

Service Case Study AAI and eduGAIN for EIDA

Overview

The European Integrated Waveform Data Archive (EIDA) is a distributed data centre established to provide a secure archive for seismic waveform data and related metadata gathered by European research infrastructures [EIDA]. It provides the geoscience research community with transparent access to these archives via a single web portal. EIDA is part of the Earth Plate Observation System (EPOS), which incorporates different earth science communities.

EIDA consists of 10 data centres across Europe (in the Netherlands, Germany, Switzerland, France, Italy, Turkey and Romania) currently storing more than 300TB of data from approximately 140 networks worldwide. In each network, the data is collected by broadband sensors, short period sensors and accelerometers at over 5000 stations. Seismologists can search and download existing datasets via a single web portal. On average, a single user requests around 130 GB of data per day.

Challenge

To make the data from the different sources secure but accessible, a virtual network is set up. The network interconnects all data centres and appears as one archive to the end user. The data is stored locally and all data centres exchange metadata and routing tables. The information is updated daily.

Most of the data is, after registration, freely available via a web portal. Individual users can request data sets and download them via a desktop-client: the data is encrypted and password protected, with a new password required for each data centre. However, most user requests include data sets from multiple nodes, so multiple passwords need to be submitted every time to obtain the data.

The Geo Science community is working to enhance their international collaboration and sharing of data. As a subgroup of EPOS, EIDA is looking to expand the seismological data archive through further sites joining. It also considers the interoperability with other Geoscience communities which will be combined under the umbrella of EPOS. The underlying infrastructure and services provided by EIDA therefore need to be scalable, allowing the addition of new services in the future.

The objective is to develop a scalable AAI solution, simplifying access and data downloads while complying with seismological and security standards.

Solution

An initial assessment showed that all EIDA sites connect to the GÉANT network but only one site was already part of a national Identity Federation¹. Starting in January 2015, GÉANT worked closely with EIDA to examine their federated identity management requirements.

The role of GÉANT was to gather the requirements, assess the status, scope the work required, identify a suitable AAI service solution and to develop a roadmap for its deployment. The GÉANT User Support and the GÉANT AAI experts agreed with EIDA on a three-step approach:

¹ <https://refeds.org/federations/federations-map>

- Provide expertise and recommendations that allow EIDA to develop a prototype of the service portal that allows access to all data sets using the same credentials and download the data sets via the desktop client.
- Protect EIDA service portal with a SAML Service Provider and register it with an eduGAIN member federation.
- Motivate and assist GFZ, the main node of EIDA located in Potsdam (Germany) to join the DFN-AAI federation and deploy an Identity Provider.
- Help accelerating the adoption of eduGAIN-support within the organisations that operate EIDA nodes

By November 2015, seven out of ten EIDA nodes were supported federated login to eduGAIN users for their services. The prototype of the service portal was deployed and successful tested. The goal is to further work on documentation and finish registration in DFN-AAI until May 2016 when it is presented to the EIDA technical committee. With positive feedback from the technical committee, the prototype can be transformed to the production service.

Benefits

GÉANT supported EIDA to get familiar with federated identity management and in the mid term enable federated access for all European data centres and integrate them into eduGAIN while complying with seismological standards. This allows seismologists to use the credentials issued by their home institutions to securely access data provided by EIDA. The proposed solution is scalable and allows the extension to further seismological data centres and the integration of other Geoscience communities.