DELIVERABLE

D18.1 Report on Access Statistics and Service Provision of VA1-VA5

Work package	WP18: Access to seismological products and information at EMSC
Lead	EMSC
Authors	R. Basili, R. Bossu, L. Danciu, P. Kästli M. Koymans, M. Landès, G. Lanzano, S. Lasocki, M. Locati, L. Luzi, A. Rovida, R. Sleeman, M. Sobiesiak, P. Urban, R. Vallone, and D. Wenc
Reviewers	VA1: Paul Earle (USGS); VA2: Joachim Wassermann (Univ. of Munich); VA3: Hong Kie Thio, (AECOM); VA4: John Douglas (Univ. of Strathclyde); VA5: Marcelo Assumpção, (Univ. of São Paulo)
Approval	N/A
Status	Final
Dissemination level	Public
Delivery deadline	31.10.2018
Submission date	29.10.2018
Intranet path	DOCUMENTS/DELIVERABLES

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730900.

Summary

This report presents the mid-project results and performances of the 5 virtual access (VA) activities carried out within the SERA project. Under H2020 framework, VA ensures free-of-charge access to e-infrastructures delivering widely-used services (e.g., computing or communication infrastructure, data services…)in order to facilitate scientific research. SERA project includes 5 such VA:

- VA1: Access to parametric data and earthquake products operated by EMSC
- VA2: Access to seismic waveform data operated by ORFEUS/KNMI
- VA3: Access to the European Strong Motion database, the European Archive of Historical Earthquake Data, and the European Database of Seismogenic Faults operated by INGV
- VA4: Access to earthquake hazard and risk tools and products operated by EFHER/ETHZ
- VA5: Access to data and products of anthropogenic seismicity by IGPAS

The main objectives of the H2020 partial financial support to these activities are service improvement, development of their usage and integration in the EPOS (European Plate Observing System) initiative.

Each of these reports was prepared by the VA operator and presents the current state of the services and their usages. They all have been submitted to an external reviewer. Reports are compiled in D19.1.

SERA
Seismology and Earthquake Engineering
Research Infrastructure Alliance for Europe

DELIVERABLE ANNEX: VA1

VA1: Virtual Access to seismological products and information (EMSC)

Work package	WP18 - VA1: Virtual Access to seismological products and information at EMSC
Lead	EMSC
Authors	Matthieu Landès, EMSC; Rémy Bossu, EMSC
Reviewers	Paul Earle, USGS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730900.

Table of Contents

Summary

The aim of this document is to describe the current status of the services provided by the European-Mediterranean Seismological Centre (EMSC) in terms of data quality and traffic monitoring.

EMSC is a key actor for global earthquake information and is involve in the seismological community, in scientific European projects and in communication media for the general public. Since the initial SERA proposal, EMSC is constantly evolving. The volume of collected data from seismological institutes and eyewitnesses is increasing. Moreover, the same trend is observed on the traffic monitoring of the different EMSC services. On the websites, on Twitter, on the Seismic Portal or on the mobile application LastQuake, EMSC gains popularity among seismologists and among the general public.

This effort for providing good quality data and information imposes more and more constraints on the EMSC core real time system. It has to be faster and more flexible while maintaining its reliability. This work is now in progress.

1 Introduction

This chapter gives an overview of the European-Mediterranean Seismological Centre (EMSC) and describes the work done within the first 16 months of the SERA project.

1.1 Overview of EMSC

EMSC is one of the very top global earthquake information centres. All activities are closely coordinated with EMSC members (85 institutes and observatories in 56 countries) as well as with US Geological Survey and with the International Association for Seismology and Physics of the Earth's Interior (IASPEI). Finally, EMSC is one of the pillars of the seismological services of the ESFRI[1] research infrastructure EPOS (European Plate Observing System).

EMSC provides real time earthquake information and earthquake products ranging from authoritative locations, moment tensors, global macroseismic data, to information related to earthquake's impact (qualitative impact estimates, mapping of the felt area, rapid detection of felt earthquakes, geo-located picks). EMSC also hosts the RESORCE[2] database, a reference database for specific Ground Motion Prediction Equations (GMPE) studies as well as a Quake Catcher Network (QCN) server to ease and favour the developments of citizen operated networks in the Euro-Med region.

The EMSC infrastructure can be divided in a collection part and a dissemination part (Figure 1).

- The **collection part** collates seismological data collected from national institutes and data collected from eyewitnesses that share their experiences through felt reports, comments or pictures.
- The **dissemination part** comprises EMSC classic mobile and desktop websites, the Seismic Portal for accessing and visualizing seismic data, accounts on social networks (Twitter, Facebook, Telegram) and a mobile application "LastQuake".

[1] *European Strategy Forum on Research Infrastructures*
[2] Reference database for Seismic grOund-motion pRediction in Europe

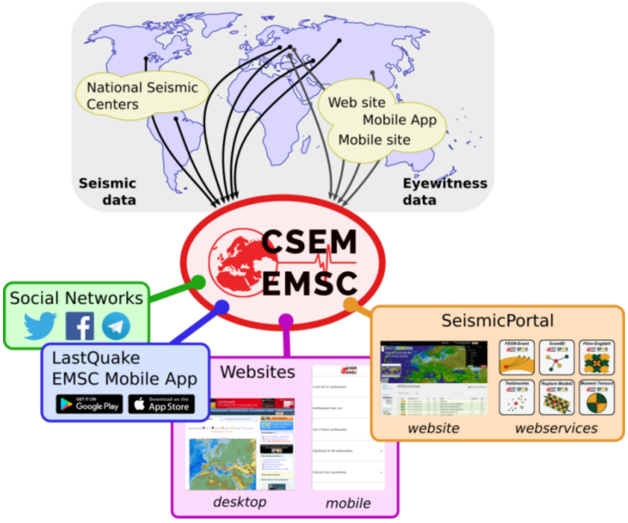


Figure 1: Overview of the main EMSC services

1.2 Description of work

1.2.1 Strengthening the EMSC communication system

The communication system of EMSC is one of the key components to extend and strengthen the EMSC services. For many years, for **data exchange with contributors**, we use emails and the messaging system PDL[3] developed by the USGS. Within the European project EPOS, EMSC has added the new HMB messaging system developed by GFZ that works on the TCP port 80**. On the user side**, the EMSC gives access to seismological data via 6 web services through the Seismic Portal. The FDSN-event exists since 2014 but the 5 others are new and have been released in the middle of 2017. In addition, the EMSC communicates real time information with a Twitter quakebot. Messages sent to Twitter are constantly evolving and participate to the increasing popularity of EMSC.

1.2.2 Accessibility for eyewitnesses and seismologists

The initial proposal to separate website for eyewitnesses from the general one has evolved. We have noticed that a higher share of eyewitness reports are done through our mobile app LQ and the mobile website (basically : through smartphones) . Although LastQuake is constantly developed, **our mobile website needs to be upgraded** and it's becoming a priority. This work will be initiated at the end of 2018. In addition we plan to move more scientific contents to the Seismic Portal. Moreover, on the hardware side, **the front end server of all EMSC services has been upgraded** thanks to the funding of the CEA that host the EMSC. This server is well designed to face huge traffic peak of user during seismic crisis without any slowdown.

Concerning the Quake Catcher Network (QCN), the EMSC has developed a QCN server available in 2016. The website is accessible here: http://qcn.emsc-csem.org/ and it allows users to visualize and query QCN data (Figure 2). With the design of the sensors, the recorded waveforms are very sparse and with the lack of funding, this project is less popular and its future is compromised. On the topic of citizen seismology sensors, we are testing the *Raspberryshake* sensor-digitizer developed by OSOP[4].

[3] Production Distribution Layer developed by USGS, https://github.com/usgs/pdl
[4] http://www.osop.com.pa/

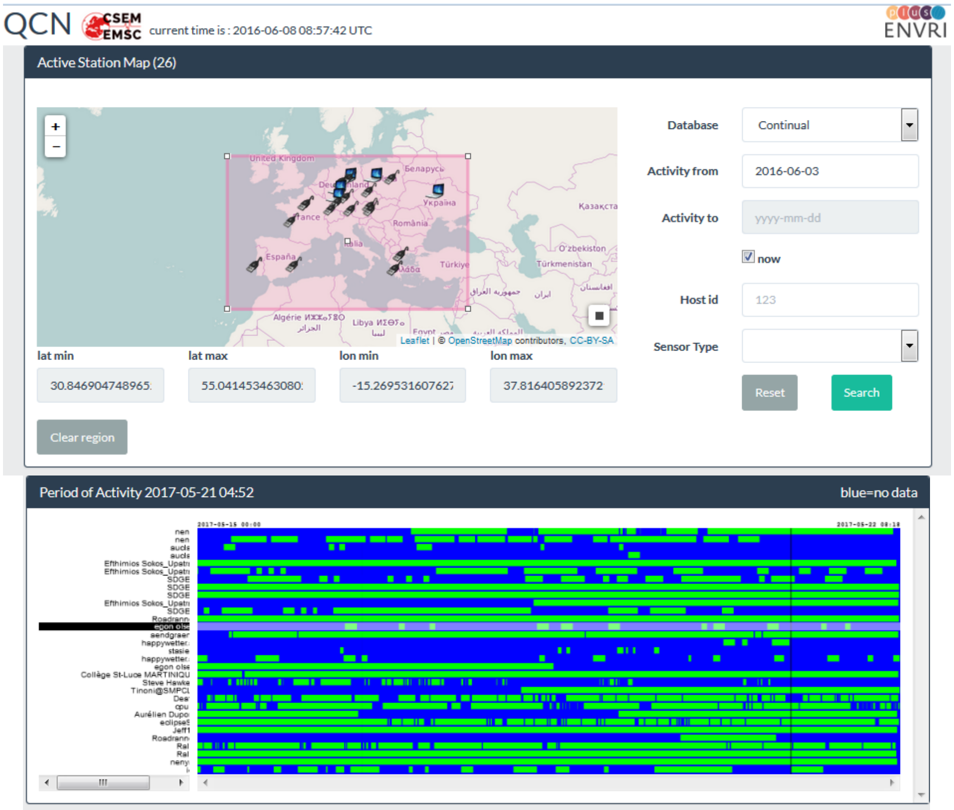


Figure 2: Graphical user interface of the EMSC QCN server. The user has the possibility to search for sensors (top) and to visualize data availability.

Within the EPOS project, the **Seismic Portal** has evolved and now contains new services that give access to felt reports, moment tensors, source models, mapping of event identifiers and Flinn-Engdahl region names. For each dataset, a graphic user interface and a web service are available.

1.2.3 Performance and reliability of the EMSC real time system

In the effort to strengthen its current services, the EMSC has begun an upgrade of its real time system in the beginning of 2018. This system is a legacy of 15 years of continuous developments and keeps historical informatics choices. The actual EMSC real time system has the constraints for fast data processing, for fast communication and for more flexibility for new developments.

2 EMSC collection system

The EMSC **collection system** collates seismological data and interfaces with global earthquake eyewitnesses in order to massively crowdsource testimonies, comments and geo-located pictures.

Note that in this section, statistics and measurements are done between January 2017 and August 2018.

2.1 Seismological data

Seismic data are collected in real time from 86 observatories (in 2017) from around the world. On average, we collect 15k seismic origins per month and 230 moment tensor solutions per month representing information for at least 4000 earthquakes monthly from all over the world (Figure 3). This set of data is the foundation of all EMSC dissemination services and is widely used by the seismological community.

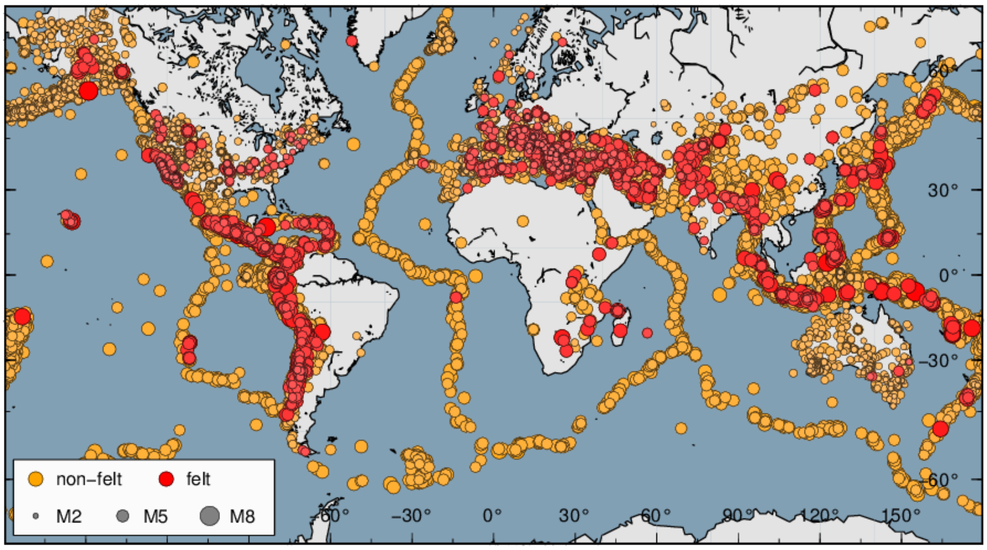


Figure 3: Earthquakes collected by the EMSC between January 2017 and July 2018. During this period, we have collected 104283 earthquakes among which 3210 are considered as felt earthquake.

The overall quality of EMSC data relies on the quality of contributors and we started **to update contacts with contributors**. In particular, we renew contact with those who don't have sent seismic information. This active collaboration is important for EMSC reliability.

2.2 Data from eyewitness community

Eyewitness data are collected through our websites and our mobile app. People who feel an earthquake and want to share their experience have the choice to evaluate the level of shaking with thumbnails (felt reports), to write a comment or to send pictures. The following data was collected between 1st January 2017 and July 2018.

The map of all felt reports (Figure 4) shows a world wide distribution. EMSCcovers almost all populated seismic regions even though in North Africa and Eastern Asia a lack of popularity can be observed. In any case, this world wide distribution gives us confidence in our ability to engage population wherever damaging earthquakes may strike.

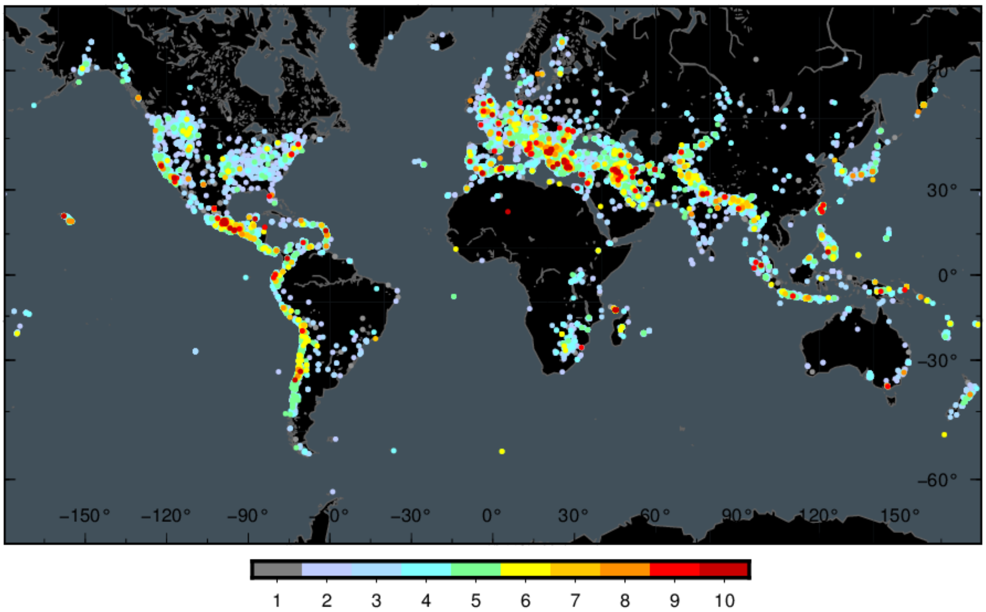


Figure 4: Distribution of the 160k felt reports collected by EMSC between January 2017 and July 2018. The colors scale represents the level of shaking reported by eyewitnesses. During this period felt reports

Felt reports is our main indicator measuring EMSC popularity with the general public. Within 19 months, we have collected more than 162k reports and 64% are collected through our LastQuake app. Although the time distribution depends on seismic activity (Figure 5), the general trend shows that we collect not only a higher number of felt reports, but also we collect them faster. For instance, at the end of June 2018, following the earthquake sequence in Mayotte we have collected more than 20k felt reports in 15 days and 70% are collected in 15min.

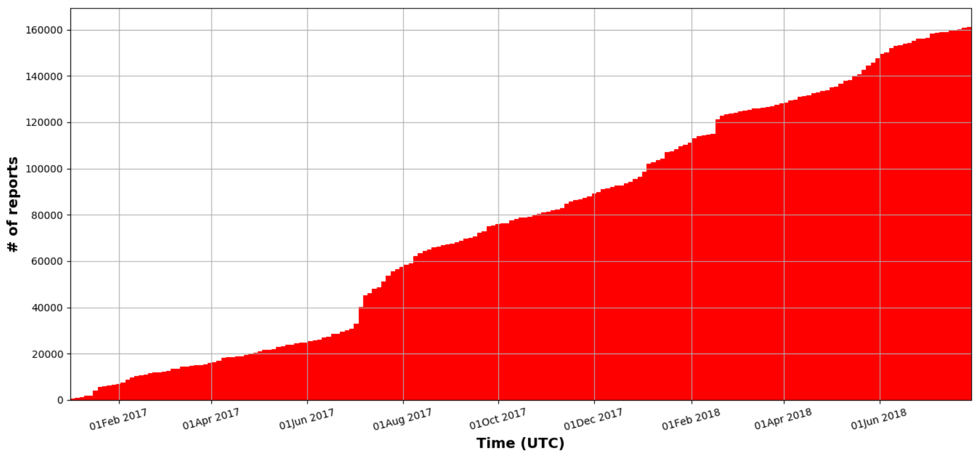


Figure 5: Monthly cumulative time distribution of felt reports collected by the EMSC.

Eyewitenesses often share their feelings with **comments**, especially during seismic sequences. We collected more than 66k comments (71% are from our app). **Pictures are a lesser popular way to share experience after an earthquake through our tools. We only collected** 310 pictures for 50 events.

3 Traffic monitoring of EMSC dissemination system

On one hand, the dissemination part targets **seismologists and researchers** with dedicated web pages (e.g. "for seismologists only page") accessible trough EMSC websites, the Seismic Portal and its visualization capacities and interactive accesses, webservices. On the other hand, the EMSC targets also **general public and eyewitnesses** with communications on social networks like Twitter, and with the EMSC mobile aplication LastQuake.

3.1 Social networks

The EMSC is present on Twitter, Facebook and Telegram. The main media remains Twitter with 75,5k followers. It has to be compared with the 25k Facebook fans and the 273 telegram members. Twitter is the media for automatic publication that inform in real time for felt earthquakes.

The number of Twitter followers is constantly increasing. On August 20th 2018, the @Lastquake account had 75,5k followers, compared to 63,2k in December 2017. The monthly evolution of views of EMSC tweets shows that the EMSC popularity on Twitter counts in million (Figure 6).

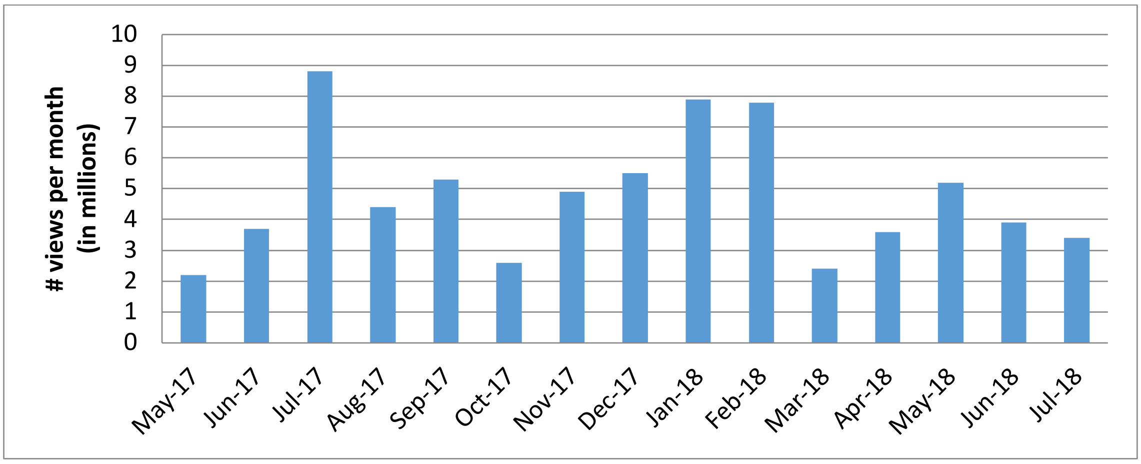


Figure 6: Time distribution of the number of views per month on the @LastQuake Twitter account of EMSC.

3.2 Mobile application: LastQuake

The LastQuake mobile application is an important component of the EMSC communication system. At the end of August 2018, there are almost 304k applications in operation, split into 117k for IOS and 187k for Android. Almost 20% of total users are active each month (Figure 7). Compared to Twitter or websites (see below in section 3.3), number may seem to be less important… However the mobile application allows a direct communication between the user and EMSC. Since each user is a potential witness of seismic activity, the engagement of user is higher than other media for sharing their experience.

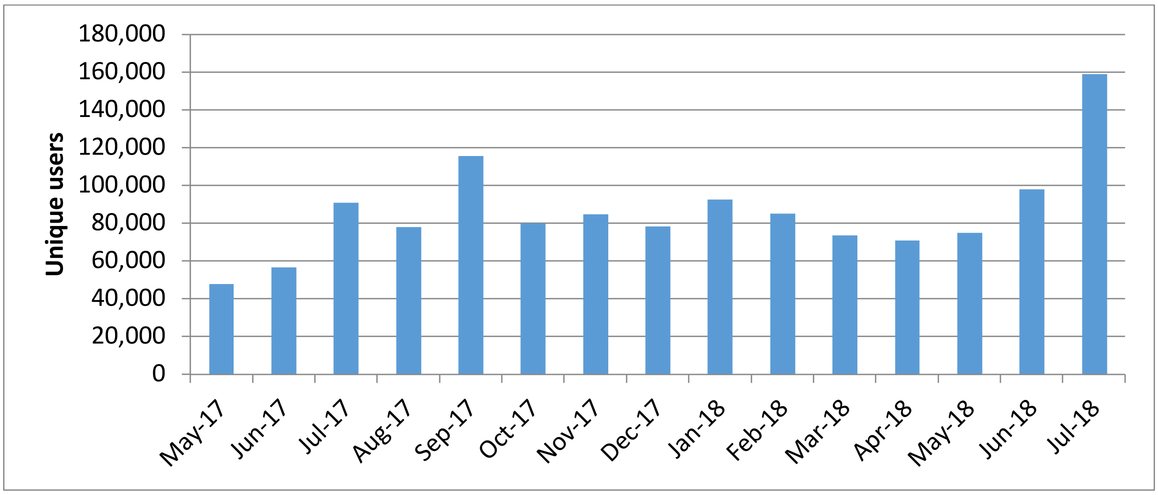


Figure 7: Time distribution of unique users per month that have launched the LastQuake application.

For EMSC, LastQuake is a worldwide success. The application is translated in 16 languages and is used in almost every country (Figure 8). The monitoring of LastQuake launches showed that it has become a system able to detect significant earthquakes in most regions.

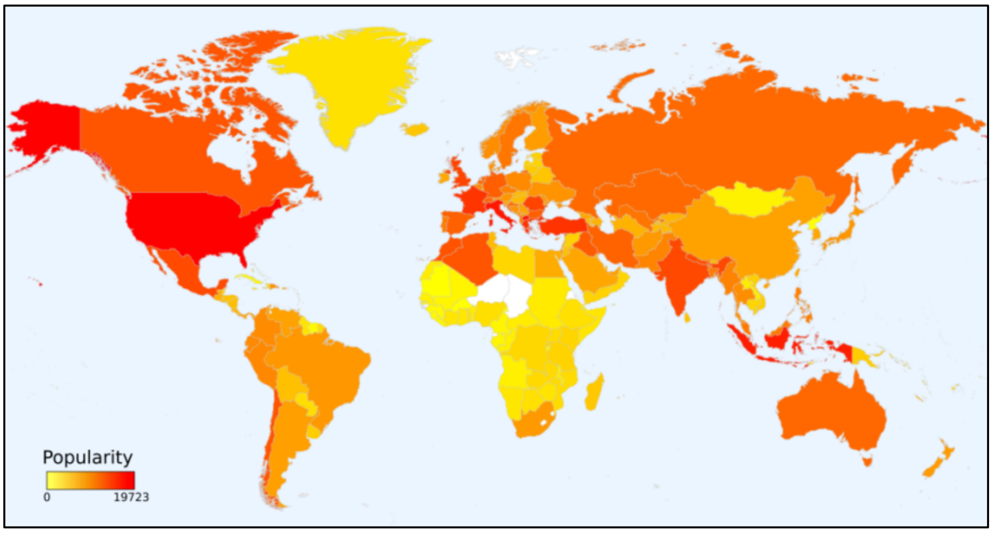


Figure 8: Spatial distribution of LastQuake users in late August 2018. The top 1 country is the United State with 19924 LastQuake applications in operation.

3.3 Desktop and mobile Websites

Desktop and mobile EMSC websites (www.emsc-csem.org and http://m.emsc-csem.org) are the traditional place to find our real time data, to search for earthquakes information and to have more information about EMSC. Contrary to Twitter or the mobile application LastQuake, websites are often used by seismologists, in particular the "for seismologists only page"[5].

As shown by traffic curves (Figure 9), they continue to be widely used with more than 800k unique visitors per month for desktop website and almost 200k unique visitors per month for mobile website. All traffic statistics estimations are performed with StatCounter[6].

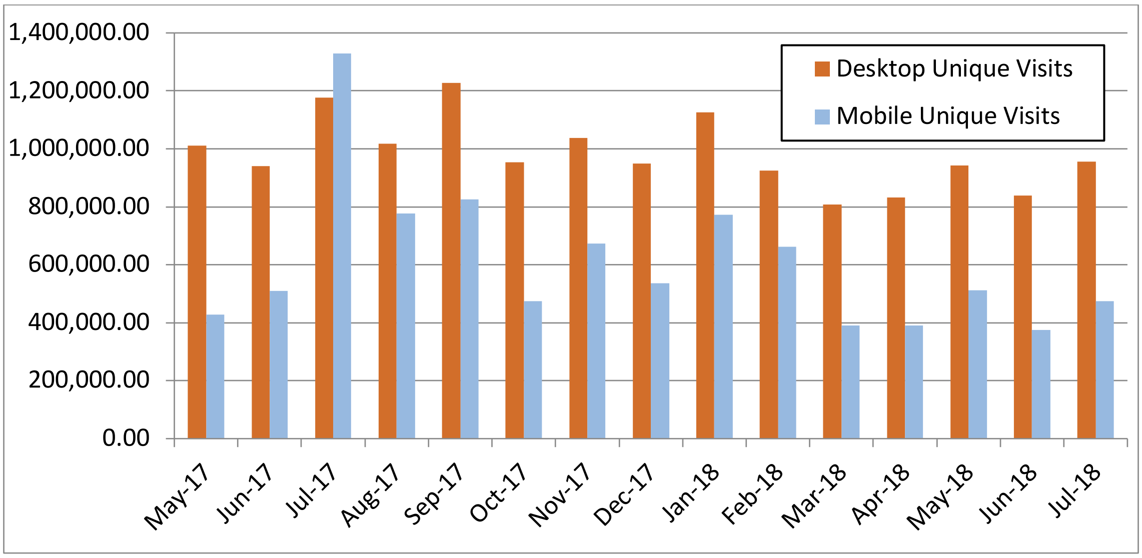


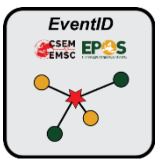
Figure 9: Traffic distribution of monthly unique user on the desktop and the mobile website of the EMSC.

