

UNAVCO and COOPEUS

The Plate Boundary Observatory (PBO), operated by UNAVCO, is the geodetic component of the US National Science Foundation-funded EarthScope Facility. A primary scientific objective of EarthScope is quantifying the three-dimensional deformation and its temporal variability across the active boundary zone between the Pacific and North American plates. To achieve this goal, UNAVCO has installed GNSS, strain, and seismic instrumentation at over 1,200 sites in the western US and manages processing, analysis and distribution of PBO data and products under EarthScope's open data policy. COOPEUS, the European Union project to strengthen the cooperation between the US and the EU in the field of environmental research infrastructures, will link EarthScope and the European Plate Observing System in joint research infrastructure enhancement activities that will ultimately advance international geodesy data sharing. (COOPEUS also links a broad set of additional EU and US based Earth, oceans, and environmental science research entities in joint research infrastructure enhancement activities.) To enhance data sharing within the US, UNAVCO and several partner geodesy data centers have implemented the Geodesy Seamless Archive Centers (GSAC), a web services based technology to facilitate the exchange of geodesy metadata and delivery of geodesy data and products to users. These services utilize a repository layer implemented at each data center, and a service layer to identify and present any data center-specific services and capabilities, allowing simplified vertical federation of metadata from independent data centers. Within the context of COOPEUS, we envision that EarthScope-EPOS partnership could leverage GSAC or similar technologies to further advance data sharing among multiple data centers.

GSAC Web Service Software

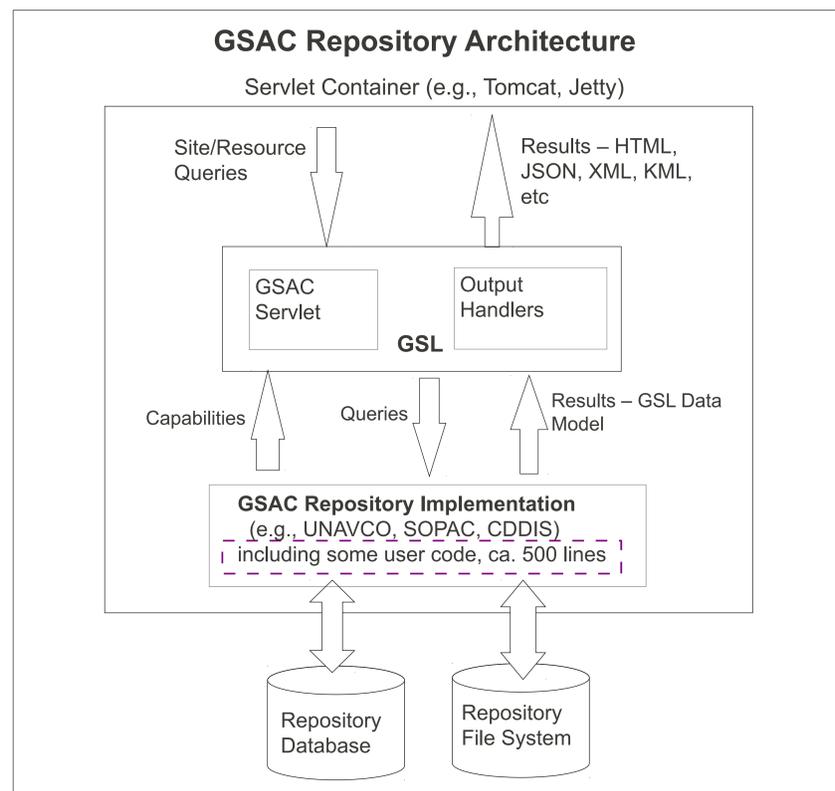
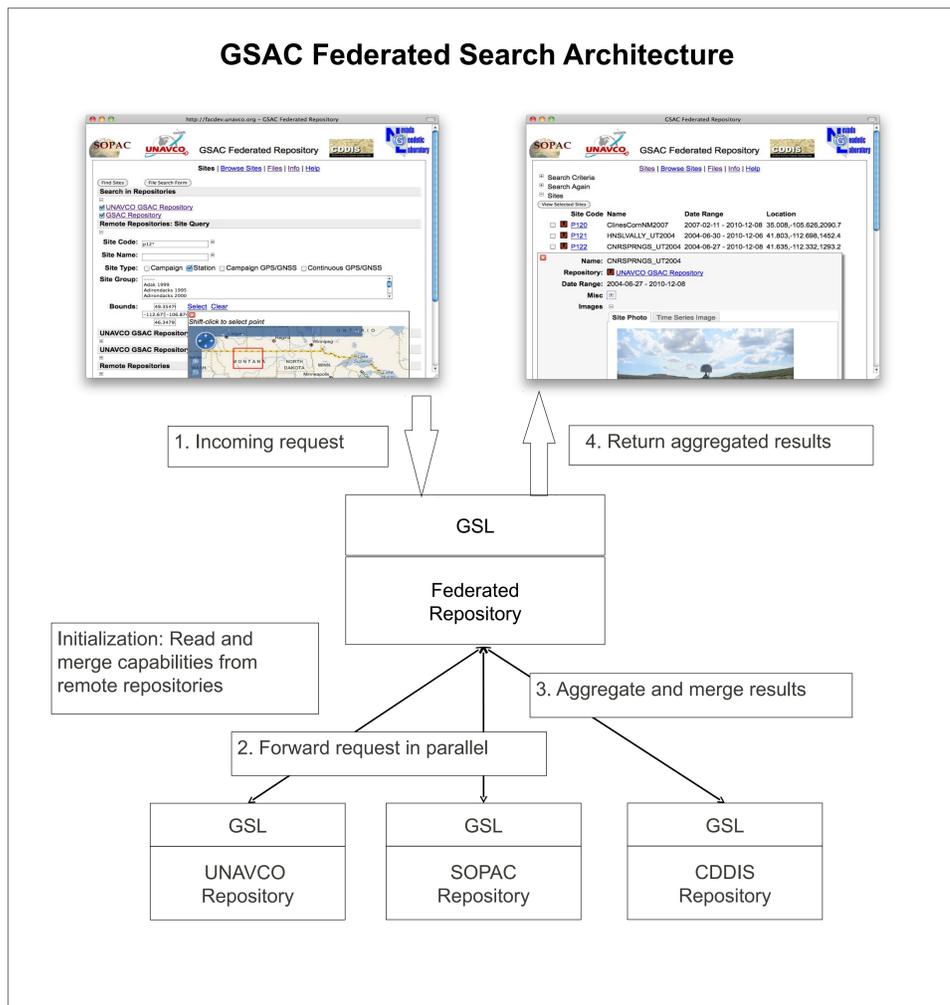
The new Web Services enabled Geodesy Seamless Archive Centers (GSAC WS) software from UNAVCO and partners SOPAC and CDDIS is a package for comprehensive user access to a broad range of space geodetic data, including GPS/GNSS, DORIS, SLR, and VLBI. GSAC WS is a suite of free, ready-to-use code that implements a web services API, provides for data downloading, and includes a browser-based GUI for data search; client software to facilitate command line interaction with the API are also part of the GSAC code. An organization with geodesy data files and information about their collections in a database can readily implement a GSAC Repository utilizing GSAC code; in many cases only minor coding would be needed to get a GSAC Repository up and running.

