GeoWS: Making Services and Data Consistent

M. Stults[1], T. Ahern[1], B. Arko[2], S. Carbotte[2], E. Davis[3], D. Ertz[4], M. Gurnis[5], J. McWhirter[7], C. Meertens[4], M. Ramamurthy[3], C. Trabant[1], M. Turner[5], D. Valentine[6], I. Zaslavsky[6]

[1]IRIS Data Management Center, Seattle, Washington
[2]Lamont-Doherty Earth Observatory, Columbia University
[3]Unidata - Boulder, Colorado
[4]UNAVCO - Boulder, Colorado
[5]Seismology Laboratory - Caltech
[6]CUAHSI - San Diego Super Computer
[7]GeodeSystems- RAMADDA- Boulder, Colorado

Project Members

IRIS, Caltech, GeodeSystems, LDEO, SDSC - CUAHSI, UNAVCO, Unidata

Project Member's GeoWS Endpoints: http://geows.ds.iris.edu/endpoints/

GeoWS EarthCube page : http://earthcube.org/group/geows-geoscience-web-services

GeoWS technical home: http://geows.ds.iris.edu/

GeoWS documents: http://geows.ds.iris.edu/documents/

Definitions

GeoWS Service - a service which has
- RESTlike interface
- standardized space, time, format query parameters
- standardized error response
- offer GeoCSV as an output format when possible

URL builder - Web application which documents and creates URLs based on respective user selections

GeoCSV output data format - is CSV format with a small set of optional keywords, e.g. field_units, field_type, etc. and a few keywords from Climate and Forecast Conventions like title, comment, etc.

GeoCSV - is compatible with **W3C - CSV on the Web Working Group (CSVW) :** http://www.w3.org/2013/csvw/wiki/Main_Page

Swagger Specification (now Open API Initiative): a specification using JSON to contain service and API documentation, see: https://openapis.org/specification

Swagger-ui: a web application that can leverage the Swagger Specification contents to provide a user interface, see: http://swagger.io/swagger-ui/

EarthCube Architecture Alignment

GeoWS services use **standardized query parameters** —> directly supports integration with brokers

GeoWS services are **loosely coupled** —> this supports System of Systems concepts of independently managed components, but operating for shared goals of information discovery and cross domain access.

GeoWS and GeoCSV concepts **establish simple, low barrier of entry**, patterns of use for web service and data handling, **however, for more complex data**, additional output formats should be standardized for netCDF, HDF5, and respective domain formats.

1 GeoWS Concepts

1 - use partner's experience to define standard approach for space-time queries in web services.
2 - investigate data delivery formats with ease of use and human readability in mind.
3 - develop representative services which allow testing and evaluation of concepts.
4 - test concepts against a range of services

2 Concepts Developed

GeoWS - a pattern for web service structure
GeoCSV - a data format that is human readable, easy to build client for, but a can have simple extensions for more rigorous data definitions
URL Builders - refine and recognize importance of tool to promote access to web service
Evaluate Swagger Specification (now Open API Initiative) - explore utility as an alternative to WADL for web service description and documentation
Evaluate Swgger-UI - a web application which leverages Swagger Specification to view service documentation and build URLs

3 Building Block Activities

by funded partners
- used existing tools
- used existing infrastructure and service framework

→ Modified or extended existing resources → GeoWS Service

GeoWS services: http://geows.ds.iris.edu/endpoints/

for unfunded collaborators
used existing web capabilities from collaborators, which are
- not standardized
- not modifiable (at this time)

→ Added resources to host proxy services and metadata storage → GeoWS Proxy Service

GeoWS proxy services: http://geows.ds.iris.edu/geows-uf/

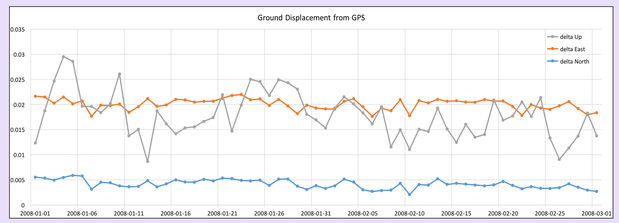
4 An example cross-domain query for ground displacement understandable to other earth scientist like seismologist

Query:
http://web-services.unavco.org/gps/data/position/P378/v3?analysisCenter=pbo&referenceFrame=igs08&starttime=2008-01-01T00:00:00&endtime=2008-03-01T00:00:00&report=long&refCoordOption=from_analysis_center