3.4 Seismic Portal

The Seismic Portal is the EMSC portal that gives access to seismic data (www.seismicportal.eu). Users have the choice to use graphical interface or web services for automatic processing.

Most of them have been developed within the European EPOS project and are now operational. To learn how to use them, some use cases are available at https://github.com/EMSC-CSEM/webservices101.

On the Seismic Portal, the EMSC provides 7 services listed below :

	The EventID web service maps dynamically event identifiers to allow the identification of a same event between different seismological institutions.

[5] https://www.emsc-csem.org/Earthquake/seismologist.php
[6] www.statcounter.com

	Web service conforms to the FDSN-Event standards and providing all the EMSC event data available. Event information can include all origins and all arrivals as desired.
	The service identifies the Flinn-Engdahl region from a geolocalisation entry point.
	Web service that gives access to the moment tensors collected at EMSC
	This service allows downloading all felt reports collected from eyewitness during earthquakes through EMSC websites and LastQuake mobile application.
	Near realtime notification of new and updated earthquake event can be received using the WebSocket protocol. Any WebSocket client can connect to our service to be notified. Javascript, Python example codes are provided to demonstrate the service.
	The web service allows recovering all rupture models from the SRCMOD database of Martin Mai (which is the database of finite-fault rupture models of past earthquakes). These earthquake source models are obtained from inversion or modeling of seismic, geodetic and other geophysical data, and characterize the space-time distribution of kinematic rupture parameters.

The **traffic monitoring of Seismic Portal** services is measured in terms of “hit per month” and “unique IP per month”. We use logs from our front end server to generate traffic measurements. Our measurements are discontinuous. In the beginning of 2018, we had an infrastructure upgrade and a configuration update at the end of May 2018 and that explain the lack of measurement at these periods.

The **FDSN-event service** is the older service available and it is operational since March 2014. In 2017, the service has gained a large popularity with almost 400k unique users per month (Figure 10). The volume of data transferred via the service is of the order of a hundreds of gigabytes per month.

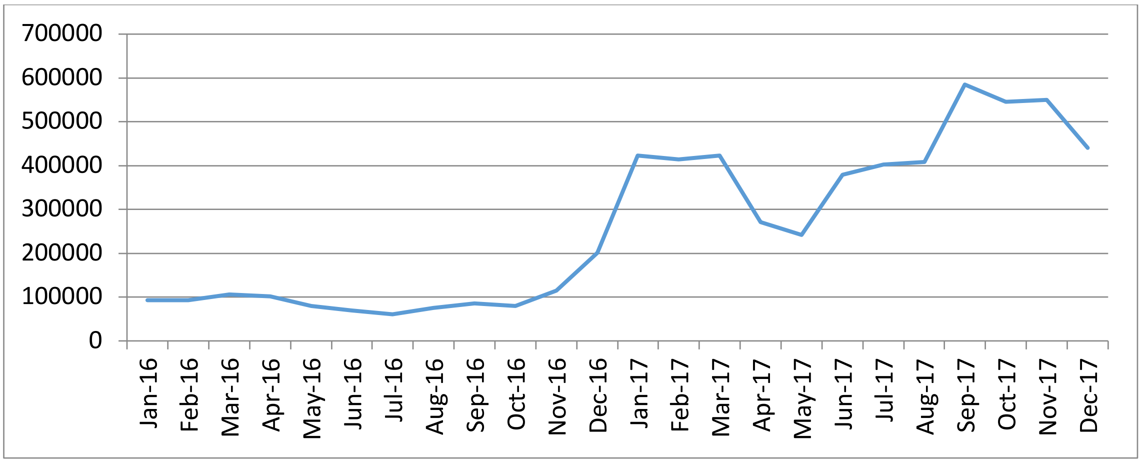


Figure 10: monthly unique user of the FDSN-event webservice available on the SeismicPortal.

For the other services we only monitored traffic for 3 months of traffic monitoring. Due to the configuration update of our front end server in May 2018, only the first 15 days of the month are available (numbers are in italic). The traffic of these new services is lower. Of course, they are new so we don't expect the traffic to be comparable as for the FDSN-event service. Moreover, these new service provides more specialized data that interested mostly seismologists.

	fdsn-event		Moment Tensor		Testimonies		Eventid	
Date	Hit	Unique ip	Hit	Unique ip	Hit	Unique ip	Hit	Unique ip
mars-18	22951768	332969	4543	389	5872	373	1311	110
avr-18	27069156	389820	3851	586	9618	597	1362	118
mai-18	*12408609*	*216970*	*2897*	*294*	*9782*	*298*	*530*	*46*

	Srcmod		Near Real Time		Flinn-Engdahl lookup	
Date	Hit	Unique ip	Hit	Unique ip	Hit	Unique ip
mars-18	371	68	9346	458	44760	17
avr-18	310	59	9888	542	53508	25
mai-18	*167*	*31*	*2817*	*255*	*2635*	*14*

4 Upgrade of the EMSC core system

Observations on traffic monitoring measurements show the need of increasing speed for delivering information to the community. For instance, some comments on Twitter and the response rate of reports show that after 30 minutes, eyewitnesses are not "hype" by the event they have felt. Aware of this constraint of fast and reliable information, the EMSC has initiated an upgrade of his real time system. It includes **a hardware upgrade of production servers** with CEA funding and it includes also **a rewrite of its software toolchains**.

The EMSC real time system has a legacy of 15 years of continuous development. It's a homemade system and with years, new features have been added without modification of the core. In addition to this complexity, this system is critical since it's also used for the French seismic alert… Every modification is very sensitive.

It's now a necessity to modernize the system. Step by step, we started to setup a versioning of the system with git[7], to identify and cleanup unused codes, to optimize the parsing of contributors data (e.g. quakeml) and to refactor the structure of the system to easily re-deploy the system. It's a long term work that may bring visible changes within a year.

[7] https://www.gnu.org/software/git/

5 Conclusion

EMSC gives access to all collected seismic and eyewitness data and takes care of traffic measurements of its services. The traffic monitoring shows the overall popularity of the EMSC among the general public and the seismological community. All components of the system, put together, participate to bring fast, reliable and good quality information.

- Some services like desktop and mobile websites and FDSN-event web service are now mature and represent important traffic measurements. New web services like moment tensor, testimonies or eventid have not yet found their users. However they have to be associated to the effort in sharing collected data to the community.
- The increasing popularity of @LastQuake Twitter account and of the LastQuake mobile application show the importance for the EMSC to address their communication to general public and not only to seismologist.
- The increasing volume of data collected adds the constraint for the long term support to update the workflow of real time system.

SERA
Seismology and Earthquake Engineering
Research Infrastructure Alliance for Europe

DELIVERABLE ANNEX: VA2

VA2: Access to seismic waveforms at ORFEUS/KNMI

Work package	WP19 VA2: Access to seismic waveforms at ORFEUS/KNMI
Lead	KNMI
Authors	Reinoud Sleeman, Mathijs Koymans - KNMI
Reviewers	Joachim Wassermann (Univ. of Munich)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730900.

Table of Contents

Summary

This report summarizes the provision of the main European services for virtual access to seismic waveform data, related metadata and products for seismology and engineering seismology, through the ORFEUS infrastructure. Through the use of standardized services (e.g. webservices) and GUI's our virtual access services are aligned and compatible with EPOS and EPOS ICS. Moreover, interactive services like StationBook and RRSM are being re-designed with modern development technologies to offer flexibility to add new functionalities (e.g. new types of data or metadata) and tailor it to on-going, changing requirements.

1 Introduction

The ORFEUS infrastructure is one of the largest infrastructures in the world that provides seismological data and derived products to the scientific research community in strong collaboration with European seismological observatories. The infrastructure is organized as a networked system of observatory infrastructures, waveform data archives and services. A key component is the federated, distributed European Integrated waveform Data Archive (EIDA) that transparently connects a number of large data centers in Europe, including the ORFEUS Data Center. This unique, federated archive serves seismological waveform data from permanent and temporary networks of broad-band sensors and strong motion sensors deployed in Europe and beyond through dedicated services. Currently, EIDA holds around 400 TB of data of about 100 permanent networks and 100 temporary networks, with a total of more than 8000 seismic stations. Through EPOS-IP, in which ORFEUS is strongly involved, and being compatible with EPOS ICS we foresee that EIDA will extend to serve other data types (OBS, NFO) to a broader user community (e.g. earthquake engineering).

Services that are being offered to the (seismological) research community to provide (virtual) access to raw waveform data and related metadata are: a) ORFEUS website, b) EIDA data portal, c) EIDA webservices, d) RRSM (Rapid Raw Strong Motion database) and e) StationBook.

Specific objectives of this work package as described in this report are:

- Coordination with NA2 activities on extending EIDA to support other types of data and to serve a broader geoscience community and the engineering and hazard communities (e.g. Implementation of data model extension; Near Fault Observatories or structural monitoring arrays).
- Services offered and developed by this infrastructure for: flexible and transparent access to raw waveform data in EIDA; related metadata (station and data quality); derived data and products; refined data discovery across EIDA; documentation and outreach.
- Access through the EIDA GUI and standardized webservices.
- Access to EIDA StationBook and the RRSM.

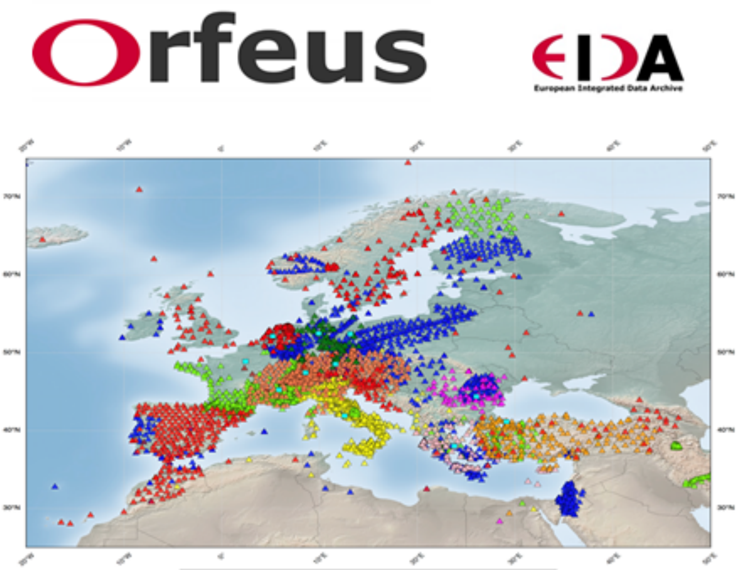


Figure 1: The ORFEUS / EIDA infrastructure provides access to raw seismological waveform data and related metadata from more than 8000 sensors throughout Europe and beyond. The federated data archive EIDA connects 10 large data archives (holdings are geographically identified by colour).

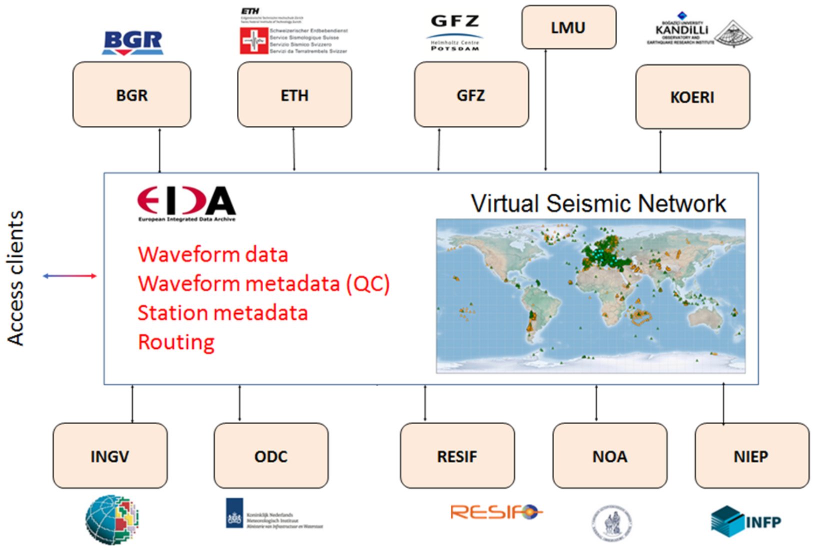


Figure 2: Standardized services that are deployed across EIDA to provide uniform and transparent (virtual) access to all data holdings.

2 ORFEUS/EIDA services

The services provided can be divided in the categories webservices and interactive services, which are listed here and further described in chapters 2.1 and 2.2.

1. Webservices
    1.1. fdsnws-dataselect
    1.2. fdsnws-station
    1.3. eidaws-routing
    1.4. eidaws-wfcatalog

2. Interactive services
    2.1. ORFEUS website
    2.2. EIDA GUI
    2.3. RRSM
    2.4. StationBook

3. In addition to the above services a number of clients have been developed at ORFEUS Data Center (ODC). These will be described in section 2.3.

2.1 Webservices

ORFEUS EIDA implements the following webservices to provide standardized and open access to seismological (waveform) data. The specifications and the usage of parameters of each service can be found at the appropriate page (through the link):

- fdsnws-dataselect - FDSN standardized webservice for mini-SEED waveform data.
- fdsnws-station - FDSN standardized webservice for station metadata.
- eidaws-routing - EIDA standardized webservice for routing between EIDA services.
- eida-wfcatalog - EIDA standardized webservice for waveform metadata.

ORFEUS EIDA consists of multiple data centers with unique data holdings and webservices. Data exposed at one data center may not be available at another, therefore the appropriate node should be selected in your request. Please consult the EIDA networks page to discover the appropriate node(s) for data requests and citation.

Within EIDA three other services are being developed and tested in order to optimize harvesting of data in a complex, federated system like EIDA and to enable users to easily collect open and restricted data:

• the EIDA federator webservice (beta version) uses fdsnws-station, fdsnws-dataselect, and eidaws-wfcatalog requests across all EIDA nodes to enable users to collect data without *a-priori* knowledge of where data is hosted.

• the EIDA mediator webservice will be designed for advanced selection of data across EIDA based on user criteria (e.g. quality parameters).

• the EIDA authentication webservice is a central authentication system that provides tokens for all EIDA services across EIDA. The authentication System connects to a B2ACCESS service (provided by the EUDAT Collaborative Data Infrastructure). This webservice is being tested.

EIDA Node	FDSNWS-Dataselect	FDSNWS-Station	EIDAWS-Routing	EIDAWS-WFCatalog
ODC	Online 1.1.1	Online 1.1.1	Online 1.1.1	Online 1.0.0
GFZ	Online 1.1.1	Online 1.1.1	Online 1.1.1	Online 1.0.0
RESIF	Online 1.1.0	Online 1.1.0	In development	Online 1.0.0
INGV	Online 1.1.1	Online 1.1.40.3	Online 1.0.4	Online 1.0.0
ETHZ	Online 1.1.1	Online 1.1.1	Online 1.1.1	Online 1.0.0
BGR	Online 1.1.1	Online 1.1.1	Online 1.1.0	Online 1.0.0
NIEP	Online 1.1.0	Online 1.1.0	In development	Online 1.0.0
KOERI	Online 1.1.1	Online 1.1.1	Online 1.1.1	Online 1.0.0
LMU	Online 1.1.0	Online 1.1.0	Online 1.0.3	Online 1.0.0
NOA	Online 1.1.0	Online 1.1.0	Online 1.1.1	Online 1.0.0

Figure 3: Status of the webservices across EIDA. Each EIDA node runs the same services (fdsnws-dataselect, fdsnws-station and eidaws-wfcatalog) and optional the routing service. When consulting this webpage (https://www.orfeus-eu.org/data/eida/webservices/) each service is queried across EIDA in order to reflect the current status.

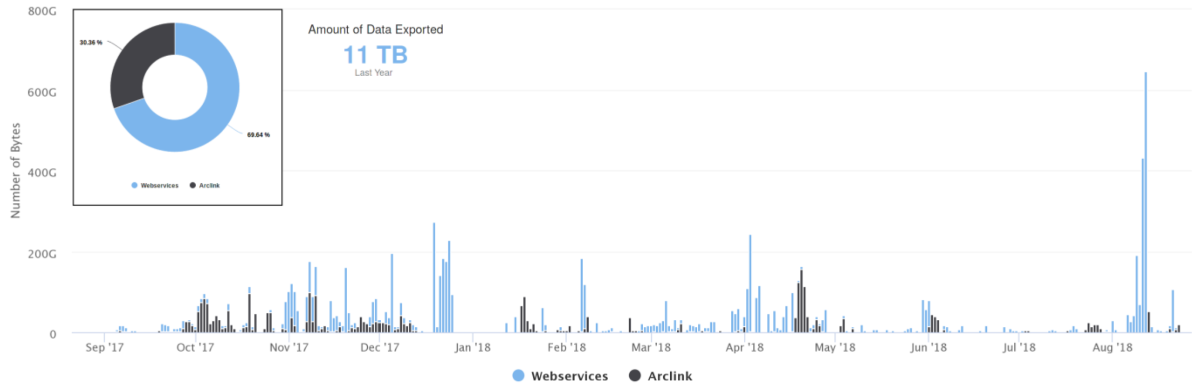


Figure 4: Amount of data (GBytes) exported by ODC in the past 12 months through webservices and ArcLink (deprecated service). A total of 11 TB has been exported to the research community.

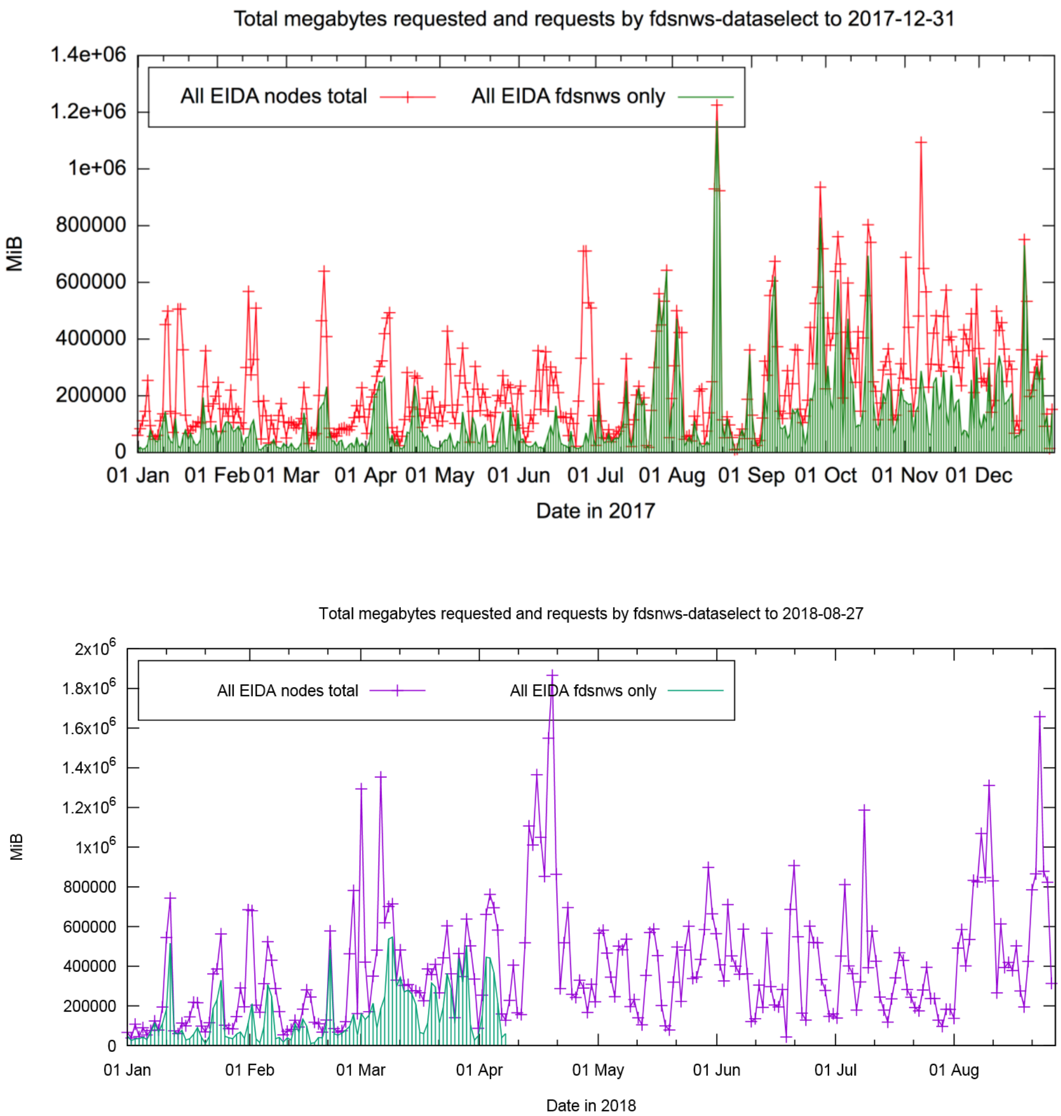


Figure 5: Amount of data (Bytes) exported by EIDA throughout 2017 and 2018.

See the Advanced Workflow Examples page or the ORFEUS EIDA Webservices Notebook for a more extensive explanation on how to use the webservices.

2.2 Interactive services

Interactive services that are provided by ORFEUS EIDA to serve waveform data and information are :

- ORFEUS website - the landing pages for all information concerning ORFEUS, EIDA and services.
- EIDA web interface - the GUI to interactively search for and download data from EIDA.
- StationBook - the GUI to access all (available) information on seismic stations across EIDA.
- RRSM web interface - the GUI to search for and collect strong motion products in near real time.

2.2.1 ORFEUS website

The ORFEUS website is the entry point for the variety of services and information to collect seismological waveform data hosted by European data archives.

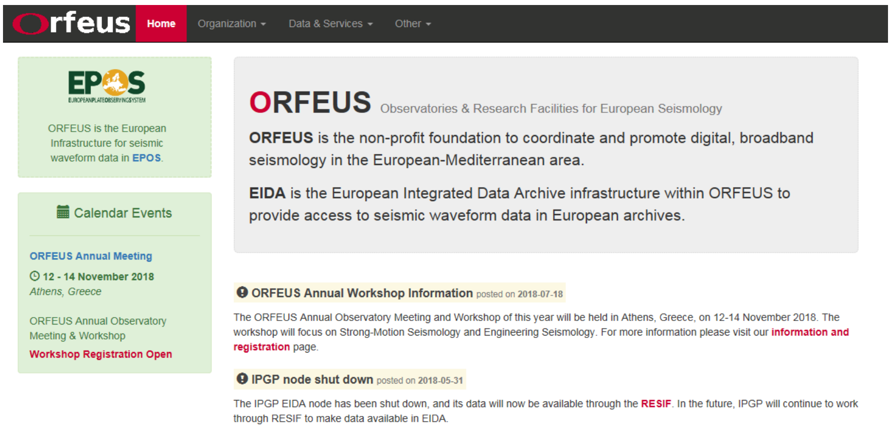
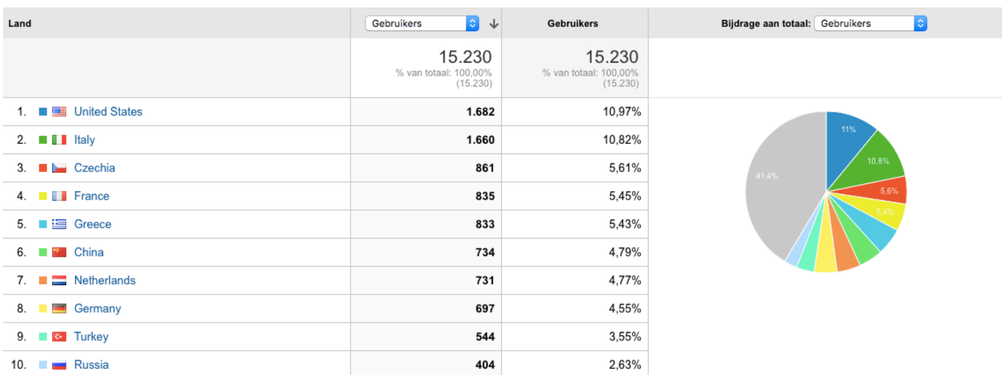


Figure 6: The ORFEUS website 'landing page'.



Land	Gebruikers	Gebruikers
	15.230 % van totaal: 100,00% (15.230)	15.230 % van totaal: 100,00% (15.230)
1. United States	1.682	10,97%
2. Italy	1.660	10,82%
3. Czechia	861	5,61%
4. France	835	5,45%
5. Greece	833	5,43%
6. China	734	4,79%
7. Netherlands	731	4,77%
8. Germany	697	4,55%
9. Turkey	544	3,55%
10. Russia	404	2,63%

Figure 7: Numbers and percentages of different users invoking the ORFEUS website. Notice the impact of ORFEUS outside of Europe (e.g. United States and China).

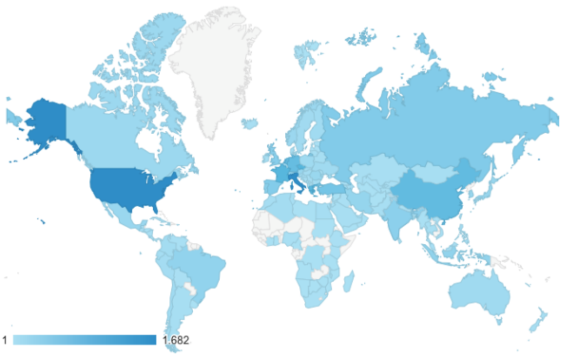


Figure 8: Geographical distribution of users across countries.

2.2.2 EIDA webinterface (GUI)

The EIDA webinterface, originally developed by GFZ, is installed at a number of EIDA nodes and demonstrates the transparency for the community to search for and download data from the federated archive EIDA.

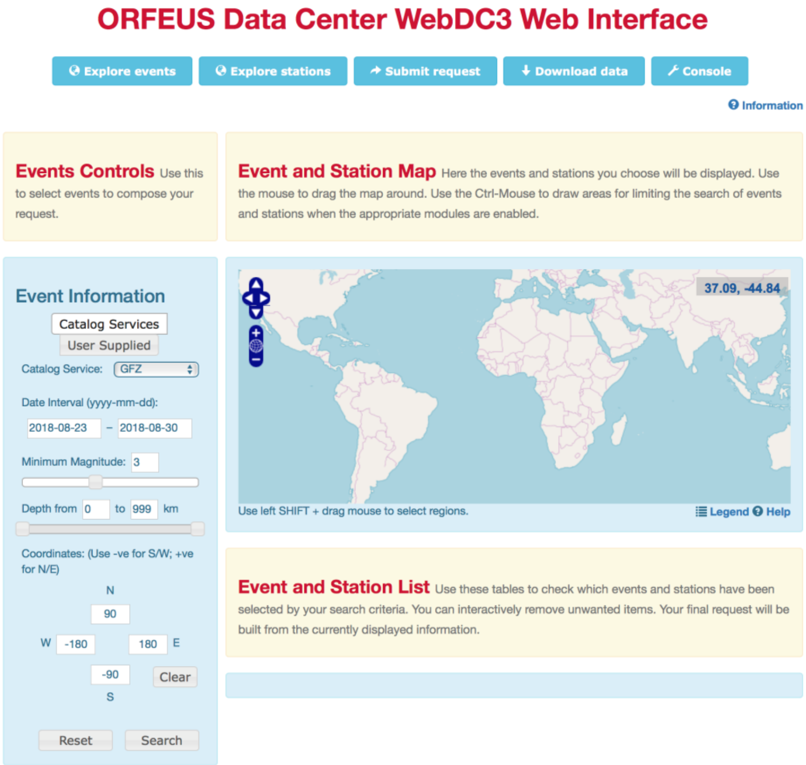


Figure 9: EIDA webinterface.

