Service Registry: Enalbing the Registering and Discovering of Geospatial Resources Worldwide," presented at the 2010 AAG Annual Meeting, Washington, DC, USA, 2010.

[4] GCMD, "GCMD Earth Science Services Keywords," NASA GSFC, Greenbelt, MD, USA2009.

[5] GCMD, "GCMD Science Keywords," NASA GSFC, Greenbelt, MD, USA2008.

[6] B. Wilson, "Service Casting: A Proposal for Advertising Web Services," presented at the 7th Earth Science Data Systems Working Group (ESDSWG) Meeting, Philadelphic, PA, USA, 2008.

[7] JPL. (2009, May 7). *Service Casting*. Available: http://sciflo.jpl.nasa.gov/scast/