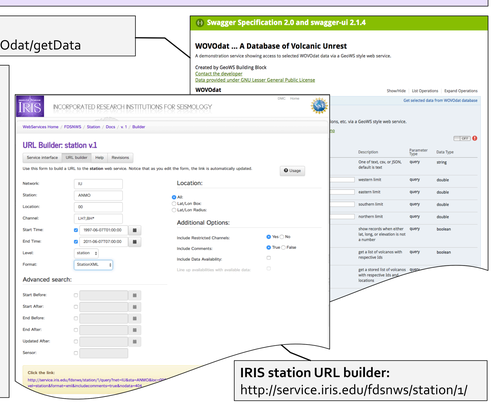
GeoCSV out:

```
# dataset: GeoCSV 2.0
# field_unit: ISO 8601 datetime UTC, meters, meters, meters, meters, meters, meters, number, number, number, degrees north, degrees east, meters, meters, meters, meters, meters, meters, meters, number, number, number, UTF-8
# field_type: string, float, float, float, float, float, float, float, float, float, float, float, float, float, float, float, float, float, string
# attribution: http://www.unavco.org/community/policies_forms/attribution/attribution.html
# Request URI: http://web-services.unavco.org/gps/data/position/P378/v3?analysisCenter=pbo&referenceFrame=igs08&starttime=2008-01-01T00:00:00&endtime=2008-03-01T00:00:00&report=long&refCoordOption=from_analysis_center
# Source File: P378.pbo.igs08.pos Date: 2016-06-21 07:51:39
# XYZ Reference Coordinate (igs08) X: -2475699.75517 Y: -3822330.2947 Z: 4450718.33167
# Applied Offset: X: 0 Y: 0 Z: 0
Datetime, X, Y, Z, Std Dev X, Std Dev Y, Std Dev Z, Corr XY, Corr XZ, Corr YZ, N latitude, E longitude, Height, delta N, delta E, delta U, Std Dev N, Std Dev E, Std Dev U, Corr NE, Corr NU, Corr EU, Solution
2008-01-01T00:00:00, -2475699.73592, -3822330.29465,4450718.32907,0.00236,0.00342,0.00356,0.80200, -0.75100, -0.81400,44.53498,237.06912,83.12267,0.00552,0.01613, -0.00931,0.00121,0.00157,0.00510, -0.02029,0.02750, -0.17109,repro
2008-01-02T00:00:00, -2475699.73852, -3822330.29869,4450718.33356,0.00243,0.00354,0.00367,0.80200, -0.75200, -0.81500,44.53498,237.06912,83.12925,0.00535,0.01614, -0.00274,0.00125,0.00162,0.00527, -0.02534,0.03542, -0.17589,repro
2008-01-03T00:00:00, -2475699.74214, -3822330.30273,4450718.33829,0.00263,0.00379,0.00388,0.81000, -0.75200, -0.81300,44.53498,237.06912,83.13638,0.00496,0.01530,0.00440,0.00132,0.00175,0.00562, -0.02119,0.01476, -0.20462,repro
```

Ground Displacements:



GeoWS proxy service - Swagger-ui example:
http://geows.ds.iris.edu/geows-uf/wovodat/1/#!/WOVOdat/getData



IRIS station URL builder:
http://service.iris.edu/fdsnws/station/1/

Unfunded Collaborators

Global Geodynamics Project (GGP) superconducting gravimeters	NCEI - National Centers for Environmental Information - **Waterlevel Tides** & meteorological satellite image (**DMSP**) data
Gravity and Magnetics - data collected by Randy Keller - UTEP	
INTERMAGNET - consortium of magnetic observatories	OOI- **Ocean Observatory Initiative - Seismic** data
StraboSpot - Structural Geo. & Tect. - U. of Kansas	**WOVOdat** - A database of volcanic unrest data from World Organization of Volcano Observatories
National Ecological Observing Network (NEON)	

Lessons Learned

Creating proxy services or modifying existing services can be fast and efficient if native service output format match desired discovery pattern.
IRIS's component **Web Service Shell** enables a repeatable, cost effect way to quickly add web services.

However, a native service with no metadata and /or having poor match of native API to desired API (i.e. discovery, data retrieval APIs, etc.)
- Creates long term need for additional infrastructure resources
- Creates long term need for curation efforts to harvest metadata, report errors when found, etc.

IRIS can add unfunded collaborator services and extend these concepts to other providers if additional resources are available

URL Builder - Lessons Learned

Storing service descriptions in Swagger Specification JSON standard is relatively easy,.
Swagger Specification is adequate for services with simple hierarchical parameter relations, but it becomes less usable for more complex interrelated parameters
Swagger-ui presents documentation and enables URL building with expect level of effort for services with moderate text output only, large text streams or binary output is not handled successfully without code modification
Custom web development is recommended for best users experience for more complex service descriptions