A new design of the EIDA interface (still in development) is being prototyped to provide an EIDA access point allowing users to browse stations, events and request waveform data using extensive filtering mechanisms. By design one of the aims of the new EIDA interface is to make it modular and open for extensions (e.g. new data types and metadata models). With modern development methodologies favouring micro-services, loosely coupled modules and separation of concerns it will be easy to add new functionalities and tailor it to rapidly changing requirements.

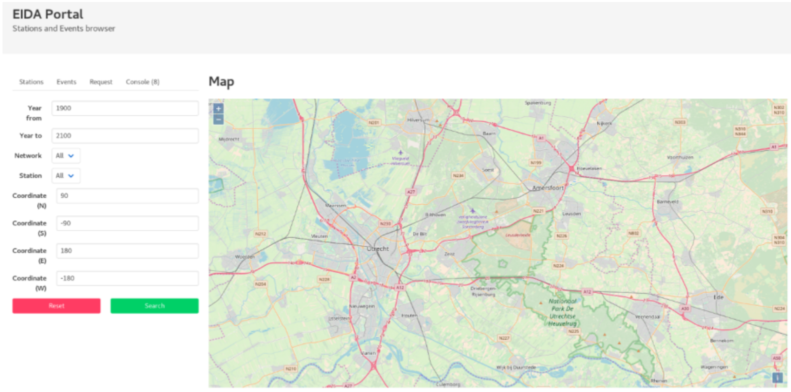


Figure 10. New EIDA interface home page prototype (still in development).

2.2.3 Station Book

Station Book is a web application built on top of FDSN and EIDA web services with its own backend logic and database. It is intended to be the complete, interactive catalogue of EIDA stations containing extended station metadata not covered by the FDSN specification (operator notes, descriptions, photos, comments). Users can register themselves but need to be given write access to a network to be able to start editing its stations. Both backend logic and underlying structure of the StationBook has been redesigned and upgraded significantly (e.g. Python3) this enables flexibility to incorporate new types of extended station metadata (e.g. new database tables). Data ingestion is in two ways:

- common network and station metadata: collected automatically from EIDA.
- station and site characteristics: added/edited by network operators.

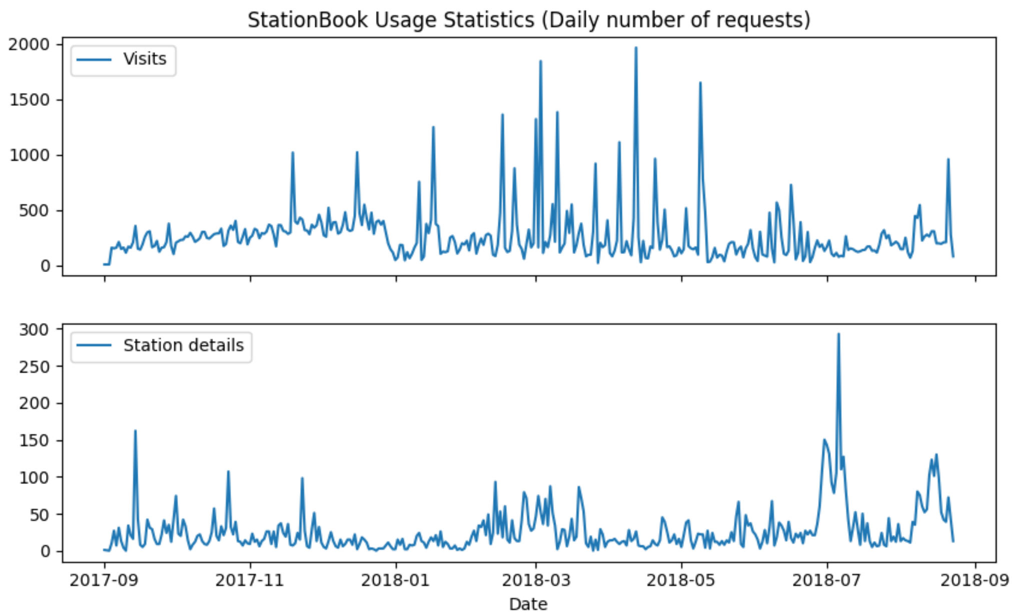


Figure 11. Daily number of requests made to the StationBook in the past 12 months, with on top the number of daily visits to the portal and below the number of requests for detailed station information.

As the new StationBook was significantly renewed we provide below the important (technical) changes in the StationBook and a detailed description of the GUI in Appendix 4.1

- Backend written in Django 2.0 (Python 3)
- Frontend in Bootstrap 4 + OpenLayers
- Using Routing WS + FDSN WS from all EIDA nodes for data sync (basic station data)
- Has its own cache table for basic station data (Network, Code, Name, Latitude, Longitude, Elevation, Status, Start, End, Created)
- Channels and instrumentation data obtained in real time from appropriate node via FDSN WS
- Additional data (Owner, Morphology, Housing, Borehole etc.) stored in dedicated tables
- Lots of possibilities for extensions (additional data and functionalities like bulk data upload, comments, etc.)

2.2.4 RRSM

The RRSM portal allows users to query earthquake information, peak ground motion parameters, response spectral amplitudes and to select and download earthquake waveforms within minutes after an earthquake with magnitude ≥ 3.5 occurring in the European-Mediterranean region [1]. Earthquake information is provided by the EMSC and all on-scale seismic waveform data available from ORFEUS EIDA is considered for fully automated processing. Real-time RRSM processing started in June 2014. Offline reprocessing was carried out for all M ≥ 4.5 events that occurred since January 2005, and all M ≥ 3.5 events since January 2012.

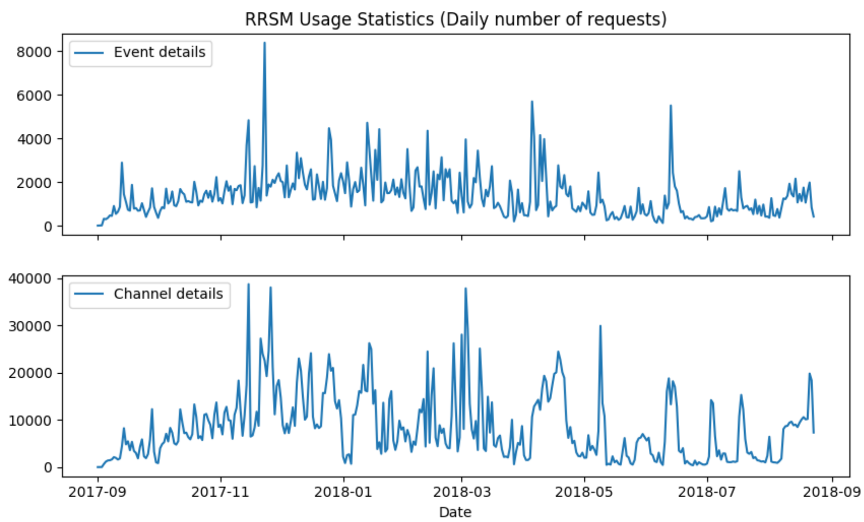


Figure 12: Daily number of requests made to the RRSM in the last 12 months. On top the number of requests made for generic event information. The bottom graph shows the number of requests for detailed strong motion values.

The RRSM GUI is a visual representation of information available in the RRSM database and through RRSM web service. Just as with the StationBook the RRSM backend logic and underlying structure has been redesigned and upgraded significantly. We provide below the important (technical) changes in the RRSM as well as a detailed description of the GUI in the Appendix 4.2:

- o Backend written in Django 2.0 (Python 3)
- o Frontend in Bootstrap 4 + OpenLayers for maps
- o Using ODCWS RRSM web service

2.3 Clients

ORFEUS Data Center developed a number of specific clients (https://www.orfeus-eu.org/data/odc/) to display features like data latency, event waveforms and data quality parameters.

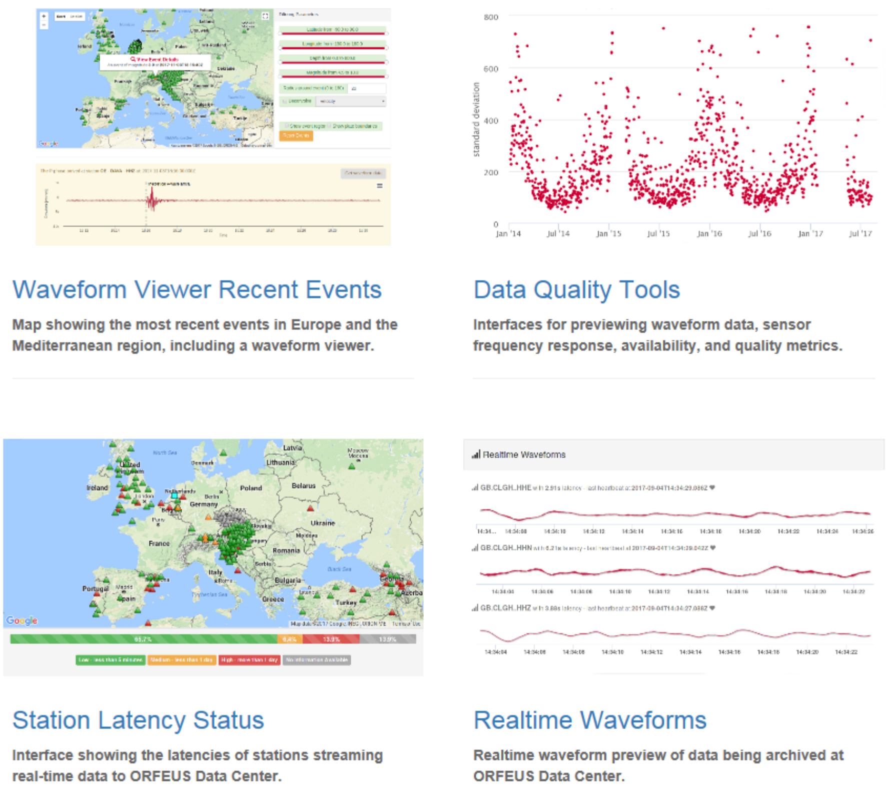


Figure 13: Example clients at ODC displaying a) earthquake waveforms, b) data quality parameters, c) data stream latencies and near real-time waveforms. Also clients to display data availability, instrument responses, data export statistics and statistical values of the data samples are available.

Besides visual interfaces to waveforms, the EIDAWS-WFCatalog [2] provides an API that exposes a waveform metadata catalogue for the seismic archive at an EIDA node. The WFCatalog Webservice provides detailed information on the waveform data like quality parameters (derived from data record headers, e.g. timing quality and header flags) and statistical values (derived from the sample values, e.g. rms). The WFCatalog can serve as an index for data discovery (e.g. Mediator) as it has support for range filtering on all available metrics. The quality parameters are continuously calculated and stored in the WFCatalog database, enabling fast and efficient querying of these parameters (no on-the-fly calculations on the waveforms are needed). This enables for example fast computation of PDF’s for seismic data over long time frames (e.g. 1 year).

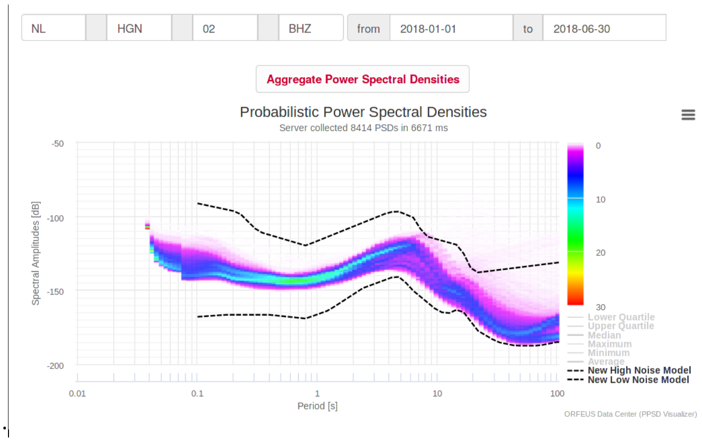


Figure 14: Example of the Probability Density Function calculated from 6 months of continuous seismic data, using the pre-calculated Power Spectral Density values within the WFCatalog database.

References

[1] Introducing the European Rapid Raw Strong-Motion Database. C. Cauzzi et. al., 2016, Seismol. Res. Lett. 87, 4, doi: 10.1785/0220150271

[2] Trani, Luca & Koymans, Mathijs & Atkinson, Malcolm & Sleeman, Reinoud & Filgueira, Rosa. (2017). WFCatalog: A catalogue for seismological waveform data. Computers & Geosciences. 106. 10.1016/j.cageo.2017.06.008.

3 Appendices

3.1.1 Station Book

The StationBook has 3 types of users:

- Guest users (non-registered users with read only access)
- Registered users (read only access directly after registration)
- Admins (users which can give registered users write access to network/station metadata)

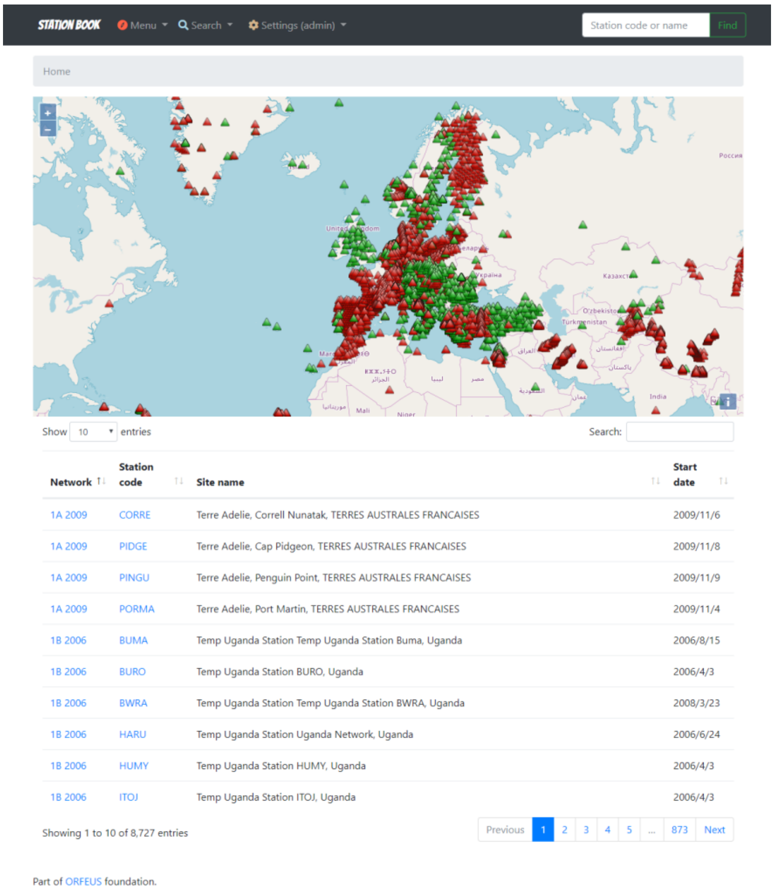


Figure A1: Station Book home page

The navigation bar located at the top of the screen contains a menu which allows users to quickly navigate to one of following pages:

Via "Menu":

- View all nodes (list of EIDA nodes hosting FDSN web services used by the Station Book)
- View all networks (available networks)
- Recent changes (list of recent changes of the stations metadata)
- Quick links
- About Station Book

Via "Search":

- Search (extended station search form)

Via "Settings":

- My account (editable account details e.g. name, e-mail, telephone)
- Change password
- Log out
- Admin panel (only when logged in with administrator rights)

On the right side of the navigation bar a "quick search" textbox can be found which can be used to quickly navigate to station by typing its code or site name.

The networks page contains a list of all networks archived at EIDA. Networks can be quickly filtered using the search box located above the table. The network details page contains network details and lists all stations belonging to this network. Stations can be quickly filtered using search box located above the table.

The station details page presents station information from the FDSN web service and extended information stored within the Station Book.

FDSN information (read only for all users):

- FDSN-Station Data
- FDSN-Station Instrumentation Data

Extended information stored within the Station Book (writeable for users with write access to the network from which given station originates):

- Basic Data
- Owner Data
- Morphology Data
- Housing Data
- Borehole Data
- Photos

Extended information can be edited using “Edit data” dropdown menu located below the map.

Photos can be added via station gallery page which can be accessed using “Media” dropdown located below the map.

The station details page provides also access to two modal windows which can be opened using buttons located on the left side of the station search box in the navbar:

- Information window with explanation of morphology classes, ground types and station sensor types
- Station change log showing recent changes of the station extended information

The station gallery page presents photos uploaded by the station operators. Additional photos can be added using “Upload photos” button located on the right side of the page (available only for users with write access to the network from which given station originates).

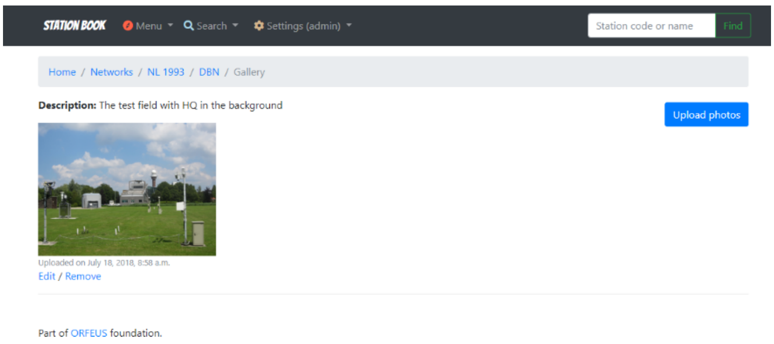


Figure A2: Station Book station gallery page

3.1.2 RRSM

The navigation bar located at the top of the screen contains a menu which allows users to quickly navigate to one of following pages:

Via “Recent events” users can filter events which occurred in various time frames. Menu entries allow users to show events from last 24 hours up to last 10 years.

Via "Search":

- Search events
- Search peak motions
- Search combined
- Custom search

By default, the home page renders events which occurred in the last month. Clicking on event Flinn-Engdahl region in the "Events" table will focus the map on given event. Clicking on event origin time in the "Events" table will show the event details page.

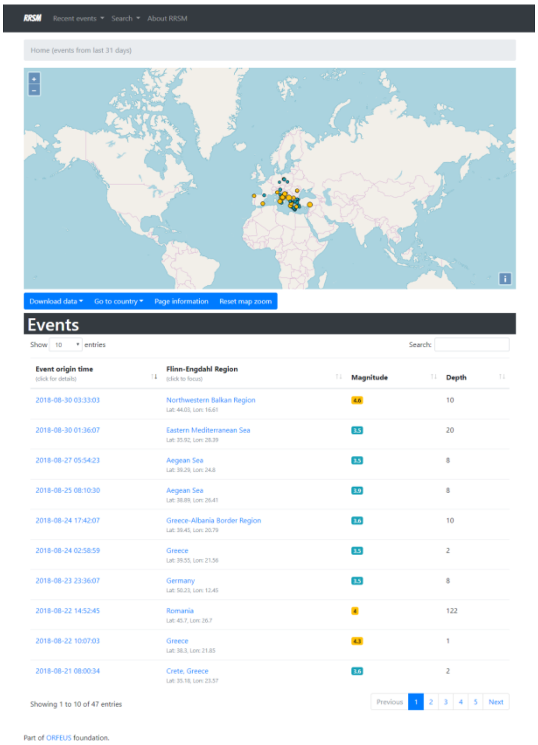


Figure A3: RRSM home page

The search events page allows users to search for events based on event characteristics.

The search peak motions page allows users to search for events based on PGA and PGV characteristics.

The search combined page allows users to search for events based on magnitude, station location and PGA/PGV characteristics.

Custom search page allows users to search for events based on custom characteristics which can be enabled and disabled using the button list located above the form.

Event details page presents the event information obtained via RRSM web service:

- Location of the event
- Locations of stations which have been triggered by the event
- PGA vs epicentral distance graph
- PGV vs epicentral distance graph
- Earthquake information
- List of stations with maximum recorded PGA and PGV values, epicentral distance and elevation

There is a button group below the map which allows user to preview RRSM web service response used to generate the page, download ShakeMap XML and Processed Waveforms, navigate to WebDC3 interface and show page information modal window.

Clicking on the network and station code in the "List of stations" table will navigate to event station stream page.

Event station stream page presents the event station streams information obtained via RRSM web service:

- Location of the event
- Location of given station
- Pseudo-Spectral Acceleration graph
- Displacement Response Spectra graph
- Station information
- List of streams

There is a button group below the map which allows user to preview RRSM web service response used to generate the page, navigate to WebDC3 interface and show page information modal window.

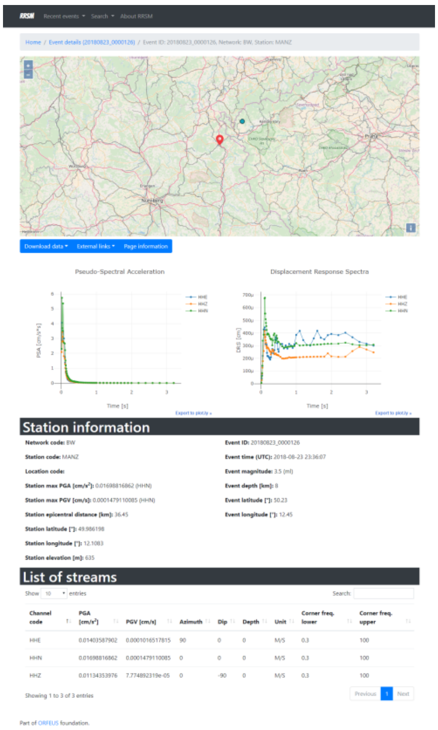


Figure A4: RRSM Event station streams page

DELIVERABLE ANNEX: VA3

VA3: Access to data and services for engineering seismology (INGV)

Work package	[WP 20 Access to data and services for engineering seismology (VA3)]
Lead	[INGV]
Authors	[L. Luzi, G. Lanzano, M. Locati, A. Rovida, R. Basili, R. Vallone, INGV]
Reviewers	[Hong Kie Thio, AECOM]

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730900.

Table of Contents

Summary

The access to data and services for engineering seismology include: i) European Strong Motion Database (ESM); ii) European Archive of Historical Earthquake Data (AHEAD); and iii) European Database of Seismogenic Faults (EDSF). This work package will strengthen and coordinate its currently separated, and intrinsically diverse, services to provide optimized access to data and tools for the seismological and the engineering seismology communities.

This deliverable describes the service provision and the access to the three services, in terms of number and type of users, number of visited pages and data download.

1 Chapter 1: Description of the services offered by the VA3

1.1 Common access to services for engineering seismology

A common portal has been created to access data and services for engineering seismology, represented by the three databases described in paragraph 1.1.1 to 1.1.3 (http://sera-va3.rm.ingv.it/, see Figure 1).

In order to assist users, a glossary with the terms common to the three services has been built (http://sera-va3.rm.ingv.it/index.php/glossaries), where terms related to engineering seismology, European institutions and data products are stored. It is planned to extend the functionalities of the existing services and to provide experimental integrated services that enable the interaction among the three data sources, as, for example, one or more study cases.

A webgis is in the making, providing a synoptic view of the data provisions with thematic map overlays, the possibility of explore geographic relationships between and among items and perform geographic selections and queries (Figure 2). The communication of the SERA-VA3 includes a twitter account (SERA-VA3, see Figure 3), where news about the portal, current projects and seismic events are routinely posted.

