

**Deliverable No.: 5.1** Current status report (WP5)

**Prepared by:** KONINKLIJK NEDERLANDS METEOROLOGISCH INSTITUUT (KNMI)

**Author:** Torild van Eck



**Project acronym:** COOPEUS

**Project title:** Strengthening the cooperation between the US and the EU in the field of environmental research infrastructures

**Grant Agreement number:** 312118

**Funding Scheme:** FP7.INFRASTRUCTURES-2012-1

**Project website address:** [www.coopeus.eu](http://www.coopeus.eu)

## COOPEUS WG5: Current status report (Deliverable 5.1)

This is a living document that will be updated with on-going work, specifically EPOS WG1 and collaborative work with US colleagues. References are mentioned as background material.

### 1. Introduction

The following two sections provide an overview of the current status of the research infrastructure for the seismological research community in Europe and the US. It also indicates the current known planning in both regions. It is a deliverable within the COOPEUS project: *“Report describing the current status and the background in regard to data sharing, accessing data, linking data archives, and addressing data provenance issues. The report will summarize the findings in task 5.1 and task 5.2”*. However, our second objective has been to provide a short guide to on-going EU and US developments improving seismological services and products for the EPOS and EarthScope communities.

#### *Seismological Research Infrastructure in Europe*

At present, two main services providing Europe-wide access to seismological data exist, they are:

- ORFEUS: access to broadband seismic data repository, in many cases in near realtime, quality control and other information. Distributed, open waveform access via EIDA to complete waveform archives for all data (including data from short period (limited) and acceleration networks, mobile and aftershock deployments, and ocean and sea floor deployments (limited) and infrasound).
- EMSC: rapid earthquake information and earthquake catalogue, associated origin information (picks, magnitudes) based on automated and manual analysis by national or regional networks.

Both organization provide a comprehensive overview of all permanent networks and mobile resources in Europe and its surroundings. These can be found on the ORFEUS and EMSC websites.

A third service that provides access to seismic hazard and risk data, products, models and tools is being built at the moment (EFEHR), also serving as the European regional component of GEM. EFEHR continues to serve the developments of SHARE and incorporates relevant developments of NERA.

Additional services are provided through AHEAD (macroseismic and historical archives), SISMOS (digitized historical records), and a number of other national services. It is the intention to integrate those within EPOS WG1 (seismological observatories and research infrastructures). The European computational seismology community is working on integrating modeling products such as e.g., fault models, velocity models, synthetic waveforms, tomographic models.

A comprehensive overview of the EPOS seismologic research infrastructure is being compiled within EPOS WG1. This overview is part of a more comprehensive European planning consisting of a European Earth Science roadmap ([www.bgs.ac.uk/EarthScienceEurope](http://www.bgs.ac.uk/EarthScienceEurope)) and an EPOS science plan. All these documents are planned to be finalized by September 2013. Below, a first and incomplete draft of the EPOS WG1 document is presented. It is more focused on the technical and organizational status and challenges facing a comprehensive European seismological RI service.

## 2. European seismological research infrastructure: current status and developments.

### 2.1 Background

The seismology infrastructure is generally built in each country funded by national resources and, in quite a few cases, by the industry or other private partners. Besides international coordination (ISS / ISC), which started in 1918 on European initiative, European scale coordination was established by EMSC in 1975, concerning parameter data and ORFEUS in 1987, concerning broadband waveform data. Both are non-profit organizations registered in respectively France and the Netherlands, and have different funding schemes based on voluntary contributions by national organizations and individual institute participants [reference]

Since the beginning of this century, the seismology infrastructures in Europe rapidly developed; exponential increase of openly available digital waveform data of many different flavors, real-time waveform data exchange, rapid accurate and automatic location and earthquake parameter determinations, significantly increased computational resources and obtained models. Data products (level 1), i.e. data derived from raw observations (level 0), are becoming increasingly important, as direct input to next stage products (level 2 and above) earthquake catalogues, earthquake characteristics, structural models, etc.).