The intention of GSAC is to facilitate easy creation of useful web services for search and discovery of data in existing repositories. GSAC software supports single repositories, but also becomes particularly powerful when used in the federated mode for search and download from several different GSAC-enabled data repositories. The federated GSAC system dynamically queries participating GSAC Repository directly through their locally-implemented GSAC. A federated GSAC does not maintain a database about other data holdings, nor does it mirror collections of data files.

The GSAC design utilizes a data model that is likely to be common to most geodesy data centers, and certain other Earth and environmental sciences data centers. For example, a site-based data collection paradigm is utilized within GSAC, and metadata such as site location and instrumentation installed at a site are typically elements of the database schema for such a repository. For files containing the observations, metadata about the data files that often is stored in the database typically includes start and end times, checksums and file sizes, etc. Searches and downloads of data from geodesy data centers often use the same or similar functions such as name-based or space-time queries. The GSAC implementation provides code for these common needs, plus an ability to tailor the details to implement capabilities specific to a particular repository. GSAC code provides an extensive working framework with rich functionality, and with the flexibility to implement as much or as little as desired. GSAC is ready to use middleware between your database and file system and the user needing to query metadata or access data. For a data center to use this code requires only implementation of a configuration file specifying database access methods and the particular vocabulary or fields in your database schema to use in searches, plus writing a limited amount of code to specify the database queries as desired to support the desired search functions.

Like other web services, GSAC accepts incoming URL-based requests, handles the request, and returns a result; the result may be information, or data file access. The GSAC result return format for metadata is user selectable; HTML, XML, JSON, and csv return formats have been implemented. Much of the work in making a web service into a readily usable system is creating a browser-based UI, request handling, and output handling; GSAC has code that implements these functions. GSAC provides searches for sites or for files. The GUI interface to GSAC shows results in several ways, including lists of information, tables, maps, and file access mechanisms. For data access, GSAC can return a list of URLs, formulate a wget script for downloading data, start a download using Java webstart, and other options. There is also a GSAC command line client, a Java-based tool that allows users to do programmatic searches of a GSAC repository and to download files.

A GSAC implementation publishes on its web site the repository information XML, a capabilities document for the API search. This information is used by a federated GSAC implementation to work with other GSAC repositories. Four GSAC implementations are in production: the UNAVCO GSAC Repository, the SOPAC GSAC Repository at Scripps Orbit and Permanent Array Center, the CDDIS GSAC Repository at NASA's Crustal Dynamics Data Information System, and the Federated GSAC Repository at UNAVCO. The Federated Repository at UNAVCO provides searches of the UNAVCO, SOPAC, and CDDIS repositories in parallel, with aggregation of the results. Together these repositories have some 10,000 sites and 90 million data files.



Get GSAC for Your Geodesy Archive

Please contact one of the authors.

GSAC Web Site at UNAVCO

<http://facility.unavco.org/data/gsws/>

GSAC Web Site at SourceForge

<https://sourceforge.net/projects/gsws/>

Acknowledgements

The authors gratefully acknowledge support for GSAC from NASA through the ROSES ACCESS program Cooperative Agreement NNX10AF07A..