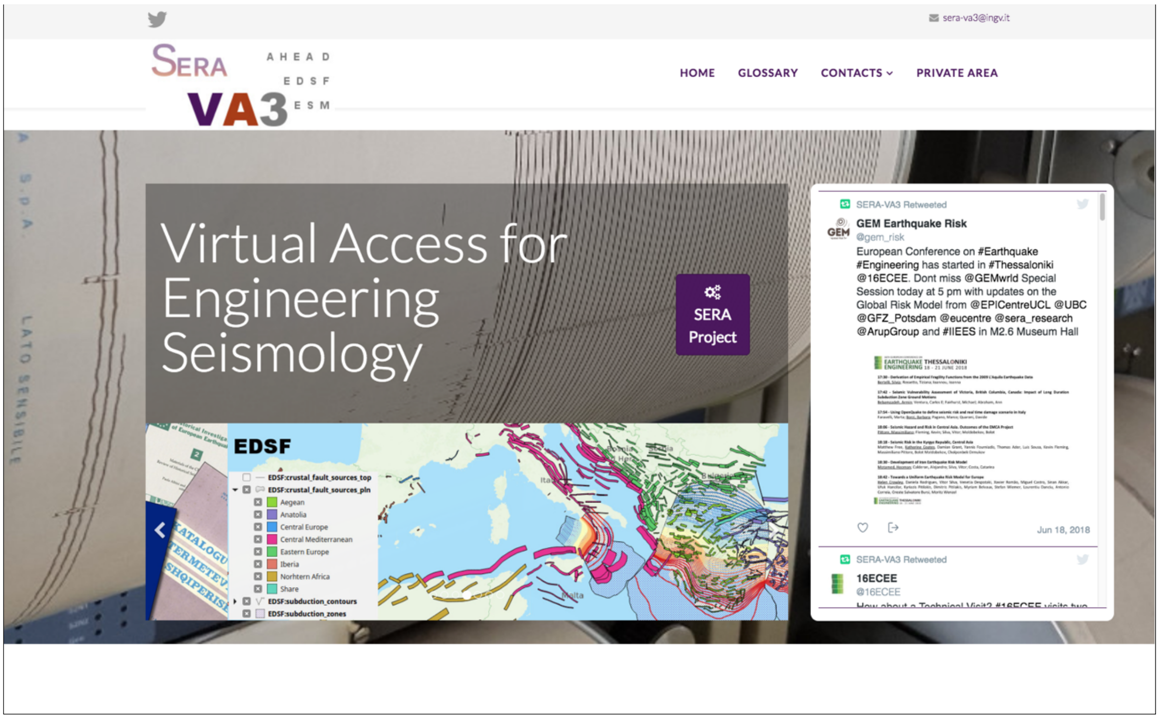


Figure 1: SERA-VA3 portal

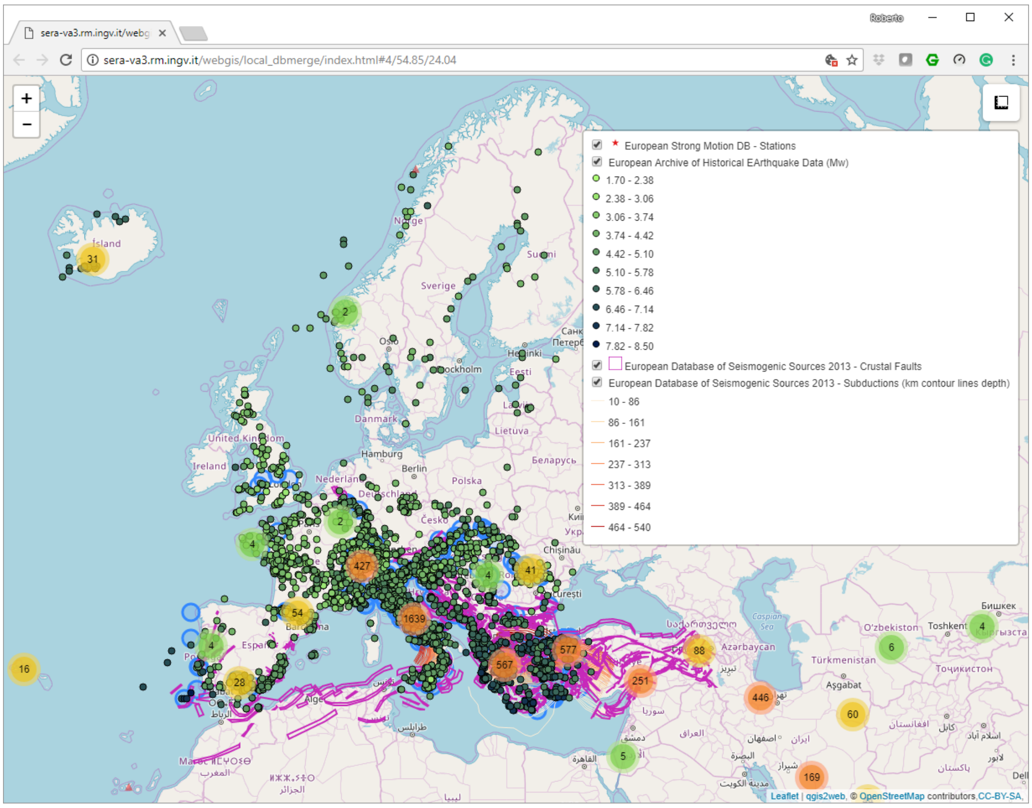


Figure 2: SERA-VA3 webgis prototype

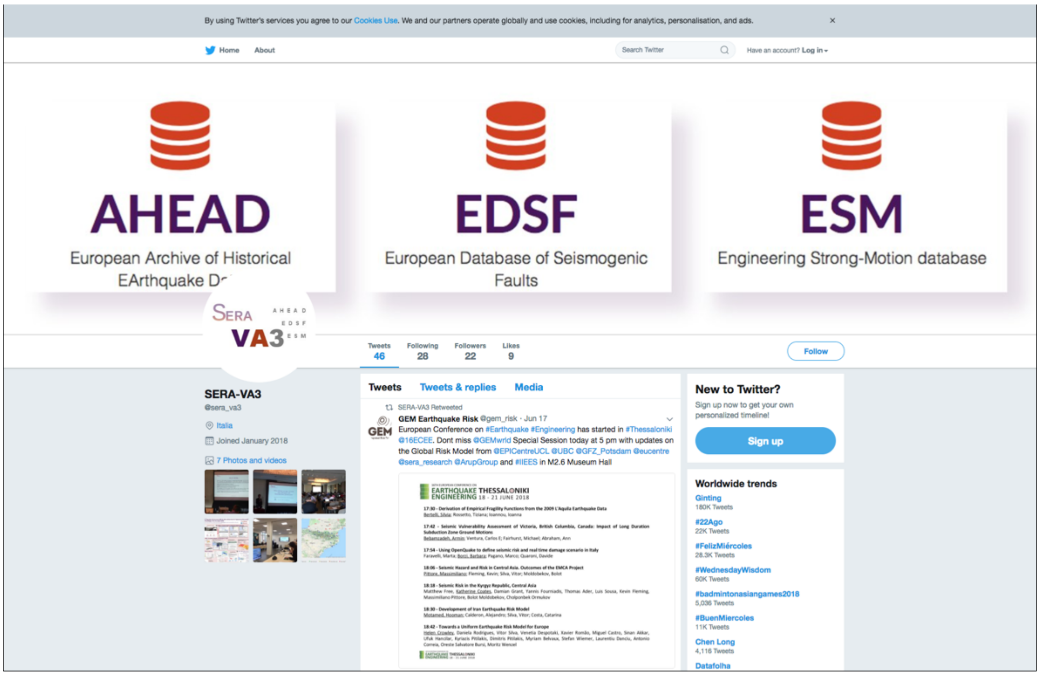


Figure 3: SERA-VA3 twitter page

1.2 European Strong Motion database (ESM)

ESM is a centralised collector of European Strong motion data, with magnitude threshold of seismic events equal to 4. It archives the waveforms recorded since 1969 by about 50 European seismic networks and provides end-users with quality-checked and manually processed waveforms. The database is updated daily with new waveforms and metadata and the number of available waveforms is about 53000 at the end of August 2018. The service is distributed and regulated under the umbrella of ORFEUS (Observatories & Research Facilities for European Seismology, www.orfeus-eu.org/) and is one of the pillars of EPOS-seismology (WP8 - waveform distribution).

The access to data is guaranteed through a web interface (http://esm.mi.ingv.it, Figure 4) and a web service that allows to download a parametric file, input for the USGS shakemap calculation (http:// http://esm.mi.ingv.it/esmws/shakemap/1/). The parametric file contains the peak values and the acceleration response spectra ordinates at three periods (0.1s, 1s and 3s), relative to one seismic event. Additional web services are in preparation to access and download waveforms and waveform parameters.

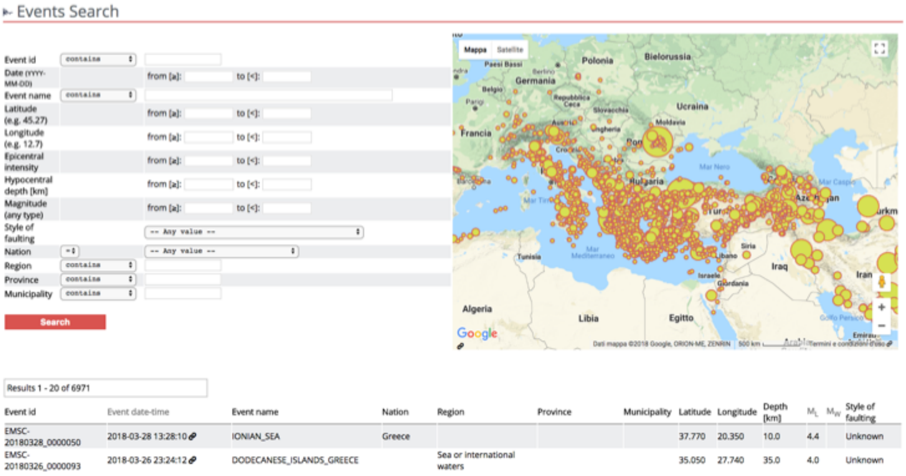


Figure 4: Screenshot of the homepage of the ESM database

Additional tools are linked to the database, such as an interface to waveform processing (http:// http://esm.mi.ingv.it/processing/, Figure 5) and a software to select a suite of 7 accelerograms compatible with the spectral shapes of the Eurocode 8 or the Italian seismic code (accessible from the database homepage). The compatibility with the European hazard map will be also checked soon after the corresponding web service will be available. A parametric flatfile is released with annual rate, for engineering seismology studies (http://esm.mi.ingv.it/flatfile-2018/). Users must register to download or process the waveforms or download the parametric flatfile.

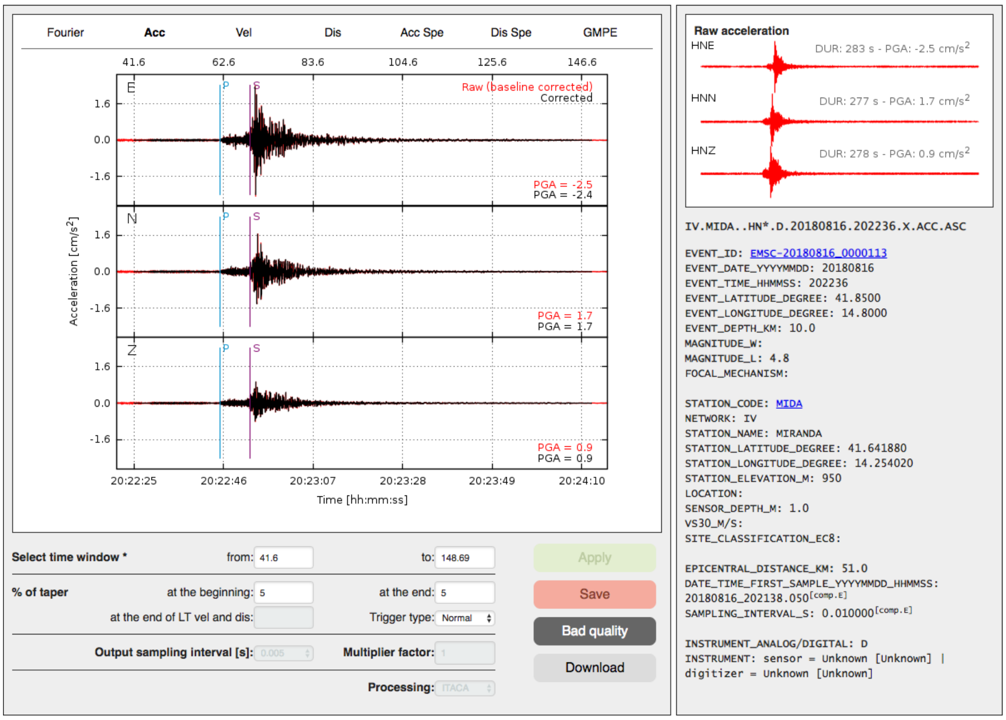


Figure 5: Screenshot of the homepage of the ESM database

1.3 European Archive of Historical Earthquake Data (AHEAD)

AHEAD, the European Archive of Historical Earthquake Data (1000-1899), is a distributed archive aiming at preserving, inventorying and making available, to investigators and other users, data sources on the earthquake history of Europe, such as papers, reports, Macroseismic Data Points (MDPs) or parametric catalogues. AHEAD consists of independent regional archives, a general repository and a collaborative inventory. At present, it mainly relies on eight regional, online macroseismic archives, which supply most of the data. It contains information on about 5000 earthquakes in the time-period 1000-1899 and provides parametric and macroseismic intensity data, derived from different sources, such as regional databases, papers, and catalogues. AHEAD establishes relationships among earthquake data of different provenance, and provides multiple macroseismic intensity datasets and interpretations for each earthquake. About 230 data sources (papers, reports, catalogues) are available to users. AHEAD is the EPOS node for distributing historical earthquake data within EPOS.

The archive is accessible at https://www.emidius.eu/AHEAD with a user-friendly web interface. The archive can be queried by earthquake, by data source or by web services.

The query by event (https://www.emidius.eu/AHEAD/query_event/; see Figure 6) allows to access the information related to each individual earthquake; the query by data source (https://www.emidius.eu/AHEAD/query_study/) allows the user to browse the list of the main studies available for the inventoried earthquakes and select them individually.

The query by web services (https://www.emidius.eu/AHEAD/services/, see Figure 7) include: i) event parameters (FDSN-event, with a user-friendly query builder); ii) OGC (Open GeoSpatial Consortium): WFS (Web Feature Service) and WMS (Web Map Service); iii) Macroseismic intensity data and iv) Bibliographical metadata.

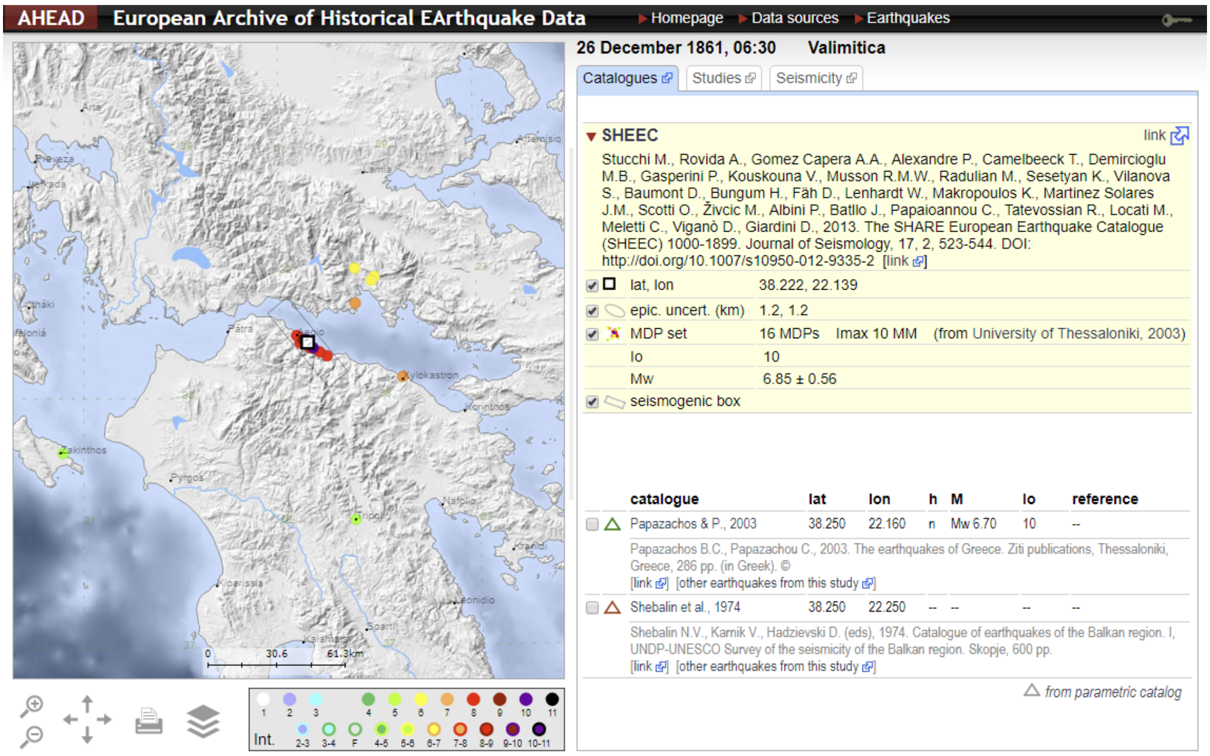


Figure 6: Example of a single event exploration.

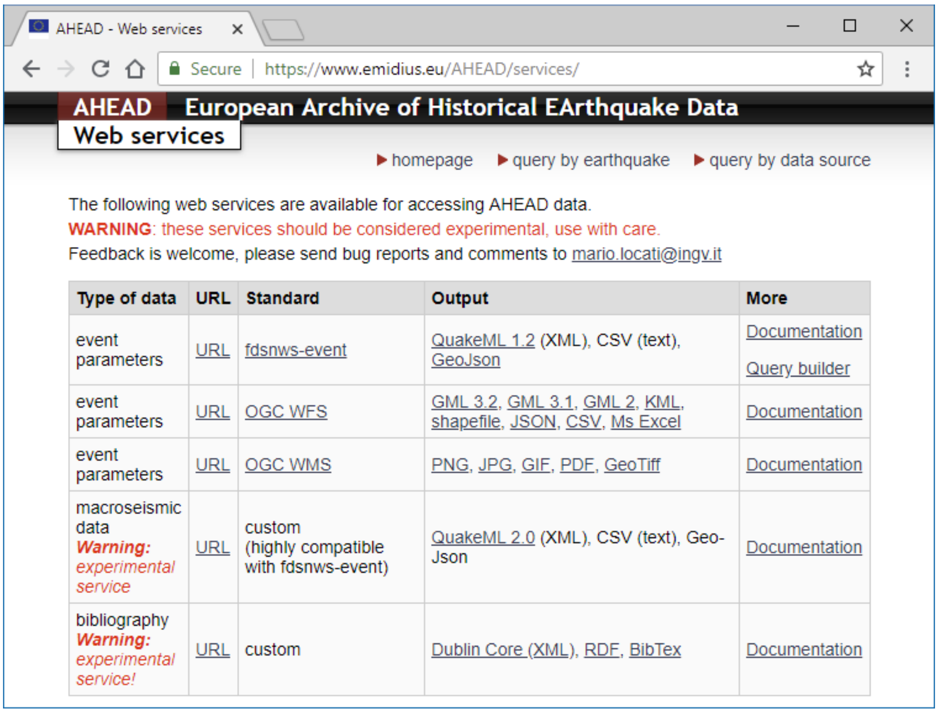


Figure 7: Web services available at AHEAD

1.4 European Database of Seismogenic Faults (EDSF)

EDSF collects and grants access to data on seismogenic faults of the Euro-Mediterranean region. It deals with 1128 crustal faults (for a total length of ~63,775 km) and 3 subduction zones (Calabrian Arc, Hellenic Arc, and Cyprus Arc; all located in the eastern Mediterranean region). EDSF provides parametric information and contextual references on geometry and behaviour of potential seismogenic faults deemed capable of generating earthquakes of magnitude ≥ 5.5. Their identification and characterization were based on papers, original data, and empirical/analytical relationships. More than one hundred scientists from several pan-European institutions contributed to the development of the database in the framework of the EU-FP7 project SHARE. The original database is being distributed through the website http://diss.rm.ingv.it/share-edsf/ since February 2013 (Figure 8), where it can be accessed through a user-friendly web interface, including a map viewer linked to parametric descriptions and references (Figure 9).

As of today, EDSF is the node for distributing seismogenic fault data within EPOS-Seismology (https://www.epos-ip.org/tcs/seismology), through the European Facilities for Earthquake Hazard & Risk (EFEHR; http://www.efehr.org/). A new web portal has been developed (http://www.seismofaults.eu/; Figure 10) providing access to the original platform, as well as to the newly implemented web services following the standard protocols of the Open GeoSpatial Consortium (OGC, http://www.opengeospatial.org/about). The following services have been implemented so far:

1. The WMS catalogue includes:
    a. GetCapability call to the collection of layers;
    b. Crustal Fault planes;
    c. Crustal Fault top;
    d. Subduction Zones;
    e. Subduction Contours;
    f. Crustal Faults (planes + top);
    g. Subduction Areas (zones + contour); and

h. EDSF whole dataset.

2. The WFS catalogue includes:
   a. GetCapability call to the collection of layers;
   b. Crustal Fault planes;
   c. Crustal Fault top;
   d. Subduction Zones;
   e. Subduction Contours.

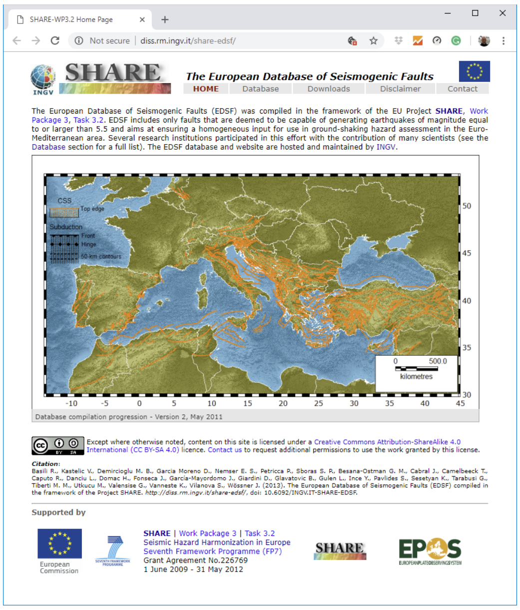


Figure 8: EDSF web site.

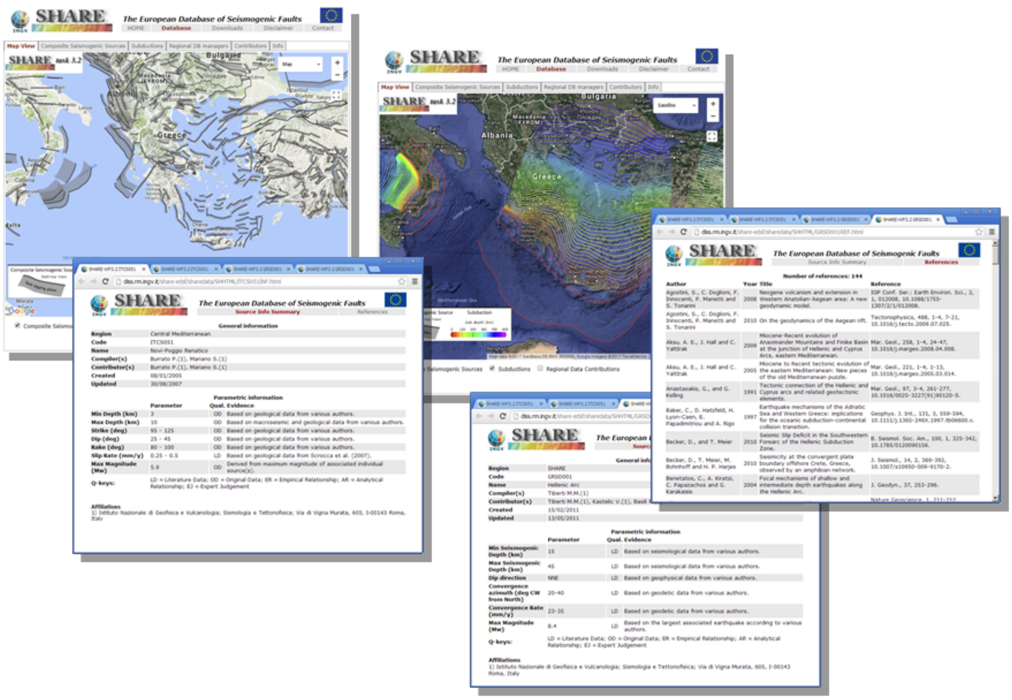


Figure 9: Navigation example of EDSF records from the original website.

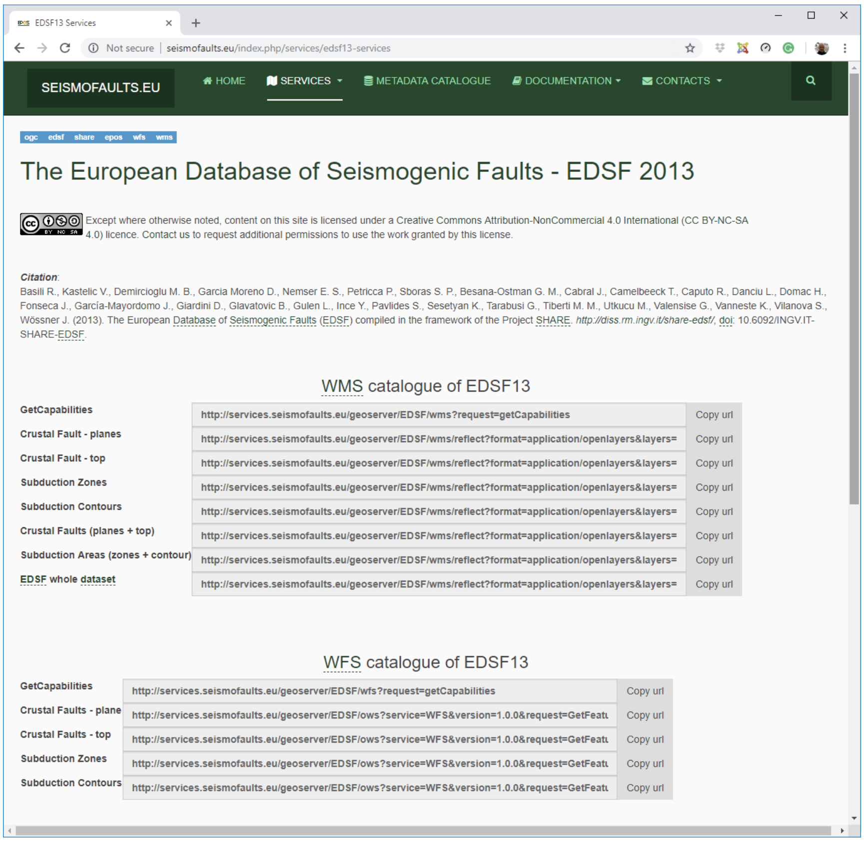


Figure 10: Web access to EDSF OGC services.

2 Chapter 2: Access statistic

The access statistics to the VA3 are generated using AWStats (https://awstats.sourceforge.io/), a free, powerful and highly customizable tool distributed under the GNU General Public License, that generates advanced web, streaming, ftp or mail server statistics.

The following sections illustrate the access distributions for the VA3 portal and the three services described in Chapter 1. Since they runs on independent web servers and distribute different types of data (e.g. waveforms, earthquakes parameters, macroseismic points, seismogenic faults), the access statistics for each service is shown separately. The reference period for this deliverable is from May 2017 (beginning of the project) and July 2018.

2.1 VA3 web portal statistics

The VA3 web portal is a very recent product, as such it is not yet known within the user community and most of its functionalities are still under development. Usage statistics (Figure 11) for this reporting phase may reflect more the navigation for development and testing rather than the behavior of the target users.

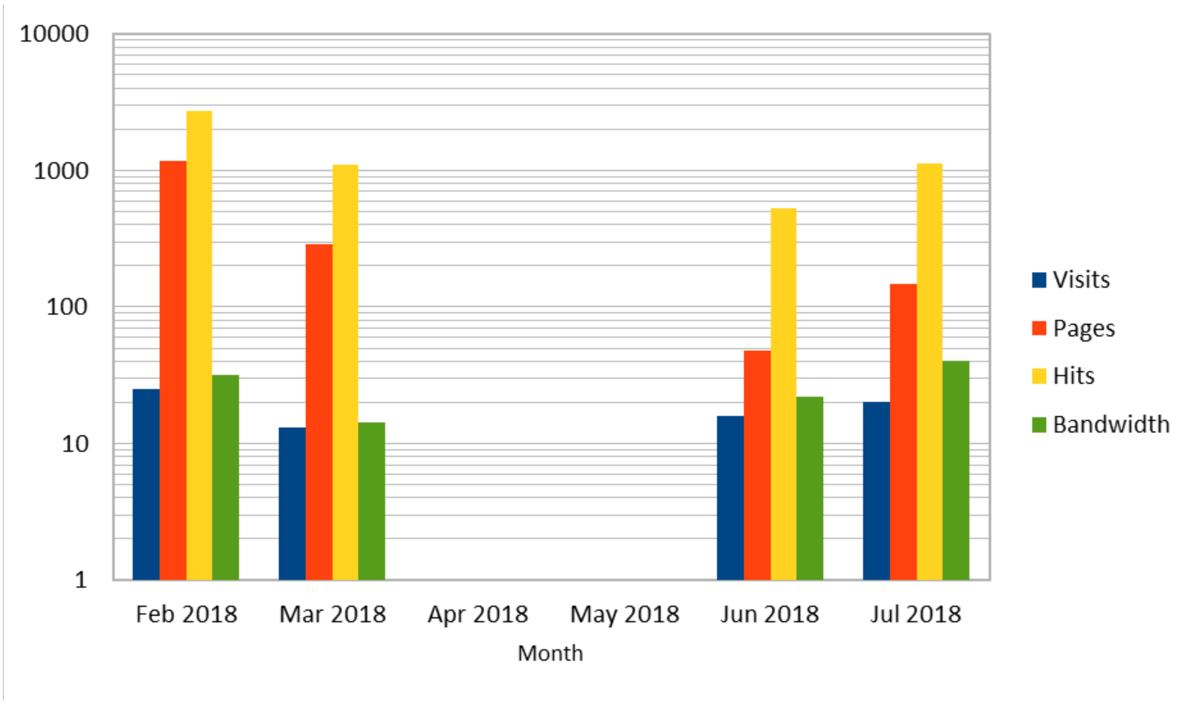


Figure 11: VA3 portal usage. Data are limited to periods of first implementation and testing.

