ESIP Discovery Cluster

OpenSearch, Datacasting, Servicecasting

ESIP 2011 Winter Meeting

Wednesday, January 5, 2011

Washington, DC

Hook Hua

Agenda

- Introduction (10-minutes)
- Interoperability Use Cases (45-minutes)
  - With discussion on interoperability
- Setup governance mechanism (30-minutes working session)
- Discussion on key issues (20-minutes)
- Atom vs RSS comparative (1-slide)
- Settle on specification of response format (10-minutes working session)

Some Background

INTRODUCTION

Discovery Services History Highlights

- **2005**: original Data Casting (with RSS) funded
- **2008**: ESDSWG Technology Infusion working group explores Service Casting
- **2009-01-06**: FROST concept proposed at ESIP, Washington DC
- **2009-07-07**: FROST demo at ESIP, Santa Barbara
- **2009-07-14**: ESIP FedSearch Cluster startup. "*Also Bastille Day! Coincidence?*"
- **2009-10-21**: Presentation at ESDSWG on Federated Search
- **2010-07-20**: ESIP, Knoxville
  - ESIP FedSearchWorkshop
  - second Data Casting
  - Service Casting
  - *Discussed overlapping Atom response formats*
- **2010-11-04**: ESDSWG, New Orleans
  - Federatedness/Casting session
    - Draft of initial ESIP Discovery Atom Response Format
    - Expand the scope of Federated search to cover Discovery services (OpenSearch, Data Casting, Service Casting)
- **2010-11-30**: First working draft document of the Discovery feed Atom response format posted on wiki
- **2010-12-07**: first Discovery Cluster telecon
- **2011-01-05**: Discovery Cluster working session (today)

The REST-style Common Response

ESIP Federated Search + Service Casting + Data Casting = ESIP Discovery

Service	Request	Response
OpenSearch Granule-level*	GET REST-style parameters	Atom 1.0 + extensions
DataCasting**	GET	Atom 1.0 + extensions
ServiceCastng	GET	Atom 1.0 + extensions

** Collection-level OpenSearch returns OSDD*
*** Consists of both collection and granule casting*

Atom 1.0

- Starting with RFC 4287 “The Atom Syndication Format”
  - Not to be confused with RFC 5023 "The Atom Publishing Protocol"
- Generic example:

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns=http://www.w3.org/2005/Atom>

 <title>Example Feed</title>
 <link href="http://example.org/"/>
 <updated>2003-12-13T18:30:02Z</updated>
 <author>
  <name>John Doe</name>
 </author>
 <id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af6</id>

 <entry>
  <title>Atom-Powered Robots Run Amok</title>
  <link href="http://example.org/2003/12/13/atom03"/>
  <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
  <updated>2003-12-13T18:30:02Z</updated>
  <summary>Some text.</summary>
 </entry>

</feed>
```

Example ESIP Extensions for Atom's Entry Tag

(Illustrative non-final format example)