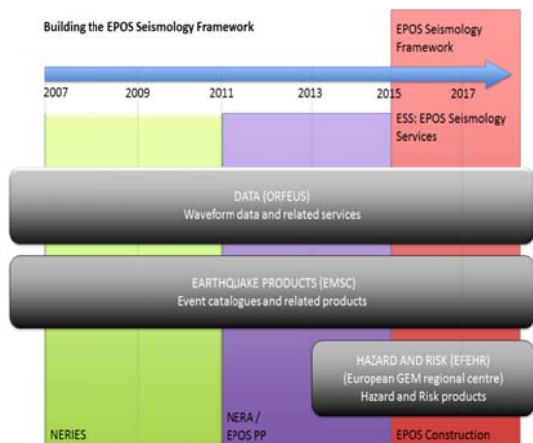
The seismological RI in Europe and its immediate surroundings operate more than 1000 broadband stations, around 2000 SP seismometers, about 1000 high quality mobile stations, more than 3000 accelerometers. More than 100 observatories and research institutes monitor and report seismicity. This excludes dedicated small network operations monitoring induced seismicity, waste disposals, NPP, etc.

Most research infrastructure accomplishments are funded on a national level. However, since around 2000, significant EU resources have enabled a stronger European integration of all these accomplishments (projects like MEREDIAN – FP5, NERIES – FP6, NERA, SHARE, REAKT – FP7). Out of these projects the EPOS concept emerged and the current European seismological Research Infrastructure.

The current European state of the infrastructures can be roughly classified in three groups:

- I. Seismological data, including primary waveforms, metadata and primary measures (arrival times and amplitude of seismic waves) and building on existing network coordination initiatives, mainly, but not only, within the ORFEUS framework. [data level 0 and 1].
- II. Seismological products, including earthquake parameters (location, magnitude), seismic bulletins and (historical) catalogues, earthquake alerts and building on existing coordination efforts, mainly, but not only, within the EMSC framework. [data level 1 and 2].
- III. Products and services in seismic hazard and risk, including base data for modeling (catalog of active faults, GMPs, building vulnerability functions, etc) and modeling tools, hazard & risk maps and scenarios, and building on on-going initiatives in SHARE and NERA and within the EFER framework [data level 2 and 3].

While this represents a simplified classification of scattered initiatives in Europe, it classifies fairly the current and hopefully its future coordination structure. The schematic planning for EPOS is shown in Figure 1.



**Figure 1.** Schematic view of the different categories in the seismological thematic services for EPOS WG1 and a possible time line for integration within EPOS.

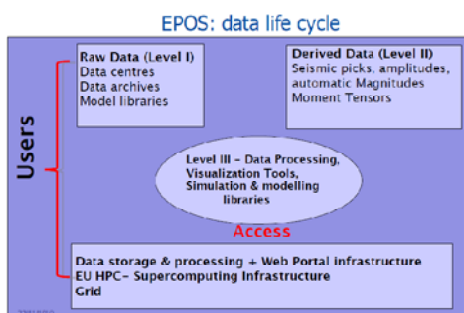
## 2.2 Current status within a global perspective

The European seismological RI initiatives are well embedded in global coordination frameworks. Digital waveform data exchange coordination between observatories and the research communities takes place within the International Federation of Digital Seismograph Networks (FDSN). In practice ORFEUS, acting as the regional coordinator, discusses with IRIS-DMC on waveform data. Earthquake information/locations global coordination takes usually place between the USGS in the US, EMSC in Europe and the International Data Centre (ISC) at a global level.

The Comprehensive Test Ban Treaty Organization (CTBTO), a politically created organization, gathers and analyze waveform data from their ~300 station global network. Although the CTBTO is located in Vienna, coordination is hampered by their political constraints. The NE Atlantic and Mediterranean Tsunami Warning System (NEAMTWS), coordinated by the Intergovernmental Oceanographic Commission (IOC), intends to build a tsunami warning system involving major national institutes in the region. The IOC being also a political, coordinates with individual countries networks.

## 2.3 Data level definitions

Currently a set of data level definitions are being identified to facilitate services and products discussions within the EPOS framework. They are intended to follow other EarthScience definitions where possible. Figure 2 provides a first indication as discussed in 2012.



**Figure 2.** First indication of three different data level definitions being worked out within the European seismological community in 2012.

## **2.4 On-going (technical) developments in Europe**

Both national, European and global developments aim at improving the seismological RI and its services. Unfortunately, driven by different funding and planning schema's, the initiatives are not naturally embedded in a coordinated framework. However, this is the current funding reality within which the community aims its utmost to coordinate the different developments.

An overview of developments as funded on-going and, generally, for the next 2-3 years framework is given within the context of a) the three categories, and b) the EPOS thematic and integrated core services.