2.2 ESM statistics

Figure 10 shows the number of unique visitors per month, which is generally larger than 1000 unit, with peaks of 2000 visitors in case of significant earthquakes. Figure 12 shows the number of repeated visits per month, which generally doubles the number of unique visits; the number of visited pages varies from 50 thousand to nearly 100 thousand per month (Figure 13), which implies 25 to 50 pages per visitor. The employed bandwidth ranges from 4 to 12 GB (Figure 14).

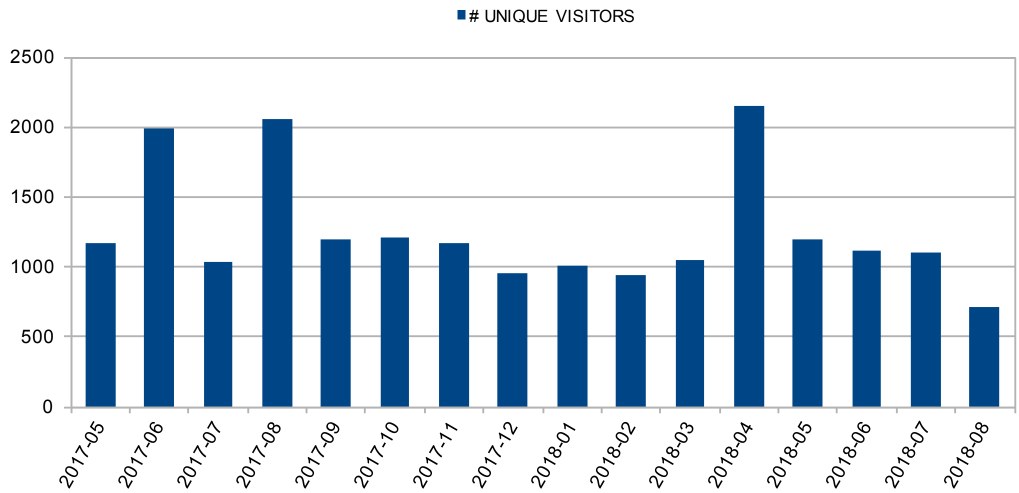


Figure 12: ESM unique visitors

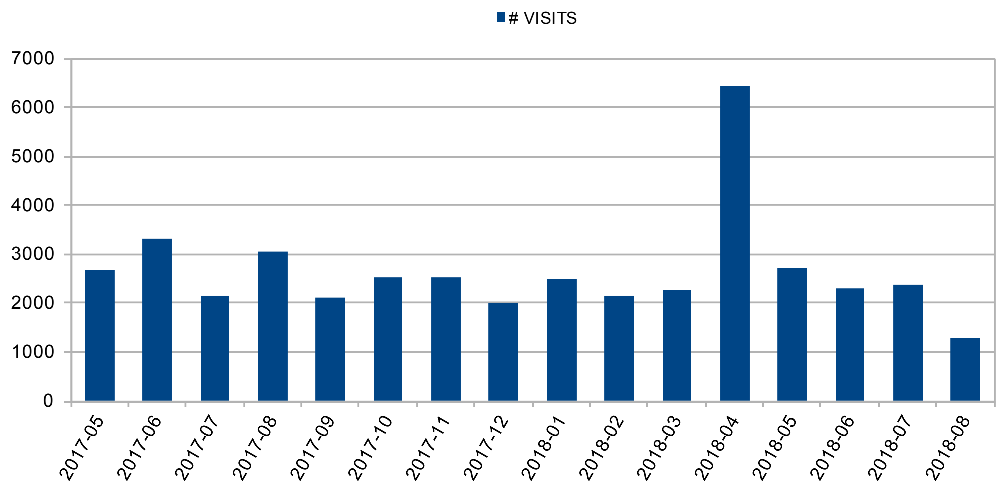


Figure 13: ESM number of visits

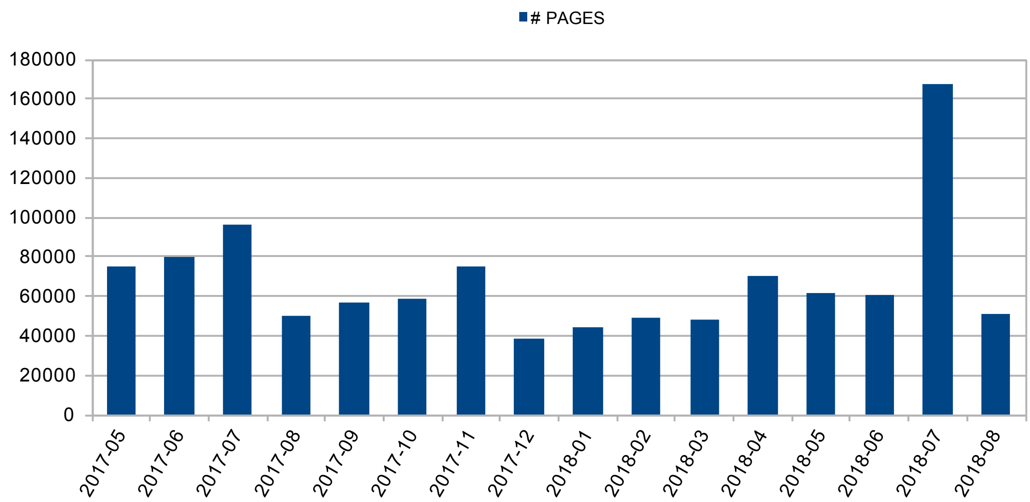


Figure 14: ESM number of visited pages

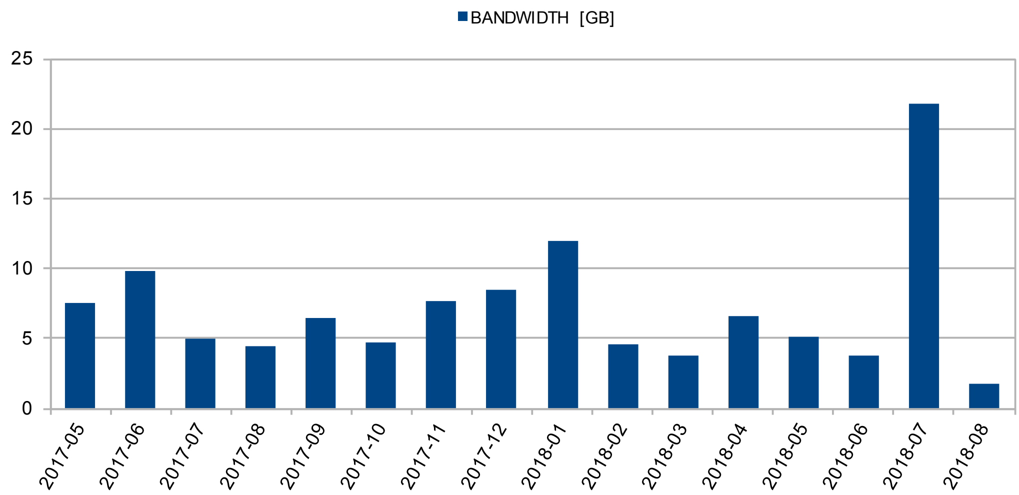


Figure 15: ESM bandwidth per month

When the ESM users are analyzed, it is evident that the ESM database is mainly used / known in Italy (Figure 15). During May-2017 July-2018 almost 50% of the pages have been visited by Italians. The rest of users are located in Europe with some visits from US and China. There is also a large number of users that could not be identified (21%). The same statistics holds for the bandwidth (Figure 17).

ESM has more than 1000 registered users. The large majority comes from the academic world of from public administrations, while a minority comes from private companies or from the consulting / freelance world.

The statistics that can be applied only to the ESM database is the count of the downloads of single records (Figure 19). The most downloaded waveform of the database (900 downloads) is the record of the Casamicciola (Ischia Island) Mw 3.9 earthquake, occurred on August 21th 2017 (station IOCA). This earthquake had a large echo in the media because of the relevant damage (and casualties) despite the low magnitude of the event. The record of the August 24th 2016 at the station Amatrice (the first shock of the 2016 central Italy sequence) has also a large number of downloads (about 750), such as other important records of the 2016 central Italy sequence. An unexpected number of downloads (500) has been monitored for the only record of the May 17th 1976, Mw 6.7 event in Northwestern Uzbekistan, at the station Gazli (Armenia), maybe because it is one of the accelerometric records of the ESM database with a PGA exceeding 1g.

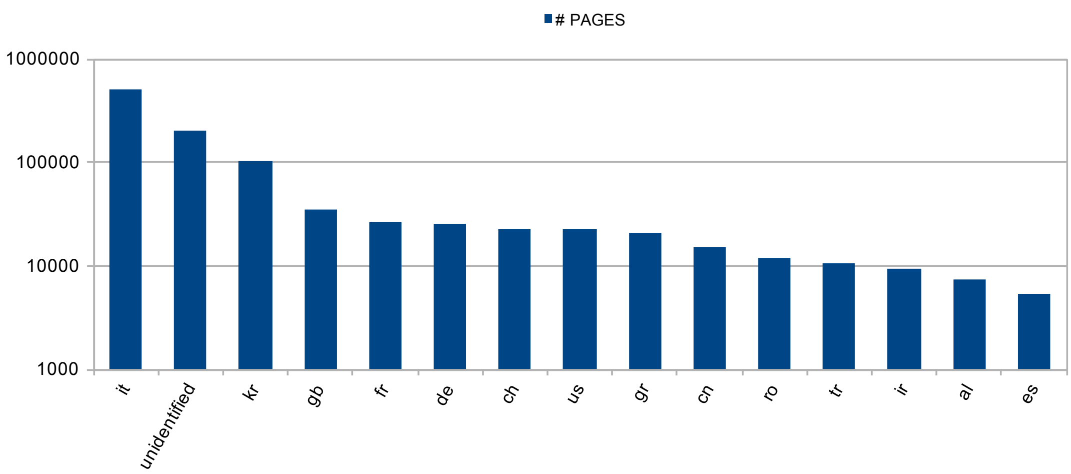


Figure 16: ESM visited pages per country

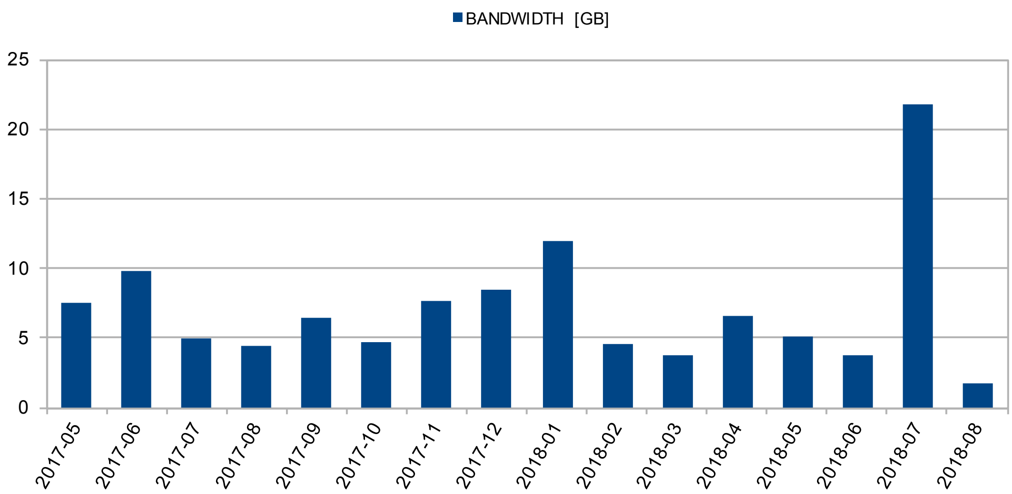


Figure 17: ESM bandwidth per country

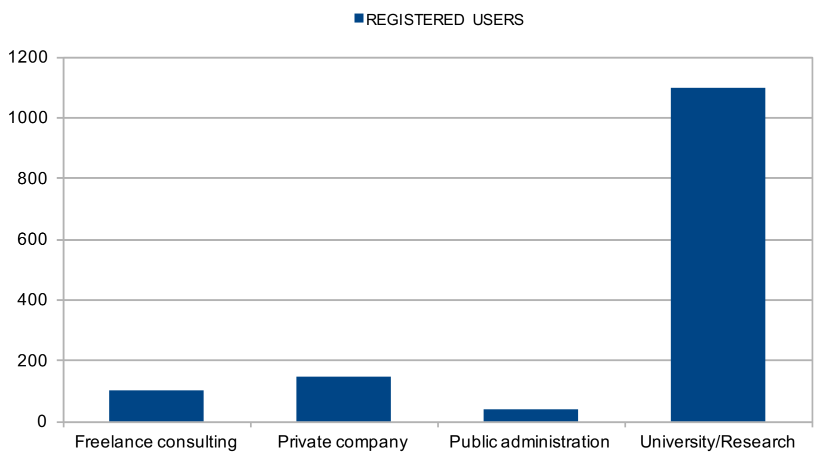


Figure 18: ESM type of users

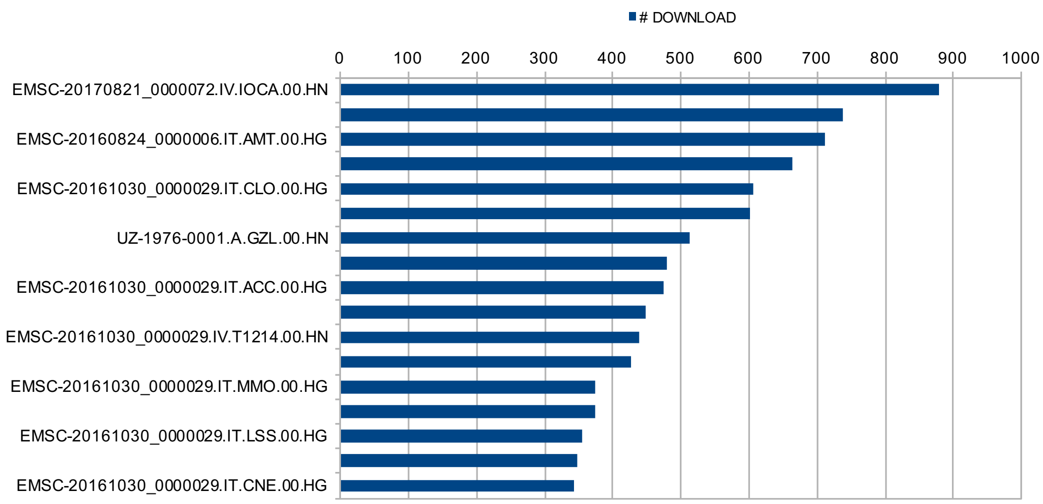


Figure 19: ESM Top 10 most downloaded waveforms

2.3 AHEAD statistics

Given the differences between web pages and web services, both in terms of number of requests and the type of users, AHEAD statistics have been split in two, to give a better inside view on each type of data access. The AHEAD web portal is well consolidated as it first opened in 2010, whereas the experimental AHEAD web services were first opened to a restricted number of early users in spring 2017, and officially launched in July 2017. The trend of visiting users (Figure 20 and 21) clearly shows that the vast majority of users access AHEAD via its web pages. However, the number of requests trend (Figure 22) shows that users are starting to massively access archived data using web services starting from spring 2018.

The comparison of data transfer efficiency between web pages and web services is demonstrated by comparing the number of requests (Figure 20) and the bandwidth usage (Figure 23). The vast majority of bandwidth is taken by transferring web pages, as they contain a series of additional information with respect to the data itself (e.g. images, graphical layout, user interface), whereas web services only carry the data itself, without any additional transfer overload.

At a first look at the web services requests (Table 1), it seems that the OGC - WMS appears the most used type of service. However, it should be stressed that the way each service works is quite different, as well their typical usage, and a plain comparison using the number of request only is misleading. In fact, the WMS only serve users with a raster image of a limited area of the entire geographical coverage they are currently viewing at a specific zoom level, and each time a new geographical area is requested (pan) or a different zoom level, the WMS send a newly generated image file to the user. The typical use of a WMS is from within a GIS such as QGIS or ESRI ArcGIS environment, and the loaded layer is commonly used as a background, reference only layer. The OGC WFS instead works by sending the entire dataset when the first requested by a user, therefore it will not generate any further requests even if the user zoom or pan the geographical area. Again, the FDSN-event web service is typically used -as far as we know- to incorporate the requested data in a more complex workflow, most probably using it as an input for a certain type of calculation.

The vast majority of requests come from servers operating in Italy (Figure 24). Among the top 10 countries, the origin of the second largest group of requests, at great distance from Italy, is Germany, then from the general “eu” domain, from Greece, Turkey, British Indian Ocean Territories ("io" domain), France, Romania, Russia, Portugal, and Switzerland.

An extensive analysis of the users collides with the recently established European General Data Protection Regulation (GDPR; EU Regulation 2016/679), as tracing and profiling users is a punishable act against their privacy. What we can say, is that the vast majority of requests to AHEAD web pages comes from the EMSC-CSEM Seismic Portal (http://www.seismicportal.eu), from an Italian weather forecast web site (http://www.meteoweb.eu), and then from Facebook.

The top 10 most requested earthquakes data and the top 10 most requested data sources are shown respectively in Table 2 and Table 3.

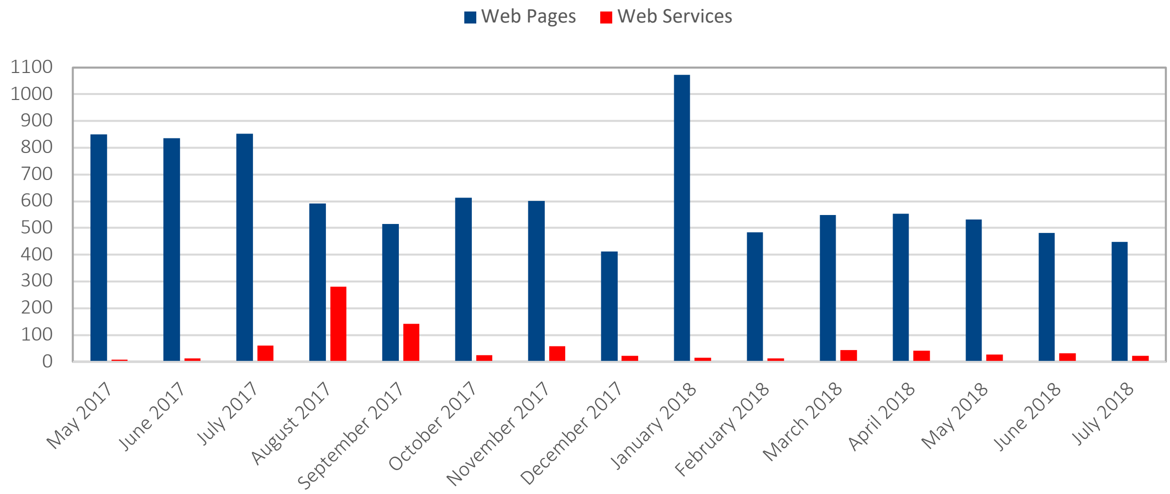


Figure 20: AHEAD unique visitors

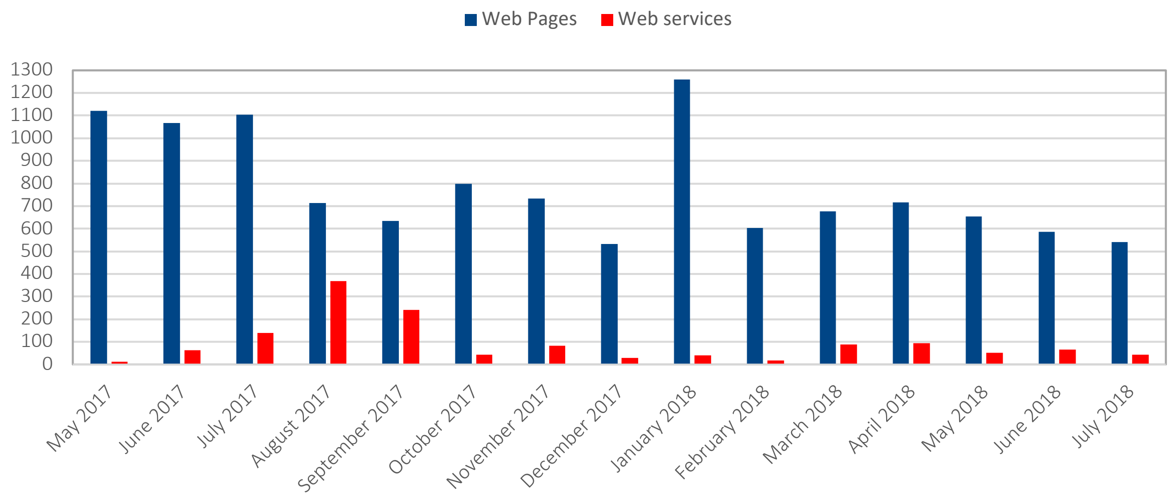


Figure 21: AHEAD number of visits.

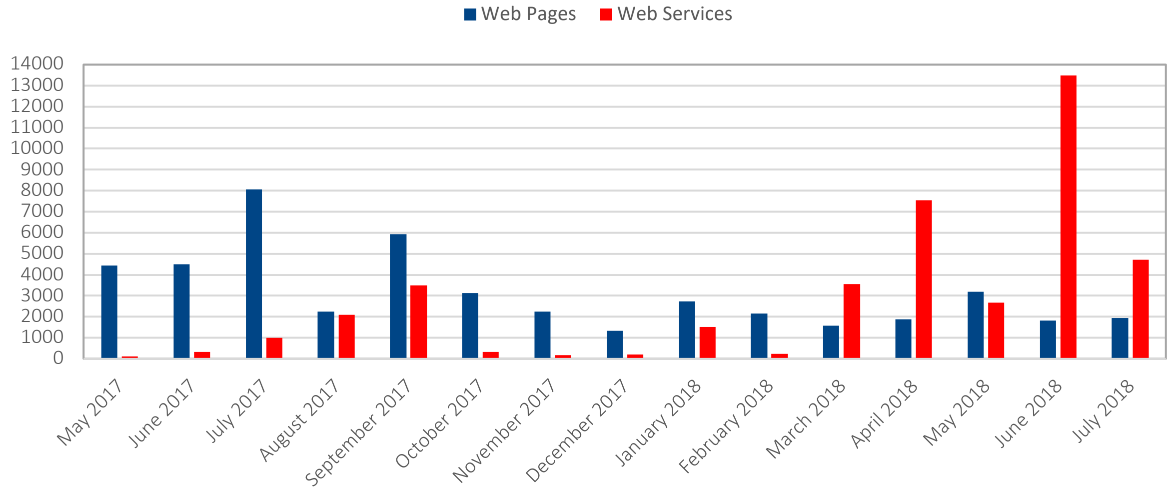


Figure 22: AHEAD number of visited pages.

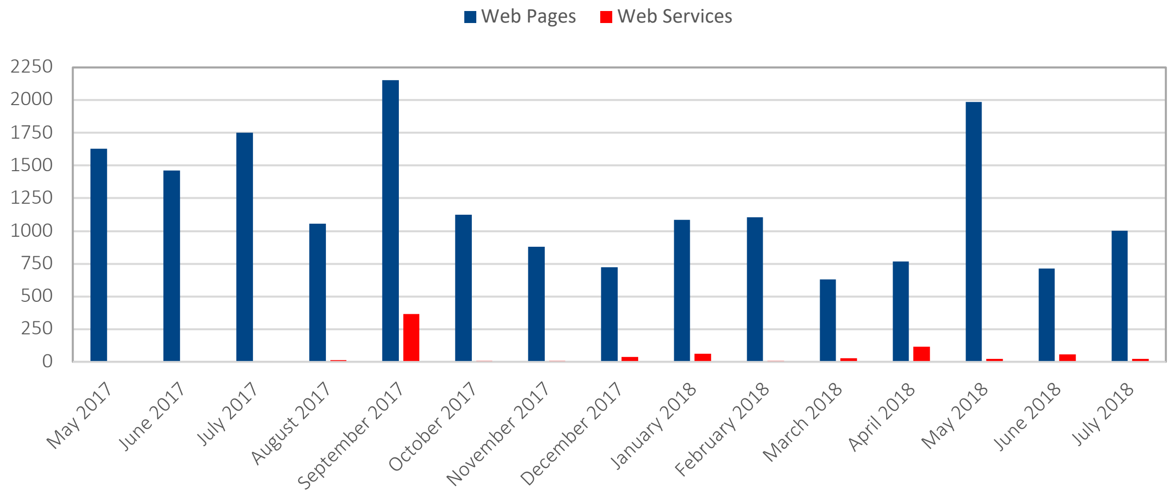


Figure 23: AHEAD bandwidth per month (in Megabytes).

TYPE OF WEB SERVICE	REQUESTS
OGC WMS	31234
FDSNWS-EVENT	1893
OGC WFS	1853
MACROSEISMIC	1040
BIBLIOGRAPHY	259

Table 1: AHEAD number of requests per type of Web Service.

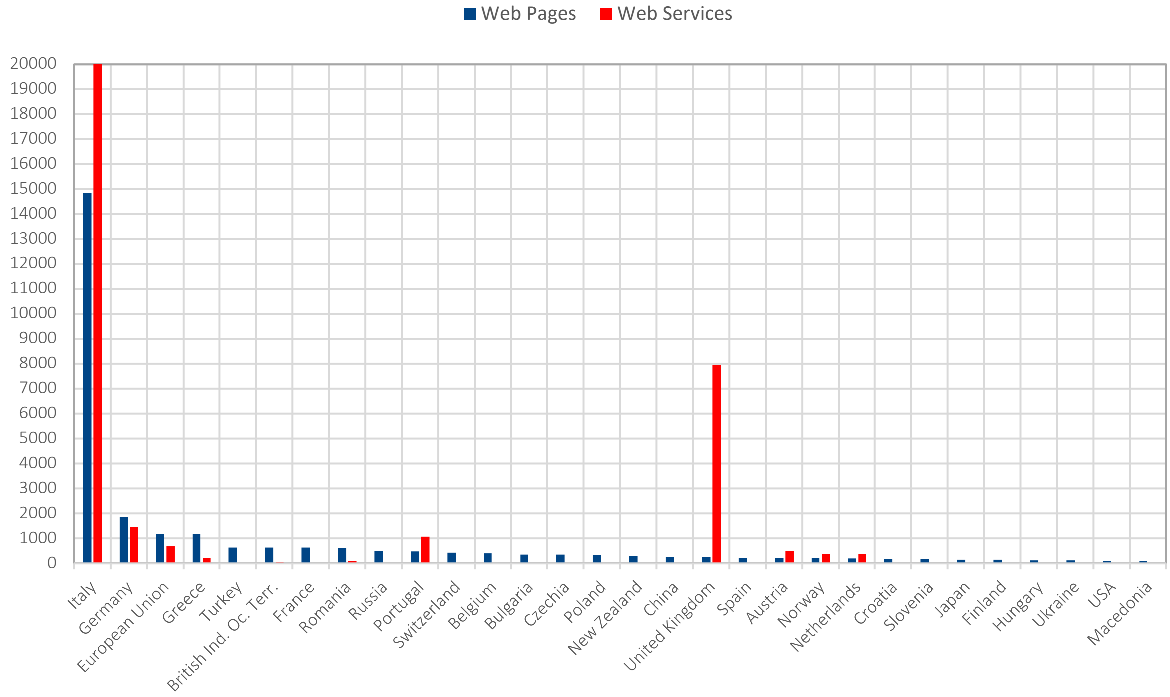


Figure 24: AHEAD visited pages per country (Top 30).