```
<entry>
 <title>regrid-modis.aqua-avg.1hr.p5deg-v1.0-2005.01.01.blk1.nc</title>
 <link href="http://wsnews.jpl.nasa.gov:8100/file/download_by_id?
     id=178df53dada229cec0dbec347c60785dde8bfdce" length="64959344" title="Granule File"
     type="application/x-netcdf" />
 <link href="http://wsnews.jpl.nasa.gov:8100/file/download_by_id?
     id=178df53dada229cec0dbec347c60785dde8bfdce" length="64959344" rel="http://esipfed.org/ns/
     discovery/1.1/data#" title="Granule File" type="application/x-netcdf" />
 <geo:box>-180.0, -90.0, 180.0, 90.0</geo:box>
 <time:start>2005-01-01T00:00:00Z</time:start>
 <time:end>2005-01-01T03:00:00Z</time:end>
 <updated>2008-12-16T15:00:26Z</updated>
 <discovery:metadata xsd="http://foo/bar.xsd">
  <discovery:dayNightFlag>day</discovery:dayNightFlag>
  <gesdisc:fooFlag>bar</gesdisc:fooFlag>
 </discovery:metadata>
 <content type="xhtml">
  Here is the description.
 </content>
</entry>
```

Onto Discovery

CLUSTER TRANSITION

FedSearch to Discovery

- Decided at ESDSWG 2010, New Orleans
- Cluster Transition
  - Old: ESIP Federated Search Cluster
    - http://wiki.esipfed.org/index.php/Federated_Search
  - New: ESIP Discovery Cluster
    - http://wiki.esipfed.org/index.php/Discovery_Cluster
- Mailing listing
  - Deprecating old FedSearch mailing list
    - http://rtpnet.org/mailman/listinfo/esip-federatedsearch
  - Subscribe to new Discovery mailing list
    - http://rtpnet.org/mailman/listinfo/esip-discovery

Mashups and more?

INTEROPERABILITY USE CASES

Example Use Cases

- *"Service casting and data (granule and collection) casting"*
  - Ruth Duer, 10-minutes
    - National Libre Portal
    - Plan to support OpenSearch's multiple response format types (Atom, RSS, KML)
    - Plan to make OpenSearch and Datacasting response the same format types (Atom, RSS, KML)
    - Both OpenSearch and Datacasting support collection and granule-level search
- *"Extend OpenSearch to service casting"*
  - Chris Lynnes, 10-minutes
- *"Coastal data visualization data casting"*
  - Ken Keiser, 10-minutes

Discussion on Interoperability

- Given these uses cases as starting point, what is needed for interoperability?
- OGC is also working on OpenSearch. We should be in contact.
- Publish/Subscribe model. Follow RFC 5023 "The Atom Publishing Protocol”?
- …

Working Session to Setup Formalism

GOVERNANCE

Governance

- First off, do we need a governance process?
- Want open and interoperable Discovery standard
- Need to move forward.
  - Multiple data centers involved in Discovery services.
  - Issues with interoperability, different centers, different standards.
  - Agree on process, and then follow it.

What We Need

- Process defined for
  - Proposing
  - Forum
  - Ratifying
  - Changing standards (as new proposal)
- Documentation
  - put transparent trail on wiki
- Appoint people? (initially)
- Look to open source community as examples
  - Java Community Process
  - "Committers" work with community, implement changes, accept input
  - Apache model

Example: Governance of the Open Provenance Model

- *(just an example, we don't have to follow this governance process)*
- OPM *editors* act as *committers* to the OPM specification(s).
  - Bootstrap with initial editors
  - New community participants will become editors over time.
- Avoid splintered community
  - Prevent branching with an open and consensus driven process

Process for Changes to OPM

All steps posted to **mailing list** and **wiki**

1. Call for change proposals.
   - Editors invite community to submit change proposals
   - Include Context: 'background', 'problem addressed', 'proposed solution', 'rationale for the solution'
2. Proposal review period.
   - Community reviews and comments on change proposal
   - Set time limit
3. Vote.
   - Community wiki account holders vote.
   - Change proposal adopted with majority of community votes **and** agreed upon by majority of editors
4. Document editing.
   - Editors provide revised version of document on wiki
5. Final review.
   - Community provides final comments before finalizations.

Custom Profiles

- "Profiles, i.e. specializations of OPM, with specific focus or for specific application domains are encouraged."
- Proposed custom profiles go through same change process
- Discovery will need this
  - e.g. GES DISC day-night flag

Discovery Cluster Governance

- …So what are we deciding?
- Initial Editors *(should be an odd number)*
  1. Andy Bingham
  2. Chris Lynnes
  3. Brian Wilson
  4. Hook Hua
  5. Ruth Duer
  6. EPA may also be interested….
- Adopt OPM governance, but with changes:
  - Each proposal shall have a deadline range, which is not extendable
- Setup initial governance template pages on wiki
- Later, propose community “standard” to ESDSWG SPG. Write RFC.

Key Issues with Discovery Services

DISCUSSION

Which Response Format to Initially Adopt?

ATOM VERSUS RSS

** Brian Wilson unable to make it to ESIP, but did send a slide.*

Why Atom Instead of RSS?

- Multiple Typed Links possible in each Entry
  - <link rel="<purpose>" type="application/wsdl+xml" href="" />
  - Links to interface, service endpoint, & human documentation
- Can embed tags from arbitrary namespaces
  - Define scast namespace, and register it
  - <scast:serviceType> = SOAP, REST, OGC.WMS, HUMAN
  - Specific services can embed more XML metadata (sciflo)
- Embed arbitrary content in 'typed' <content> tag
  - <content type="xhtml" or type="xml">
  - Use for more documentation, or more XML metadata
- Use <category> tag to characterize services
  - Use taxonomy of services, or just text keywords
- Opensearch and GoogleData standards extend Atom
- Atom got timestamps right (unlike RSS)
  - ISO YYYY-MM-DDTHH:MM:SS

Brian Wilson (JPL)

Identifying Key Issues

- Atom Syndication Format vs RSS 2.0
  - *Andy Bingham's Datacasting project has agreed to also support Atom in his Datacasting reader*
- Custom Profiles
- Custom OpenSearch request parameters
  - Allow custom value restrictions?
- Others…

Specification

DISCOVERY RESPONSE FORMAT

Settling on Initial Response Format

- Discovery Cast Atom Response Format
  - http://wiki.esipfed.org/index.php/Discovery_Cast_Atom_Response_Format
- Namspace
- ESIP extensions to Atom Response
  - Temporal extents
  - Spatial extents
  - rel links type vocabulary
  - Pagination support
  - Custom tags (with namespace!) at top-level of <entry>

Next Steps

1. Submit your Discovery services URL to the wiki page (informal)
2. Continue development on Discovery mailing list
   - Settle on ESIP request convention
     - OpenSearch, DataCasting, ServiceCasting
3. Implementations adhere to ESIP Discovery specification
4. Demonstrations of adherence
5. Interoperable mashups?
6. Federated orchestration
7. …

References

- ESIP Discovery Cluster
  - http://wiki.esipfed.org/index.php/Discovery_Cluster
- ESIP Discovery mailing list
  - http://rtpnet.org/mailman/listinfo/esip-discovery
- Initial Discovery Cast Atom Response Format
  - http://wiki.esipfed.org/index.php/Discovery_Cast_Atom_Response_Format
- The Atom Syndication Format
  - http://www.ietf.org/rfc/rfc4287.txt
- Governance of the Open Provenance Model
  - http://twiki.ipaw.info/pub/OPM/WebHome/governance.pdf