EVENTID	DATE	EPICENTRAL AREA	REQUESTS
17551101_0930_000	1755 11 01 09 30	Lisboa	1213
18571216_2115_001	1857 12 16 21 15	Basilicata	1155
17550224_0000_000	1755 02 24	Mytilene	1138
16940908_1140_000	1694 09 08 11 40	Irpinia-Basilicata	436
18550725_1150_000	1855 07 25 11 50	Törbel	425
17301023_1020_000	1730 10 23 10 20	Gubbio	108
16930111_1330_000	1693 01 11 13 30	Sicilia orientale	83
12470220_0000_000	1247 02 20	Wales	67
16880605_1530_000	1688 06 05 15 30	Sannio	62
17660522_0000_000	1766 05 22	Istanbul	58

Table 2: AHEAD Top 10 most viewed earthquakes.

DATA SOURCE	REQUESTS (WEB PAGES)	DATA SOURCE	REQUESTS (PDFS)
Ambraseys and Finkel, 1995	73	Arch. Mac. GNDT, 1995	679
Ahjos and Uski, 1992	69	St. Geof. Amb., 2002	384
Albini and Moroni, 2003	53	Castelli et al., 1996	288
Ambraseys and Finkel, 1999	44	Shebalin et al., 1974	260
Albini and Rodriguez, 2001	44	Albini et al., 2003	183
Albini and Pantosti, 2004	41	ENEL, 1995	150
Ambraseys and Jackson, 1990	40	Olivera et al., 2006	149
Ambraseys and Finkel, 1991	40	CPTI04	117
Albini and Rovida, 2010	40	Camassi, 2001b	111
Soysal et al., 1981	40	Barbano et al., 1996	107

Table 3: AHEAD Top 10 most viewed data sources (literature).

2.4 EDSF statistics

EDSF usage statistics are subdivided into two categories: 1) web pages, and 2) web services. Web pages regard only the visits to the portal http://www.seismofaults.eu/ which introduces to the various forms of distributing seismogenic fault information. Web services regard the use of the data distributed via OGC standard protocols. The use of the data though the original EDSF website, including visits to the map viewer, connected pages, and file downloads, are not counted here.

The server that distributes these data is hosted in the INGV data center in Rome, which follows security and EU General Data Protection Regulation (GDPR; https://eugdpr.org) rules that limit the monitoring of some users' information. Since the client IP is hidden, we cannot count the number of unique visitors and the geographic distribution of visits. To overcome these and other limitations, a different hosting solution is under examination.

About web pages, the monitoring with AWStats started in January 2018. Before that date we used Google Analytics tools (https://analytics.google.com) which provides access statistics in a different way. For compatibility of information with the ESM and AHEAD statistics, those data ore not reported here. The only statistics that can be easily compared between the two tools is the number of visited pages; Figure 25 shows that although the numbers are in the same order of magnitude, there are remarkable differences on how page visits are counted. Note also that Google Analytics cannot monitor the OGC web services.

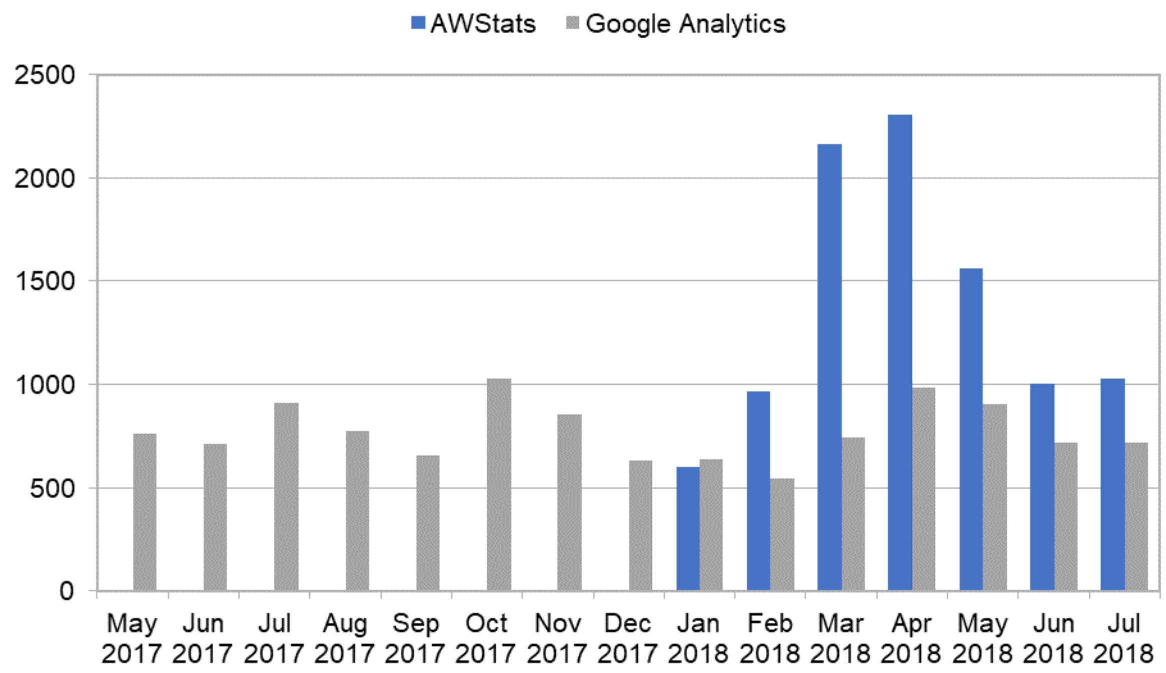


Figure 25: Comparison between AWStats and Google Analytics in monitoring the number of visited pages.

Figure 26 and Figure 27 show the number of visits and of visited pages, respectively. The distribution of visits seems to be rather uniform during the observation period for both web pages and web services, with a factor of two in favor of web pages. The distribution of visited pages instead if very different. Usage of web services shows a striking increment since April 2018. This increment is related to the EPOS validation activity that started in that period. To verify the reliability of published web services, EPOS adopted a monitoring system (NAGIOS; https://www.nagios.com) which attempts a connection every five minutes to four different EDSF layers. This activity alone accounts for about 35,000 visited pages per month. Excluding this period for web services, the average number of pages per visit is seven for web pages and 63 for web services.

Regarding the bandwidth usage (Figure 28), we notice little difference between web pages and web services despite the differences in the amount of visited pages. This circumstance can be explained by the characteristic lightweight of EDSF data files. The increment of bandwidth used by web service since April 2018 reflects the increment of visited pages for the EPOS validation.

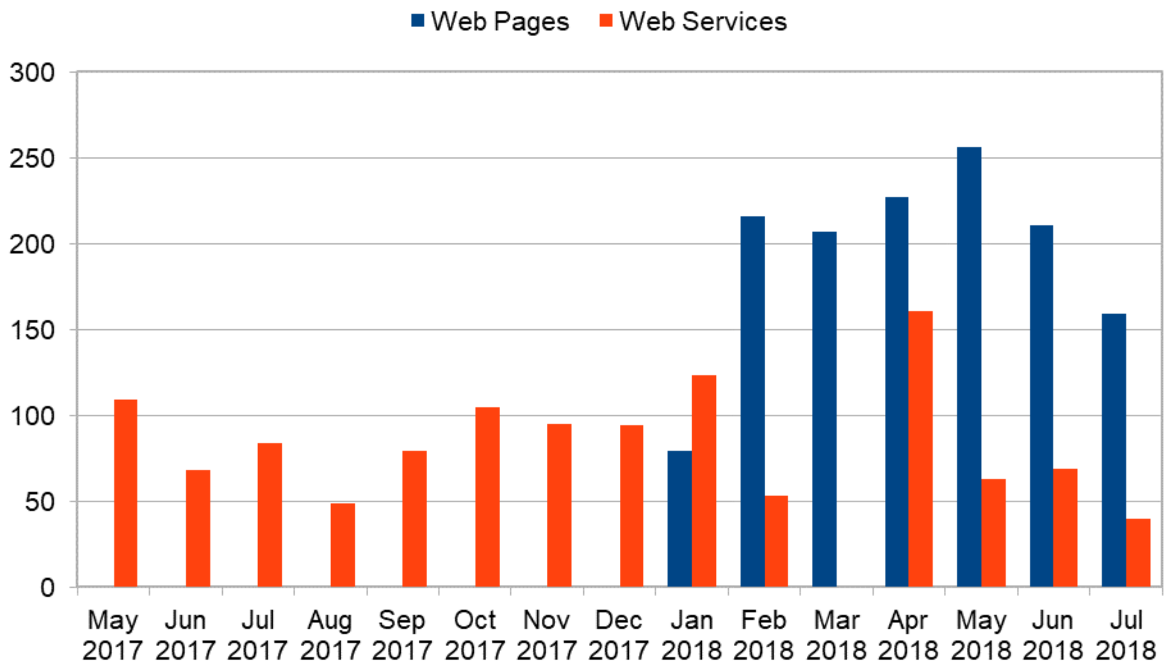


Figure 26: EDSF number of visits.

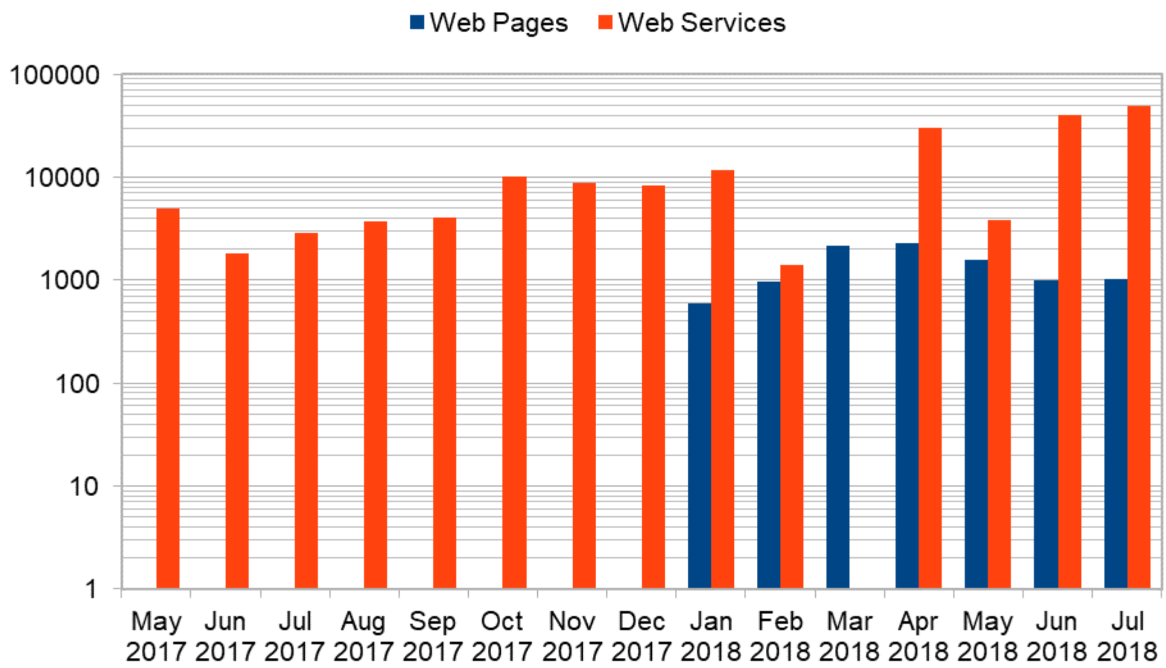


Figure 27: EDSF number of visited pages.

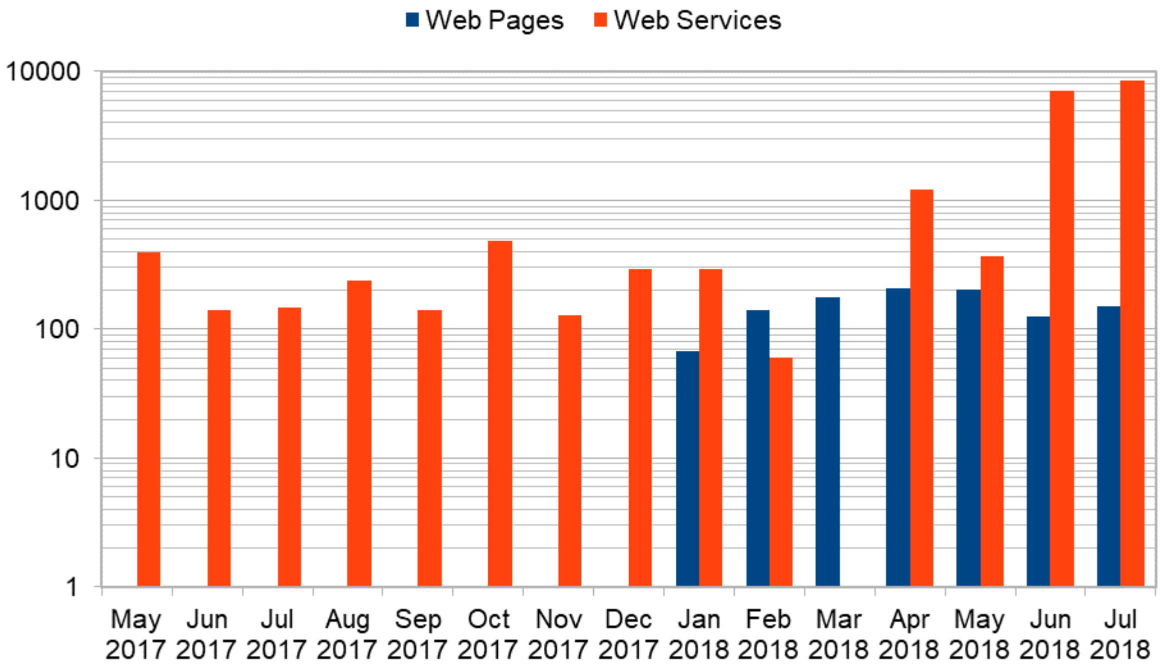


Figure 28: bandwidth per month (in Megabytes).

The breakdown between crustal faults and subduction zones shows that crustal fault pages are viewed three times more frequently than subduction zone pages as a WMS, whereas they are requested equally frequently as a WFS. Table 4 shows this breakdown. Note that the number of visited pages in this case is smaller than the total because there are generic requests (e.g. GetCapabilities) that do not distinguish between the individual layers.

The use of web services can also be expressed by requests. Differently from pageviews, requests are queries of any kind (e.g. GetMap, GetLegend, GetFeature, GetFeatureInfo). Table 4 reports the request breakdown by category: WMS, WFS, and OWS. The latter includes all requests made through a query that does not distinguish between the first two categories. This analysis shows that WMS requests are five times as many as WFS requests. The OWS requests represent instead just a little fraction of the total.

Table 4: EDSF visited pages breakdown by type and web service.

TYPE OF WEB SERVICE	CRUSTAL FAULT	SUBDUCTION
OGC WMS	61958	8313
OGC WFS	22184	8408

Table 5: EDSF number of requests per type of Web Service.

TYPE OF WEB SERVICE	REQUESTS
OGC WMS	98929
OGC WFS	19017
OGC OWS LOGS	786

Seismology and Earthquake Engineering Research Infrastructure Alliance for Europe

DELIVERABLE ANNEX: VA4

VA4: European Facilities for Earthquake Hazard and Risk (ETH)

Work package	WP21 – Virtual Access VA4: European Facilities for Earthquake Hazard and Risk
Lead	ETH Zurich, Switzerland
Authors	L. Danciu, P. Kästli, Swiss Seismological Service, ETH Zurich, Switzerland
Reviewers	J. Douglas, University of Strathclyde, UK

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730900.

Table of Contents

Summary

This document provides a brief description of the SERA Virtual Access (VA): European Facilities of Earthquake Hazard and Risk (EFEHR – www.efehr.org). The document provides an overview of the main services and gives insights on the usability of these services. The usability is quantified by statistics on access since the beginning of SERA project in May 2017. The web-traffic analytics of the EFEHR web-portal indicates a preference for users to access the hazard maps and uniform hazard spectra. The visitors are distributed worldwide, and often the visitors are consulting the hazard values at a specific site, rather than downloading entire sets of results and/or models. An average of 5000 requests per month are observed for all the web-services. The most used web application is the hazard map viewer. In terms of outcomes resulting from this access, there are about 500 citations of the datasets, results, models provided by EFEHR web-portal since 2013.

1 Overview of the SERA-VA4: European Facilities for Earthquake Hazard and Risk – EFEHR

European Facilities for Earthquake Hazard and Risk (EFEHR) provides open access to seismic hazard and risk models. The EFEHR web-platform is the public interface of a complex system connecting databases of relevant datasets, inputs, outputs and model results with a display portal for access, visualization and download. The portal consists of web-services linking to the main hazard outputs: seismic hazard maps, seismic hazard curves and uniform hazard spectra. Such webservices are based on various technologies summarized in the next sections and are fully operational since 2013. The risk services are under development, prioritized for early 2019, are also listed in the next section.

In terms of available hazard models, the EFEHR web-portal distributes the seismic hazard models for:

- The 2013 European Seismic Hazard Model (ESHM13, Woessner et al 2015)
- The 2014 Earthquake Model of the Middle East (EMME14, Giardini 2018)
- The 2015 Swiss Hazard Model (SuiHaz15, Wiemer et al 2015)
- The 1999 Global Hazard Map of the Global Seismic Hazard Assessment Program (GSHAP, Giardini 1999)

EFEHR web-portal provides a single access point for data, models and results. No user authorization is required. Technical description of the EFEHR web portal is summarized in the Appendix. In the next sections the relevant web services are described, as they are the key interface to access the data of the EFEHR web platform.

1.1.1 Data access applications

The EFEHR web platform mounts three stand-alone web applications for interactively discovering and retrieving hazard curves (Figure 1, upper), and hazard spectra (Figure 1, lower) and hazard maps (Figure 2).

Each of the applications implements a workflow to select a point of interest from a map and retrieve all hazard models covering that area. Given the site selection, specific parameters settings (e.g. model selection, intensity measure type, probability of exceedance, site class, hazard aggregation type) are used to query the request from the database. This allows efficient selection of the data points of interest from over 0.6 billion hazard data points. For hazard spectra there is an option to compare with the Eurocode 8 elastic design spectra for different classes of earthquakes. For hazard maps, additional elements that can be added to the map including the earthquake catalogue, seismogenic sources and/or active faults. The layer manager is fully customizable allowing the addition of the input datasets and/or control different transparency levels. A log panel is located at the bottom of the web page. It provides summary of the model query parameters and a URL link to download the selected data.

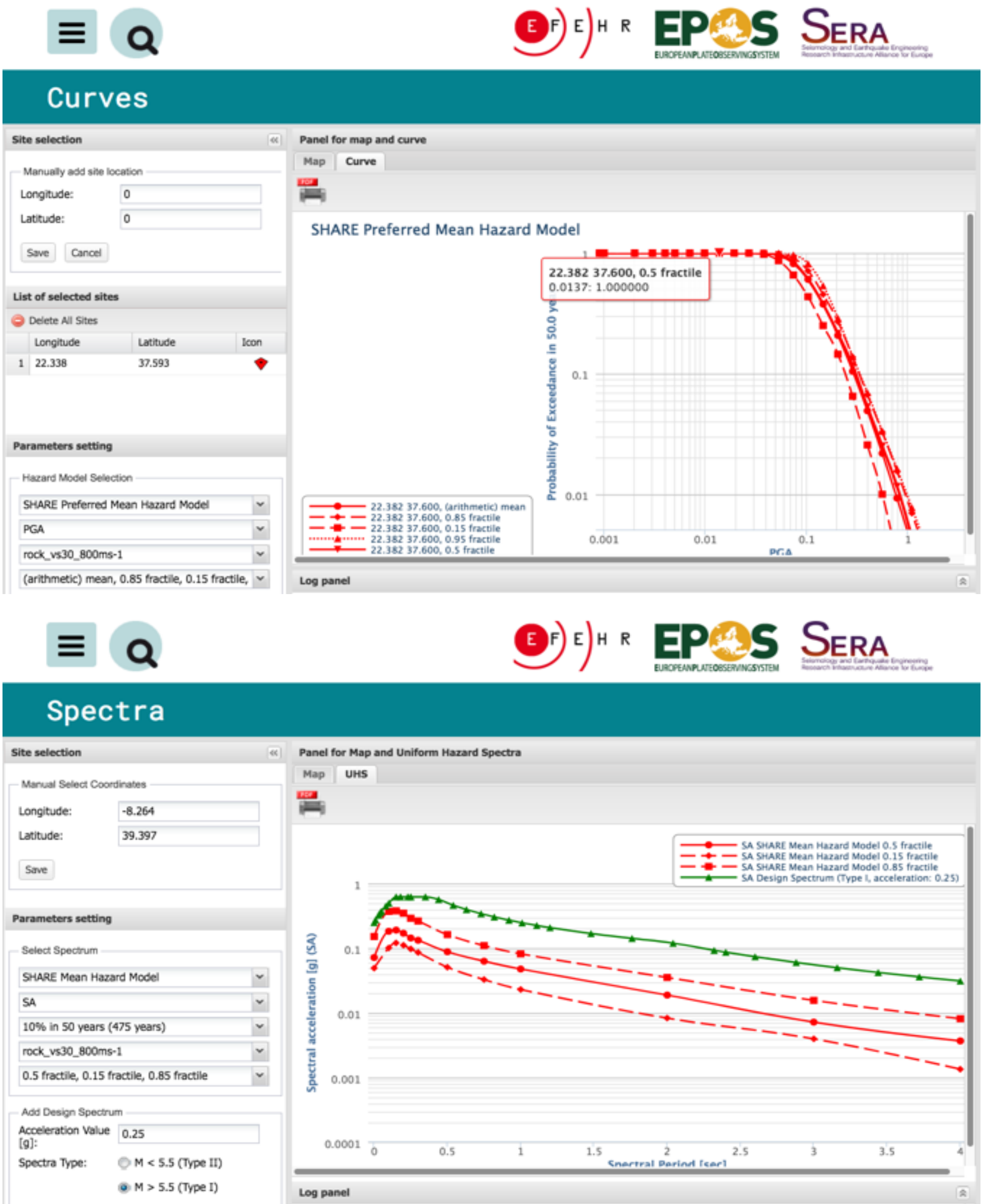


Figure 1: EFEHR web-portal: hazard curve viewer (above) and hazard spectra viewer (below)

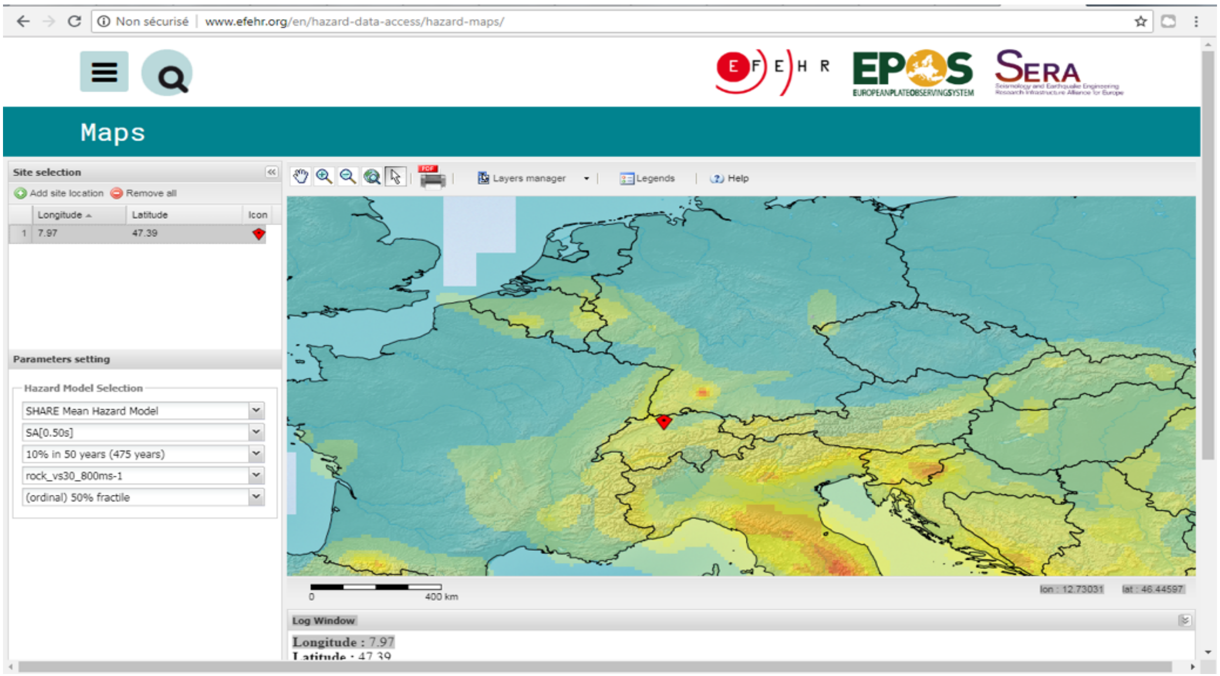


Figure 2: EFEHR web-portal: hazard map viewer application

1.1.2 Hazard related web services

Viewers retrieve hazard data (as well as metadata on available models and parameters) using a restful web service API. The API is public and can be used directly by researchers to programmatically retrieve data. A WADL (Web Application Description Language) definition allows automatic generation of web service clients for several modern programming languages. An example of such a client in MATLAB is distributed by EFEHR (http://www.efehr.org/en/Documentation/web-services/data-access-via-matlab/). Another example is a Python script developed at Royal Observatory of Belgium to programmatically access the EFEHR web services: https://github.com/ROB-Seismology/rshalib/blob/master/result/efehr.py .

For hazard maps, parameter discovery is implemented using EFEHR's REST API; maps themselves are shipped via custom services (ASCII data), file download (compressed ESRI(R) shapefiles), and OGC-standardized Web Map Services (projected map images). The full documentation of the REST (Representational State Transfer) API (Application programming interface) is available at http://www.efehr.org/en/Documentation/web-services/.

EFEHR - hazard maps to provide access to seismic hazard maps of various seismic hazard models (regional - ESHM13, EMME14, GSHAP and national-SuiHaz15). Selection by point of interest or region of interest (user-defined polygon)
Web service example: http://appsrvr.share-eu.org:8080/share/map?lat=47.0&lon=6.0&filter_map
Documentation : http://www.efehr.org/en/Documentation/web-services/hazard-map-data/

EFEHR - hazard curves to provide access to seismic hazard curves of various seismic hazard models (the 2013 European Seismic Hazard Model, the 2014 Earthquake Hazard Model of the Middle East). Selection by point of interest (user defined)

Web service example: http://appsrvr.share-eu.org:8080/share/curve?lat=47.5&lon=7.6

Documentation : http://www.efehr.org/en/Documentation/web-services/hazard-curve-data/

EFEHR - hazard spectra services facilitate access to uniform hazard spectra of various hazard models (the 2013 European Seismic Hazard Model, the 2014 Earthquake Hazard Model of the Middle East, GSHAP) for a given location, for various probabilities of exceedance (or mean return periods)

Web service example: http://appsrvr.share-eu.org:8080/share/spectra?lon=7.6&lat=42.5

Documentation : http://www.efehr.org/en/Documentation/web-services/uniform-hazard-spectra/

EFEHR - EC8 Elastic Design Spectra: allows comparison of the hazard spectrum (in acceleration or velocity) predicted by a hazard model to the Eurocode 8 design spectra, we provide a service to retrieve the design spectra (spectral acceleration or spectral velocity) for a given peak ground acceleration.

Web service example:

http://appsrvr.share-eu.org:8080/share/design-spectra/design_spectra_elastic_ec?lon=7.6&lat=42.5&spectratype=bigmag&imt=SV&design-acceleration=0.02

Documentation : http://www.efehr.org/en/Documentation/web-services/ec8-elastic-design-spectra/

1.1.3 Risk related web-services

As part of the preparation towards SERA milestone MS22 (Delivery of first European risk model for inclusion in SERA VA distribution, month 30), EUCENTRE and the Global Earthquake Model Foundation (GEM) have been working on the following risk services:

EFEHR – access to exposure data to provide access to the European exposure model that will be used in the European risk model being developed in SERA JRA4. This model will be stored in GEM's OpenQuake platform (https://platform.openquake.org), and web services (more specific, OGC Web Map Services (WMS), Web Feature Services (WFS) will be developed to allow users to access the data for a given region of interest.
Web-services: under development;
Estimated first version: Early 2019

EFEHR - access to building stock vulnerability to provide fragility/vulnerability models from a European database of fragility/vulnerability models. We are currently investigating the use of GeoNetwork (https://geonetwork-opensource.org/) and Open Geospatial Consortium (OGC) Catalog Service for the Web (CSW) and/or OAI-PMH (Open Archives Initiative) to expose this vulnerability database. In this case, the shipping format is a challenge as the information shipped is not strictly spatial and can include documentation.
Web-services: similar with OpenQuake platform: https://platform.openquake.org/vulnerability/list.
Estimated first version: October 2018

EFEHR - risk maps to provide access to seismic risk maps from the European Seismic Risk Model (JRA4). Users will be able to access, through WMS and WFS web services, the risk maps available for their region of interest. The risk maps (for the same risk metrics) will provided for up to 10 different return periods, as alongside with the average annual loss (AAL).
Web-services: under development;
Estimated first version: October 2018

EFEHR - risk curves to provide access to seismic hazard curves at a given point of interest. It is expected that there will be in the order of 2.5 million risk curves (which will provide the annual frequency of exceedance of economic loss and fatalities.
Web-services: under development;
Estimated first version: October 2018

1.1.4 Static access to data

In addition to interactive access to model results, within the EFEHR web-portal, we provide direct links to key elements and datasets used in the model building process, and input files for calculation. The main access is provided via direct download of the following elements:

- SHEEC - European Earthquake Catalogue (SHEEC v3.3, Grünthal et al 2013, Stucchi et al 2012); Data URL: http://www.efehr.org/export/sites/efehr/.galleries/dwl_europe2013/SHAREv3.3.zip

- European Database of Seismogenic Faults (EDSF, FSBG Model Version 6.1, Basili et al 2012); Data URL: http://www.efehr.org/export/sites/efehr/.galleries/dwl_europe2013/FSBGmodelv6.1.zip

- Seismogenic Source Models (ESHM13 Source Model Version 6.1, Woessner et al 2015) ; Data URL : http://www.efehr.org/en/Documentation/specific-hazard-models/europe/seismogenic-sources/

- Strong Motion Dataset (SHARE GMPEs Dataset, Yenier et al 2010, Delavaud et al 2012); Data URL: http://www.efehr.org/export/sites/efehr/.galleries/dwl_europe2013/SHARE_StrongMotionData.zip

- ESHM13 Input Files for OpenQuake (ESHM13 OQ Input Files, Woessner et al 2015); Data URL:http://www.efehr.org/export/sites/efehr/.galleries/dwl_europe2013/SHARE_OQ_input_20140807.zip

- EMM14 Input Files for Middle East (EMME14 input files for OpenQuake, Danciu et al 2016); Data URL:http://www.efehr.org/export/sites/efehr/.galleries/dwl_Middle-East-2014/the2014_emme_model_oq_input_files_ver07.zip

All the above datasets and input files are provided for full transparency of model building and replication of the hazard calculation and reproducible results. In the next update of the EFEHR platform these static datasets will be replaced by corresponding web services, currently under development within the EPOS/SERA project. This update will contain the risk services as well and it is foreseen to be finalized in the end of the SERA project.

2 EFEHR Web-portal Usage and Metrics

2.1.1 General user characteristics

Visitors of the EFEHR web platform come from all over the world. Most visitors of the EFEHR web platform already know about it beforehand:

- 88% the page either from bookmark or from directly typing its web address.
- 7% come from references (www.share-eu.org, scientific publications, Wikipedia etc.)

- 5% come from search engines

A typical visitor (median) accesses 7 pages (either web pages or viewer applications). 10% of the visitors come back to the web portal within the same month. Also, most of the visitors access either the hazard at a specific site or just view the maps, rather than downloading the data.

The web traffic of EFEHR is summarized in Figure 3 in terms of usage information (the number of page loads). Note that the multiple consecutive page views within the same portal during the same day is quantified as one visit.

We started monitoring the exact provenance of the requests only in May 2018. Within the observational time window of June to August 2018 the top five countries were from Germany, Russian Federation, Republic of Serbia, India and China. The spatial distribution of the page load per country within Europe is given in Figure 4. The top five countries in terms of accessing content on EFEHR web-platform are Germany, Russian Federation, Republic of Serbia, Italy and France.

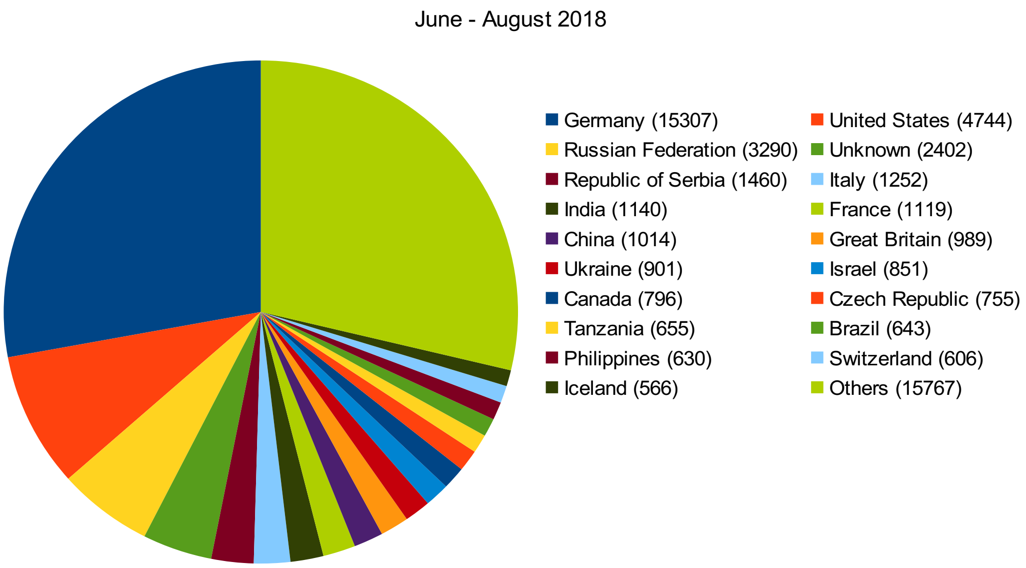


Figure 1: Pie Chart of number of accessing the EFEHR web-portal per country between June to August 2018.

The observation period is probably too short to give an unbiased view on the spatial distribution of the public interest. High access rates from countries (e.g. Tanzania and Philippines) with limited model coverage in EFEHR web-platform (GSHAP only). We do not have an explanation of the interest and usage of the EFEHR services and data for these countries. Users from India or China might be interested in the hazard products of EMME14, whereas users from the Russian Federation might be interested in both ESHM13 and EMME14. This summer, the hazard model for Russian Federation has been completed as a joint collaboration between ETH and GEM. Albeit this hazard model is not yet available on EFEHR, hazard-relevant documentation on EFEHR might have triggered the attention of some users. However, this is just an attempt to explain the access of the EFEHR web-platform and a more detailed survey is needed. Such a survey and monitoring system must be implemented.

The figures above correspond to “http page loads”. Here, loading a data application counts as one page load, independently of how many data, e.g. hazard curves, are downloaded afterwards. However, to reset the application after one single download, a user may either use the application’s reset button or just reload the entire application. Inadequate scripting of interactive access may contribute to some peculiarities in numbers of requests per country.

Figure 2: Spatial distribution of the user access of EFEHR web-portal per country in Europe between June to August 2018.

2.1.2 Data access (hazard results)

This section provides the information on how the web service API was used: operations to either retrieve hazard data, or meta-information (availability of data, covered parameter space). Most data access refers to hazard maps (five times more interactions compared to both hazard curves and hazard spectra, Figure 5). Source of the requests are interactive use of the viewer applications on the EFEHR portal, as well as direct, scripted service requests

The time-history charts in Figure 5 show access rates to the different web services per topic, from January 2016 to July 2018. Spikes can have different reasons:

- Spikes in hazard curves and hazard spectra are typically due to scripted trials to harvest entire result set for hazard models, grid point by grid point, and (for hazard spectra) probability of exceedance per probability of exceedance. These harvesting events are few; typical users just download data for their current site of interest. This documents a modern understanding of the service architecture by many users: rather than trying to get everything to their local storage, whether required or not, they trust in the availability of the services and implement a lean just-in-time access to the data that are actually required.
- As maps are offered via the global OGC WMS service interfaces, they may easily be included as layers to basically any other web mapping portal worldwide. We guess that a large share of our requests come from this type of usage for selected hazard maps, rather than from users interactively evaluating the parameter space of a hazard model on our web application before looking up a map. A hint for this is the drop of access rate in December 2016: at that time, we

moved our infrastructure, and, while maintaining the DNS name, the IP address of the service provider changed. Thus, we had a large drop in requests coming from portals which configured the path to our hazard layers by IP. At the same time, we lost customers coming from these remote applications to our own viewer (e.g. by data source links and references). The few months following that period predominantly depict requests coming from our own hazard map viewer only.

The share of data accessed via the interactive viewers, compared to direct programmatic access (e.g. programmatically querying EFEHR's database) using the Representational State Transfer (REST) API, varies widely, along with the overall usage (Figure 7). The few users requesting typically large numbers of data sets within short time using client software are widely responsible for the spikes, while the number of interactive accesses is similar, but more evenly distributed in time and coming from more users. A typical (median) user submits 30 service requests in a single session.

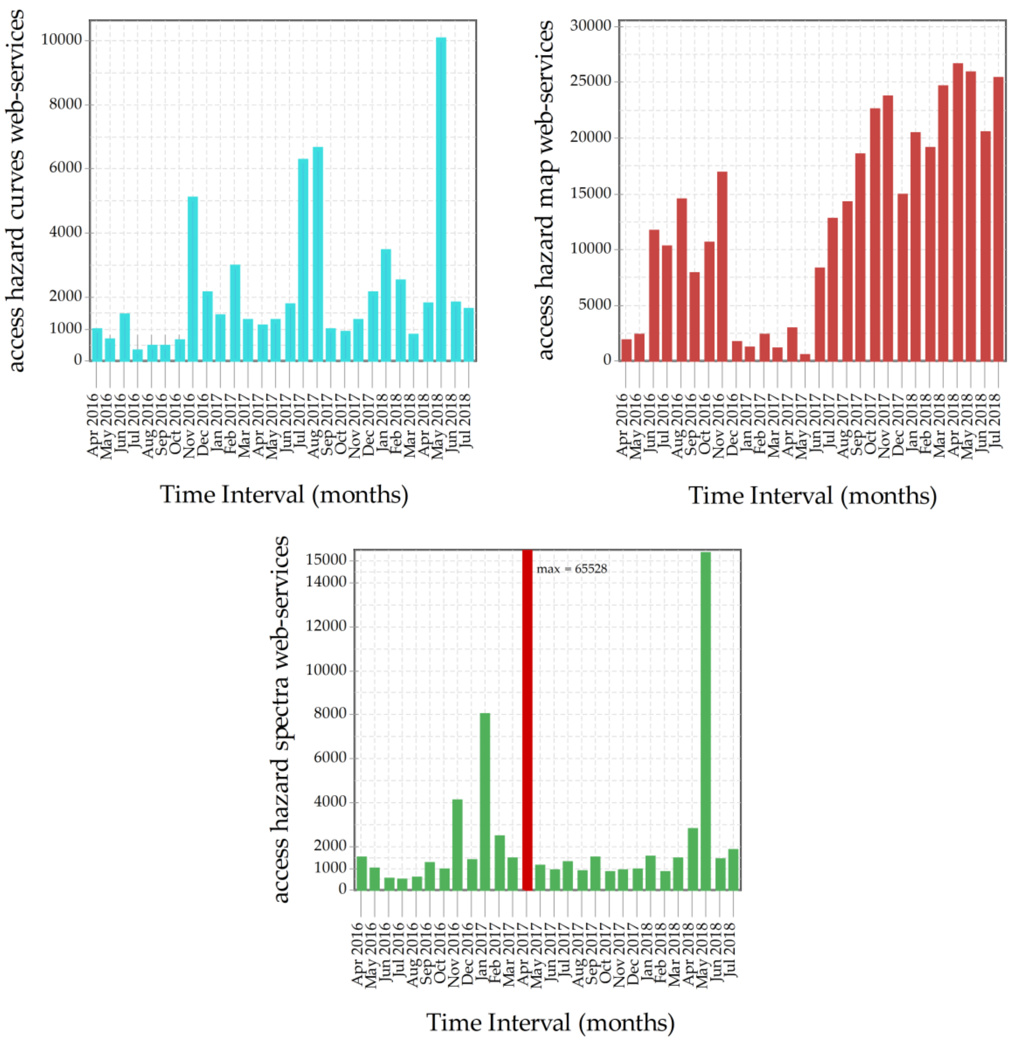


Figure 3: Frequency of web service calls between January 2016 and August 2018 – the histogram indicates that the most frequent web-portlet is the hazard map, followed by the hazard spectra and hazard curves

2.1.3 Impact Assessment

The data, models and results provided by the EFEHR platform has a strong impact on various communities from seismology to engineering, public and decision makers. To quantify the impact of the EFEHR usage to the community, the relevant citations are queried from various scientific portals and listed below. Note that the use of any dataset provided by EFEHR web-platform should be cited with the Giardini et al (2013) – online resources reference. However, there are many publications that refer to data obtained from EFEHR web-portal as data of SHARE Project (www.share-eu.org) without any reference or citation to neither Giardini et al (2013) or Woessner et al (2015).

With Google Scholar, the following numbers are counted from 2013 when the ESHM13 data and models were released through the EFEHR web-platform:

- 82 citations of ESHM13 online resources available via EFEHR web-platform as Giardini et al (2013)
- 144 citations of the ESHM13 data sets as Woessner et al (2015)
- 110 citations of www.efehr.org
- 170 citations of data and models of www.share-eu.org

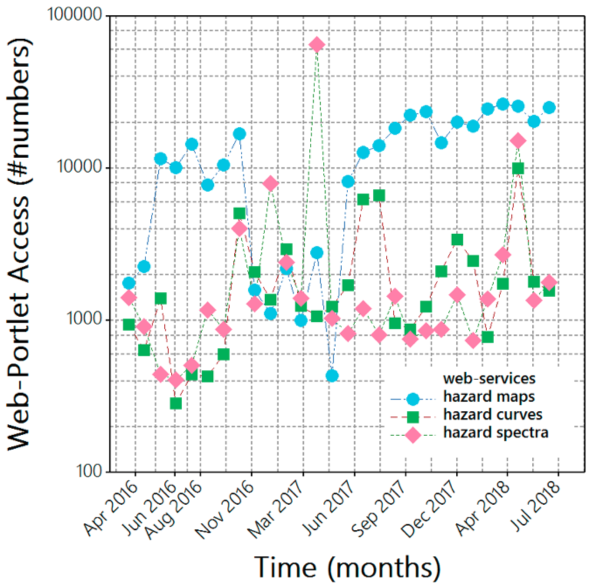


Figure 4: Web service access rate between April 2016 and June 2018

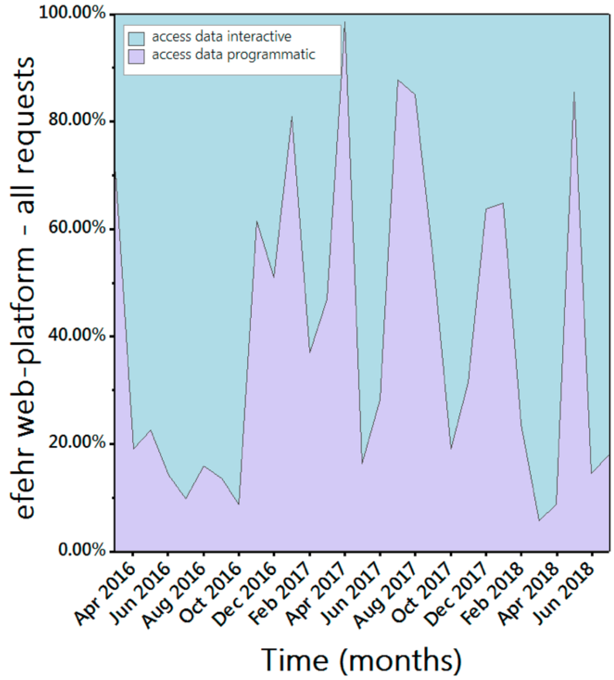


Figure 5: Percentage ratio between the programmatic versus the interactive access of the EFEHR services within April 2016 and June 2018

2.2 Other: EPOS/SERA Metadata and Data Policy

All existing web-services of EFEHR will be mapped through the EPOS DCAT-AP (https://github.com/epos-eu/EPOS-DCAT-AP) by the end of 2018. For all webservices and data we are also working on adhering to the EPOS metadata format, ensuring that DOIs can be assigned to all products as well as CC BY SA v4.0 open data licenses, in order to meet the requirements of the EPOS Data Policy. A preliminary EFEHR Data Management Plan is ongoing, and some aspects are given in the next section.

2.2.1 EFEHR Data Management Plan

Data Summary

EFEHR data and models are collected from completed scientific projects for long-term archiving, documentation, accessibility and use in research. Data covers the following domains:

- Observed seismicity / earthquake catalogs
- Seismicity models, including parametrized active faults, seismic zones, and gridded seismicity models, ground motion prediction models, and logic trees of them
- Resulting probabilistic earthquake hazard data, expressed in hazard maps, curves and spectra with reference to different
- Hazard model documentation
- Expert information and best practice documents for probabilistic hazard assessment.

All data comes with original research reports, but homogenized data formats and representations (Natural hazards' Risk Markup Language (NRML), OGC Web Map Services (WMS) response, geospatial vector data format (shapefile, https://en.wikipedia.org/wiki/Shapefile) for geographic information system (GIS) software, open standard to represent seismological data (QuakeML, https://quake.ethz.ch/quakeml/), a variant of eXtensible Markup Language (xml) and tabular ASCII data). Documentation, products and services are mostly addressed to a well-informed public, including researchers and engineers.

FAIR Data Policy

Entire datasets identified by *doi* identifiers and full respective metadata sets. For access to individual data points, a set of EFEHR-specific discovery services (RESTful web service API) is offered. The API is documented in WADL and explanatory text at http://www.efehr.org/en/Documentation/web-services/. An EPOS-DCAT-compatible secondary documentation is in preparation and planned to be available by the end of 2018

Data is accessible as bulk downloads - entire hazard models in OpenQuake input NRML format and maps (ESRI shape format), as individual data points, hazard curves and spectra discoverable and retrievable from a web service API in NRML format, and (for spatial data) via OGC standard WMS (web map service) interface. Thus, all data is in well documented community or industry standard formats. Standard access services are used as far as available; otherwise standards have been defined for EFEHR. All data holdings are freely accessible to unregistered users under the license agreed on with the originating project / initial provider. EFEHR tries to homogenize agreements to the Creative Commons - CC BY SA v4.0 (https://creativecommons.org/licenses/by-sa/4.0) open data license.

Allocation of Resources

Standard technical operation, knowledge transfer, expertise and a basic infrastructure for quality-controlled seismic hazard/risk assessment are covered by EFEHR being one of the long-term strategic pillars of the European Plate Observatory EPOS, and backed up by ETH/SED. IT support and maintenance is covered by standard SED IT operations and the respective maintenance and 24/7 service team. Project specific scientific support, research task, and development of new services follow a long-term plan, but are financed on a project basis.

Data Security

Data preservation and disaster recovery is granted by two daily off-site backups of both data holdings and virtual service infrastructure, with a preservation time of 3 months. Database integrity, service stability and access control is granted by a 3-layer system architecture (database/data holdings <-> access services <-> web layer) with firewalled interconnects and full logging on the upper two layers.
Service continuity is supported by a 24/7 it monitoring & intervention team at best effort. However, there is no formal service availability level is guaranteed.

Ethical Aspects

EFEHR does not hold individual or personalized data in its scientific content, nor request or log such data from users.

3 References

Basili R., et al., (2013) The European Database of Seismogenic Faults (EDSF) compiled in the framework of the Project SHARE. http://diss.rm.ingv.it/share-edsf/, doi:10.6092/INGV.IT-SHARE-EDSF.

Giardini G (1999), The global seismic hazard assessment program (GSHAP)-1992/1999, Annals of Geophysics 42 (6)

Giardini D. et al., (2013) Seismic Hazard Harmonization in Europe (SHARE): Online Data Resource, http://portal.share-eu.org:8080/jetspeed/portal/, doi: 10.12686/SED-00000001-SHARE, 2013.

Grünthal, G., Wahlström, R., Stromeyer, D. (2013), The SHARE European Earthquake Catalogue (SHEEC) for the time period 1900-2006 and its comparison to EMEC. Journal of Seismology, 17, 4, 1339-1344, doi: 10.1007/s10950-013-9379-y.

Danciu L., Sesetyan K., M. Demircioglu, M. Erdik and D. Giardini (2016) OpenQuake input files of the Seismogenic Source Model of the 2014 Earthquake Model of the Middle East (EMME-Project), doi:10.12686/a3

Delavaud, E., et al. (2012) Towards a ground-motion logic tree for probabilistic seismic hazard assessment in Europe, J. of Seismology, doi:10.1007/s10950-012-9281-z

Stucchi et al., (2012) The SHARE European Earthquake Catalogue (SHEEC) 1000–1899 (2012). Journal of Seismology, doi: 10.1007/s10950-012-9335-2.

Yenier E, Sandikkaya M.A, and Akkar S (2010), Fundamental Features of the Extended Strong-Motion Databank prepared for the SHARE Project, Deliverable D4.1.

Wiemer S, Danciu L, Edwards B, et al (2016) Seismic Hazard Model 2015 for Switzerland (SUIhaz2015), Official Report of the Swiss Seismological Service). doi: 10.12686/a2.

4 Appendices: Details of EFEHR Portal Design

The EFEHR web-portal consists of several components summarized hereinafter. The key components are:

- Database server – with Postgresql 9.3 (https://www.postgresql.org), postgis extensions and daily backup
- Java/Tomcat/OpenCMS (http://www.opencms.org/en/development/installation/server.html) based web content management system
- Standalone interactive data viewers in html/javascript (Ext Js, GeoExt, OpenLayers – see http://presentations.opengeo.org/2012_javascript/javascript/concepts.html)
- Map server (https://live.osgeo.org/archive/6.5/it/overview/mapserver_overview.html) implementing the OGC web map service standard.
- A REST (Representational State Transfer) Web Service provider implemented in Java (https://docs.oracle.com/javaee/6/tutorial/doc/gijqy.html)

SERA
Seismology and Earthquake Engineering
Research Infrastructure Alliance for Europe

DELIVERABLE ANNEX: VA5

VA5: Access to data and products of anthropogenic seismicity of IGPAS

Work package	WP22 VA5: Access to data and products of anthropogenic seismicity of IGPAS
Lead	IGPAS, ACC CYFRONET AGH
Authors	Monika Sobiesiak, Stanisław Lasocki, Paweł Urban, Dominika Wenc
Reviewers	Marcelo Assumpção, University of São Paulo

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730900.

Table of Contents

Summary

The IS-EPOS platform (https://tcs.ah-epos.eu/) is a web-based utility which provides a unique collection of anthropogenic seismicity data, paired with industrial production data in hydrocarbon extraction, geothermal energy exploitation, underground mining, water reservoir impoundment and experimental data. The platform supports research through dissemination of the collected data to a wider community and through the development of specific software tools (applications) for statistical and waveform analysis. The platform structure offers an own web-based workspace for each user together with respective data handling, processing, resource management and visualization tools.

1 Introduction

The ever growing need of energy and natural resources for industrial production have stirred at the same time a growing awareness of its negative consequences: Damaging and disagreeable effects like induced or triggred seismicity in areas which are usually aseismic, water level changes, groundwater contamination or even landslides (Froude[1] et al., 2018) are just some of the observed phenomena felt by populations or the wider public. As cause of these changes in Earths' structure, new exploitation methods like unconventional hydrocarbon exploitation through fracking methods, enhanced geothermal energy exploitation and CO2 sequestration, but also the "classical" underground and surface mining activities like conventional hydrocarbon extraction and the impoundment of surface reservoirs for liquids are widely recognized (Davies[2] et al., 2013, McGarr[3] et al., 2002, among others). In Figure 1, an overview on the published number of hazardous induced or triggered seismic episodes with a broad magnitude range between M1 and M7.6 assigned to the various anthropogenic activities are given. This histogram further shows, that induced or triggered micro- or large scale seismicity ranks among the most recognized and important responses to anthropogenic activities.

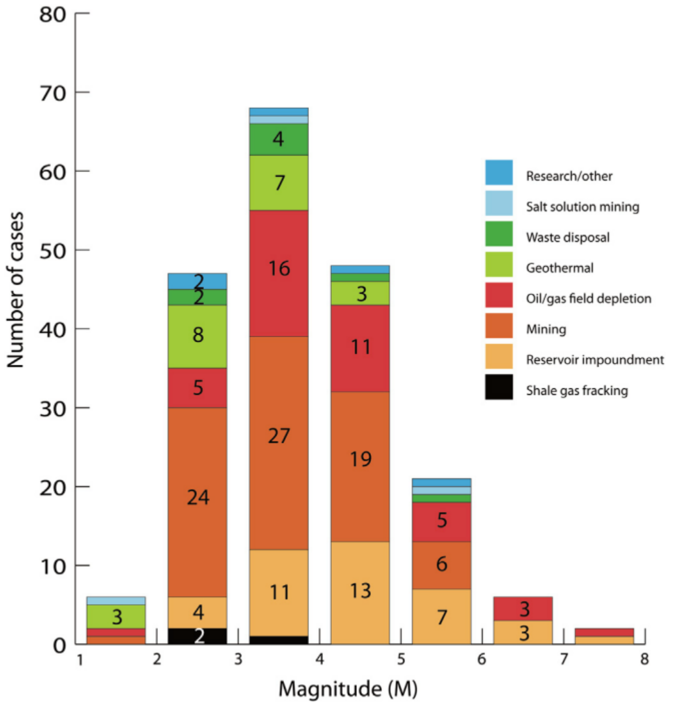


Figure 1: Published anthropogenic hazard cases since 1929, modified after Davies et al., 2013.

These increasing interactions between anthropogenic activities and concerned population requires more and more first-hand objective information about the observed response processes and open quality communication and exchange between involved groups such as scientists, engineers, industry, governmental and public entities, politicians and last but not least, the public.

Research in anthropogenic seismicity and hazard has been intensified in the last 30 years which led to new insights into the processes behind induced and triggered earthquakes associated with the above mentioned anthropogenic activities. New methods in monitoring, processing and analysis of data sets which take into account the specific conditions of monitoring underground, detection of micro-seismicty, analysis of source mechanisms of small earthquakes, often with low signal-to-noise ratios, could be brought on the way (Grigoli[4] et al., 2017, McGarr[3] et al., 2002). However, there are still pressing issues to be solved like improved hypocenter locations for small scale seismicity, real-time solutions for control systems in mines, discrimination of induced, triggered and tectonic seismicity to help on liability debates.

The IS-EPOS platform of Research into Anthropogenic Seismicity and other Anthropogenic Hazard (**https://tcs.ah-epos.eu**) is designed to serve the needs for research, information and knowledge transfer between science and industry, as well as expert information and education for the interested public. Within the above described versatile context, the platform is a unique collection of anthropogenic seismicity data sets combined with production data such as water injection rates , well head pressure and temperature. For investigations on causes for response processes, this data is indispensable. Such combined data sets are still sparse as industry data often is restricted because of private ownership or liability concerns. Therefore, dissemination of such kind of data through the platform to a wider community is a valuable contribution to the research in this field.

The platform, initially a product of IS-EPOS Polish national project, is presently being further developed in the framework of EPOS IP H2020 as part of the Thematic Core Service of Antropogenic Hazards (EPOS TCS AH, www.epos-ip.org). Within projects like SERA, the platform provides virtual open access to its resources, supporting in such a way research and investigation topics as targeted in SERA Joint Research Activities (JRA1-5).

1.1 Concept of the platform

As stated in the Introduction, the conceptual idea behind the platform is to serve as a utility which fosters interaction and exchange between the protagonists, science, industry, public and decission makers. This includes transfer of knowledge like scientific results i.e. from industry to science and vice versa, information on actual topics and events to the public and decission makers. In such a way, common projects could be enhanced where results contribute to the needs of all protagonists. This basic concept is sketched in Figure 2.

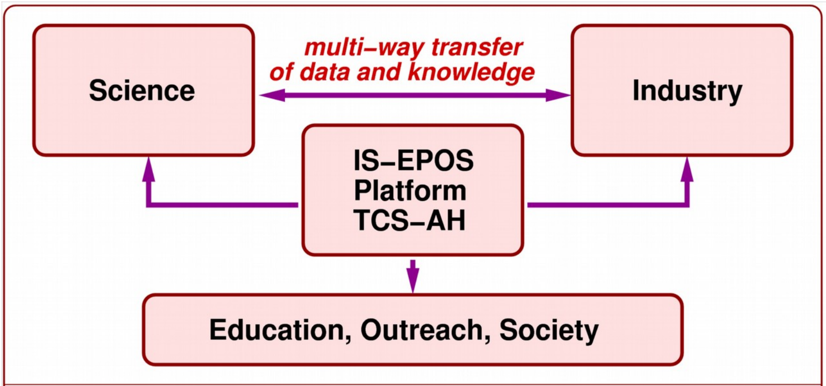


Figure 2: Sketch of the platform concept, afer Leptokaropoulos, 2017, EGU

1.2 Structure of the platform

Three major components form the structure of the platform: 1) the comprehensive data base of combined seismological, industrial and geologic data merged into episodes, where one episode describes the response to a single anthropogenic activity, 2) a pool of software and program packages to support research and analysis of respective data where each enclosed software contribution is called 'application', and 3) the web-based workspaces for each user to carry out respective studies and investigations with tools and data from the platform or own uploaded data. This structure is sketched in Figure 3. Furthermore, the platform hosts a comprehensive document repository of publications, manuals and overviews which are linked to the respective episodes and applications.

Physically, the platform is located at the Academic Computer Center CYFRONET at the AGH University of Science and Technology (ACC CYFRONET AGH) in Kraków, Poland, where it is technically developed and maintained through a team of IT scientists and engineers.

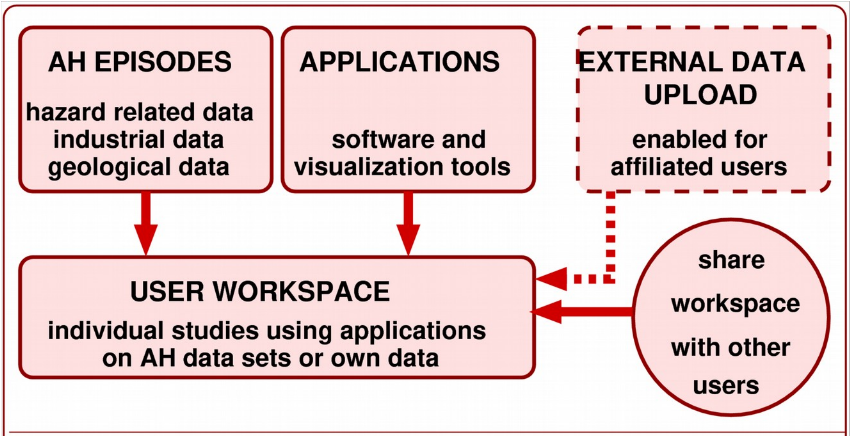


Figure 3: Major components of the platform structure after after Leptokaropoulos, 2017, EGU

1.3 Accessing the platform

IS-EPOS platform is accessible through https://tcs.ah-epos.eu. Many platform resources are open to the public. However, to make use of the platforms' full service, the user should go through the registration process which is three-fold:

- **As anonymous user:** access to information about available Episodes and applications is provided and it is possible to browse the document repository. However, it will not be able to view the data itself nor use the applications.
- **As registered user:** access to data from the episodes is provided, but the user still will have no access to the workspace nor data download.

- **As registered user with institutional or project/group affiliation:** access to the workspace is given where applications can be used. It is possible to download and upload data.

2 Components of the Platform

As mentioned before, the structure of the platform consists of three major components, which are 'Episodes' (anthropogenic hazard specific data sets, 'Applications' (software and tools for analysis and handling of data), and the 'Workspace' (for each user to combine data and applications to work on individual research targets).

2.1 Episodes

In total, there are 21 episodes containing seismological, industry production data and geological data (see Table 1). 14 of these episodes are open to all registered users and 7 episodes have still restricted access for members of respective projects. Each data set represents an episode related to an anthropogenic activity site. These sites are located in 11 different countries worldwide.

There are permanently new episodes from other sites and countries coming in, new data can always be added to the platform. However, certain quality tests are applied before the new data set is uploaded as well as formats are controlled or transferred when necessary to respective required formats.

Episode	***Type of activity***	***Country***
Bobrek Mine*	longwall mining for coal	Poland
Czorsztyn*	shallow water reservoir	Poland
GISOS-Cerville	underground solution mining (exp.)	France
Groningen Field*	convent. Hydrocarbon production	Netherlands
Gross Schoenebeck*	geoth. Energy production experiment	Germany
LGCD*	underground copper mining	Poland
Lubocino*	shale gas exploitation	Poland
Preese Hall*	shale gas exploitation	Great Britain
Pyhäsalmi Mine	in-situ underground laboratory	Finland
Song Tranh*	deep water reservoir	Vietnam
Starfish	underground gas storage	France
The Geysers*	geothermal energy production &	USA
The Geysers Prati 9/29	treated wastewater injection	USA
USCB*	underground coal mining	Poland
Val d'Agri	water reservoir	Italy
Val d'Agri Field	conventional hydrocarbon extraction	Italy
Wysin*	shale gas exploitation	Poland

CARBFIX	geoth. energy production, CCS	Iceland
Monteynard	water reservoir	France
Oklahoma*	hydrocarbon extraction, wastewater inj.	USA
St. Gallen	geothermal project	Switzerland

Table 1: All episodes which can be found on the platform. Red star marks the episodes which are fully open to all users to whom virtual access is provided.

2.2 Applications

Programs, software and software packages are called 'Applications' on the platform which are ready-to-use for data analysis. The user needs to transfer or upload a data set (whole episode or part of an episode) to their workspace and do the same with the application selected for individual studies. When the application is executed on the workspace, all results and resulting graphics will be displayed on the screen and stored as data files in the workspace. As an example, in Figure 4 (after Chapter 2), you can see a 3D distributuion of seismicity in Bobrek Mine (Poland). A list of application names and the number of documents and publications referencing the respectice application, are given in Table 2.

In general, the applications are sub-divided into four groups:

- data handling applications (include catalog filtering, extraction of parameters, seismic phase picking, conversion of formats, download and upload tools etc.)
- data processing applications (include autocorrelation tool, cross correlation, tool for focal mechanism and moment tensor calculations, inter-event time distribution analysis, estimation of maximum credible magnitude, etc.)
- resource management applications (include data catalog format change, i.e. from matlab to ascii, export of matlab files as XLSX spreadsheet, etc.)
- visualization applications (like histograms for mining front advances, 3D seismicity hypocenter distributions of which an example can be seen at the end of this chapter in Figure 4, seismicity related to mining front advance, etc.).

Table 2: Names of applications and number of references (in brackets) of the respective application

- Coefficient of randomness (2)
- Completeness Magnitude Estimation (7)
- Cross correlation (2)
- Earthquake interactions-georesource scale (3)
- Earthquake interactions-mainshock scale (3)
- Earthquake swarm: reshuffling analysis (3)
- Effective Stress Drop Estimate (2)

- Fracture Network-Models Mechanical Stresses (7)
- GMPE (5)
- Inter-event Time Distribution Analysis (5)
- Magnitude Conversion (6)
- Moment Tensor Inversion (7)
- Risk Assessment (5)
- Seasonal trends (1)
- Seismic Hazard Assessment (9)
- Source Location (4)
- Spectral Analysis (3)
- Stationarity test (4)
- Stress and strain changes induced by fluid injection and temperature change driven by geothermal injection (1)
- Stress Inversion (5)
- Autocorrelation

Most of the software codes behind the applications are open source which is the general policy for software contributions on the platform. A number of researchers from different international institutions have contributed to the application list.

2.3 User’s Workspace

The *workspace* is the user’s individual area to work on with the data sets with software tools provided by the platform. However, the user also has the option to download data to an own device or upload additional data to the workspace from an external account. Thus, the workspace provides to the user a web-based tool for data analysis which can be used interactively on episode data from the platform or on own data. In order to start an analysis, both, the selected data plus the chosen application have to be transferred actively to the workspace. Here, software tools and data sets, can be stored in different directories which are defined by the user. The applications are activated in the workspace, where results are displayed as single values in files or as graphs shown after a succesful run. Both output options can be either stored in the workspace directories or can be downloaded to own devices. Summarizing the advantages, the workspace provides:

- fast overviews on data properties and parameters
- upload and download tools
- storage for data and results
- flexible usage of applications
- detailed analysis of data
- visualization of data and results.

In order to facilitate collaboration for projects or also individual cooperations between researchers, a share function has been implemented. With this share function, users can provide access to selected parts of their own resources to other researchers on the platform. This access can be bi-directional.

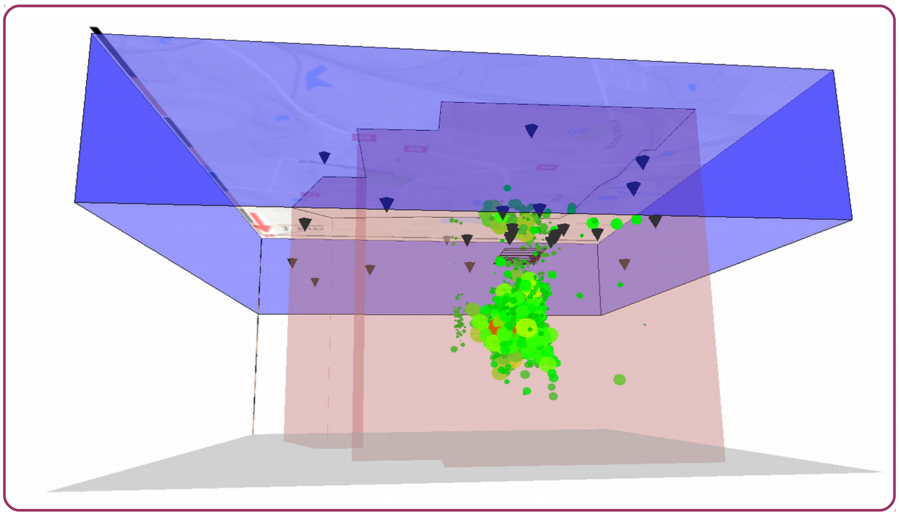


Figure 4: Example of clustered induced seismicity from Bobrek Mine (Poland). The light red colour marks the outer limits of the mine, the light purple rectangle gives one layer of the velocity model, green, yellow, and red circles mark the earthquake hypocenters. The figure was generated with the help of the 3D visualization tools of the platform.

3 Development of the platform

By the start of the SERA project, IG PAS and ACC Cyfronet AGH ensured a virtual access to the IS-EPOS platform. Until that time, the number of registered users was in total 564. From the start of SERA to 31st of August 2018, 191 additional registrations of users were counted. Thus the total number of users up to date is 755. This 30% increase in registered users was certainly influenced by promoting the platform in workshops on induced seismicity like COST Action TIDES (TIme Dependent Seismology) in March 2018 in Bologna, Italy, the SERA JRA1 meeting in January, 2018, in Zurich, the demonstration of platform functioning in April 2018 at the EPOS exhibition booth at the EGU in Vienna, the presentation of the platform at the same EGU conference in the session 'Integrating data and services in solid Earth sciences' and the joint SERA JRA1/JRA2 workshop in September 2018 in Kraków. The list in the following is giving several important numbers on the actual status of user engagement:

- 755 platform users (564 in mid 2017)
- 22888 displayed files (9541 in 2017)
- 394 downloads of waveform files (175 in 2017)
- 622 downloads from workspace (145 in 2017)
- 841 files added to workspace (354 in 2017)
- 480 uploaded files to the platform (209 in 2017)
- 5 shared files

These numbers demonstrate an increase in activity on the platform in all sectors, like usage of data, use of workspace and research activities. The number of shared files is still low which is due to the just recently implemented tool.

3.1 User statistics

The following histograms and the maps of the global distribution user registrations and logins demonstrate a usage of the platform in daily schemes. In total, users come from over 28 countries on all 5 continents.

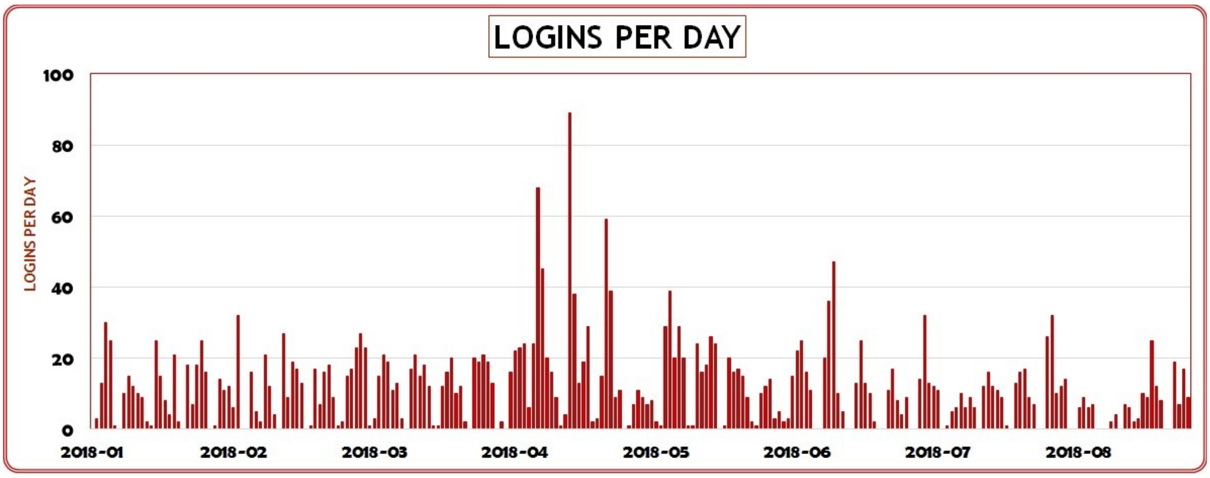


Figure 5: Daily use of the platform. The peak in April was due to the demonstrations of the platform during EGU.

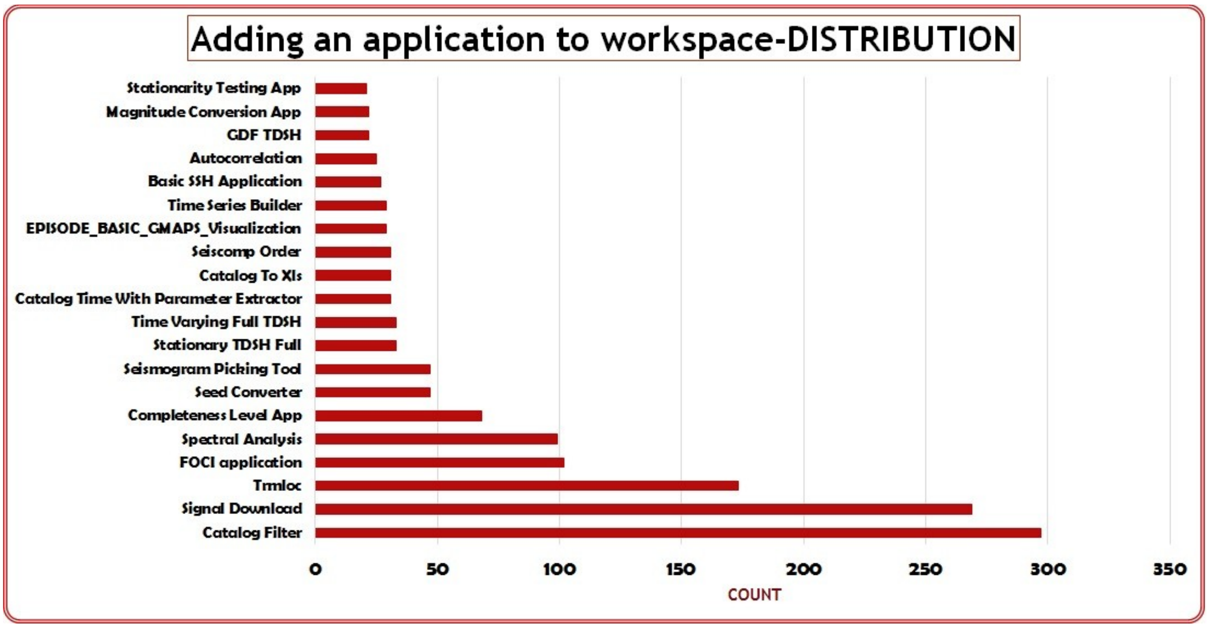


Figure 6: The histogram shows the most frequently used applications.

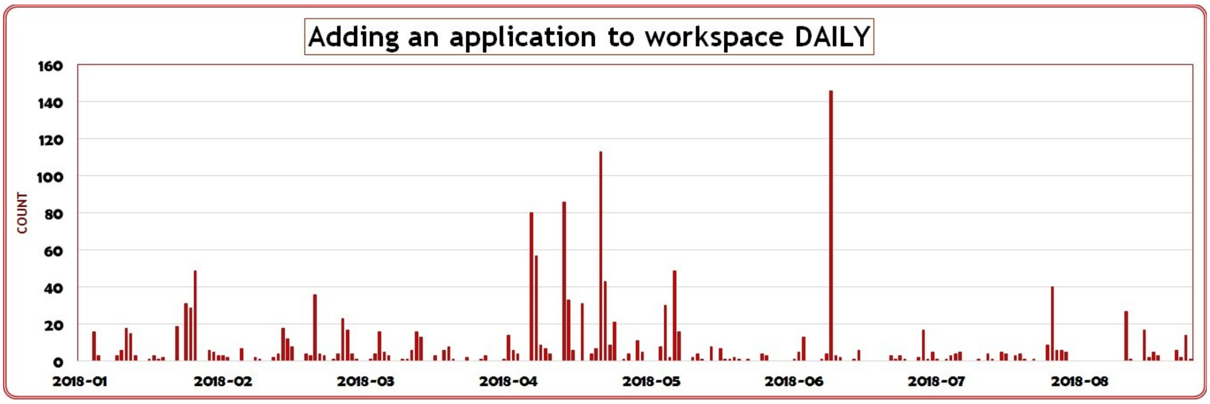
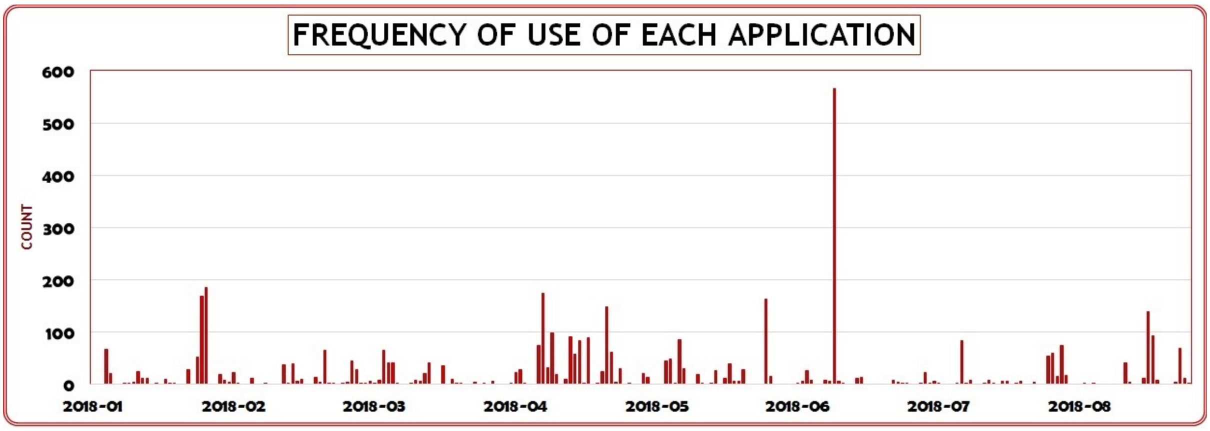


Figure 7 and 8 (below): The histograms are showing how users transfer applications daily to their workspace (upper). The histogram below reflects the daily frequency of use of the respective applications.



3.2 Concluding remarks and outlook

The statistics as shown in the former Chapter point out that the IS-EPOS Platform of Anthropogenic Hazard has become a considered tool in the area of anthropogenic seismicity and hazard cases. All important statistical values experienced a significant increase during the last 12 months. To summarize the advantages of the platform use, the following listing is given. The platform provides:

- an overview on anthropogenic hazard cases with induced or triggered seismicity involved
- tools for downloading open data and uploading own data to the respective workspaces
- options to change formats both for waveform files as well as matlab result files
- open source applications
- share function for sharing data or actual results with other researchers on the platforms
- visualization tools for 3D animation to achieve information on the spatial and temporal distribution of seismicity.

With the aim to further enhance the flexibility of the platform for advanced requirements of the users, the following issues will be targeted in the future:

- providing a tool for creating shared workspaces among members of specific projects
- facilitating interactive work on developing new applications
- possibility for downloading applications for stand-alone use on own computers.

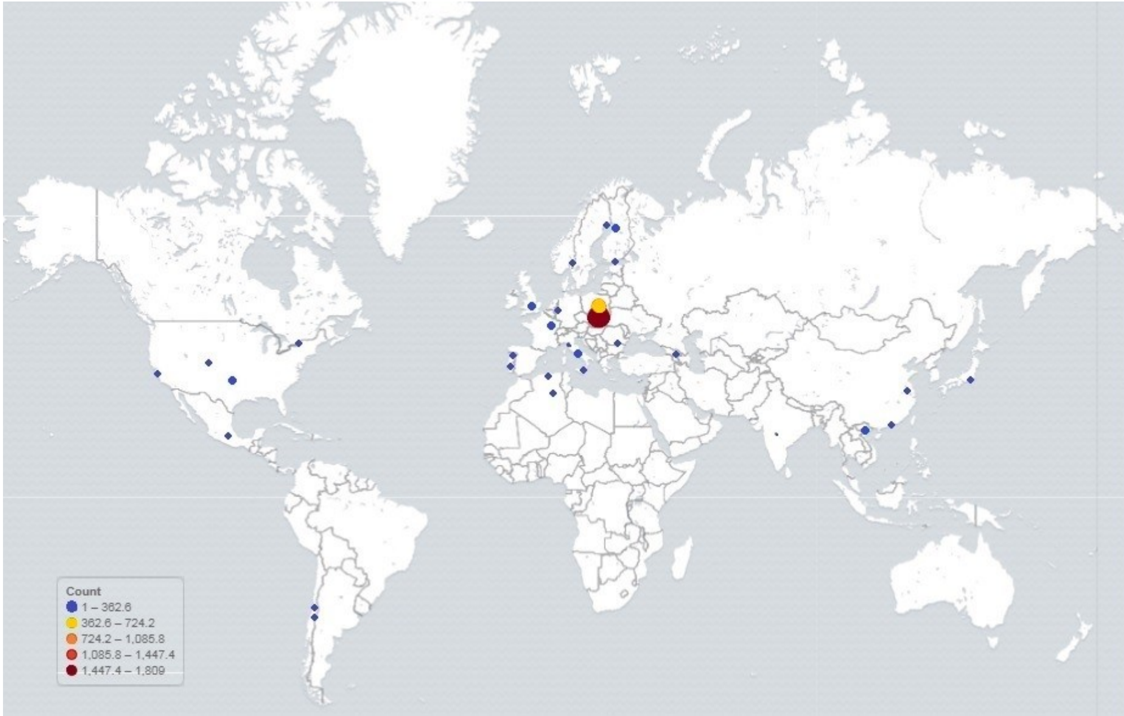


Figure 9: Map of global user logins

1 References

1) Froude, M.J., and Petley, D.N., Global fatal landslide occurrence from 2004-2016, Nat. Hazards Earth Syst. Sci.,18, 2161-2181, https://doi.org/10.5194/nhess-18-2161-2018, 2018.

2) Davies, R., G. Foulger, A. Bindley, P. Styles, Induced seismicity and hydraulic fracture for the recovery of hydrocarbons, Marine and Petroleum Geology, 45, 171-185, http://dx.doi.org/10.1016/j.marpetgeo.2013.03.016, 2013.

3) McGarr, A., D. Simpson, and L. Seeber, Case histories of induced and triggered seismicity, Int. Geophys. Ser., 81(A), 647-664, 2002.

4) Grigoli, F., S. Cesca, E. Priolo, A.P. Rinaldi, J.F. Clinton, T.A. Stabile, B. Dost, M.G. Fernandez, S. Wiemer, and T. Dahm, Current challenges in monitoring,

discrimination, and management of induced seismicity related to underground industrial activities: A European perspective, Rev. Geophys., 55, doi:10.1002/2016RG000542, 2017.

5) Leptokaropoulos, K., M. Staszek, S. Cielesta, D. Olszewska, P. Urban, and G. Lizurek, Time Dependent seismic hazard in Bobrek coal mine, Poland, assuming different magnitude distribution estimations, Acta Geophys., doi: 10.1007/s11600-016-0002-9, 2017.

Liability claim