### **2.4.1 Developments within the seismological RI categories**

#### *Category I: Seismological data; waveform data and related products*

Key developments are SeisComp3 (important monitoring network acquisition and analysis package) being widely distributed and promoting (real-time) waveform data exchange. To ensure access to an increased amount of data the European Integrated waveform Data Archive (EIDA) has been created under the umbrella of ORFEUS. Data access is improved by the implementation of ArcLink and standardized Webservices. The data holdings have been extended beyond broadband data to include data from other sensors, temporary deployments, acceleration and strong motion data, historical seismograms, ocean and sea floor networks, synthetic seismograms, and infrasound data. The archiving extension to other types of data is mainly realized through the NERIES and NERA ([www.nera-eu.org](http://www.nera-eu.org)) projects, Date service improvements are the VERCE project ([www.verce.eu](http://www.verce.eu)) addresses the is working on both the data intensive and computational intensive services. The COOPEUS ([www.coopeus.eu](http://www.coopeus.eu)) EC project provides some funding for US-EU cooperation, mainly between EPOS/ORFEUS and EarthScope ([www/earthscope.org](http://www/earthscope.org)), IRIS ([www.iris.edu](http://www.iris.edu)) and UNAVCO ([www.unavco.org](http://www.unavco.org)).

#### *Category II: Seismological products; event catalogues and related products*

Key developments are real-time products (earthquake alerts, shakemaps, ...) European earthquake bulletins and catalogues, both historical (SHEEC - [www.emidius.eu/SHEEC](http://www.emidius.eu/SHEEC)), historical data (AHEAD - [www.emidius.eu/AHEAD](http://www.emidius.eu/AHEAD)), Outreach, i.e. earthquake information, background and educational information/material has also improved significantly both on a national level and the EMSC. The REAKT ([www.reaktproject.eu](http://www.reaktproject.eu)) project is coordinating rapid alert initiatives, NERA includes school seismology, and catalogue work is done within SHARE ([www.share-eu.org](http://www.share-eu.org)) on a European scale. The EMSC ([www.emsc-csem.org](http://www.emsc-csem.org)) plays a leading coordinating role in Europe. International coordination on historical data is coordinated through ...?

#### *Category III: Services in seismic hazard and risk*

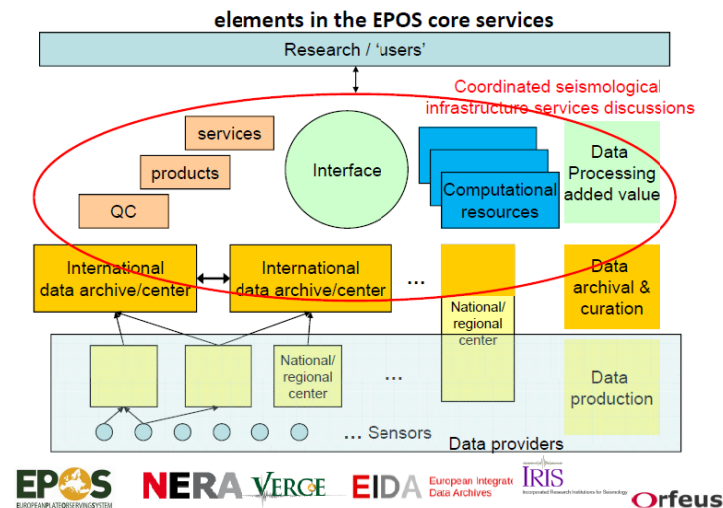
Key developments are integrated / unified European hazard and risk products as developed within the SHARE and SYNER-G ([www.vce.at/SYNER-G](http://www.vce.at/SYNER-G)) projects and carried further within NERA and GEM ([www.globalquakemodel.org](http://www.globalquakemodel.org)). Additional products & services are emerging from the multi-hazard / multi-risk focused project MATRIX (<http://matrix.gpi.kit.edu>) and the developments in time-dependent hazard and risk assessment in NERIES and REAKT. The European Facility for Earthquake Hazard and Risk EFEHR ([www.efehr.org](http://www.efehr.org)) is set up as central access and dissemination platform for these hazard and risk products. While it's long-term governance structure is yet to be established (within EPOS), it is envisioned that EFEHR also provides the coordination platform for further seismic hazard and risk related initiatives in Europe. GEM plays a key role in the international coordination.

## 2.4.2 Improving seismological services in Europe

The EPOS thematic and integrated services developments are currently funded through national and EU projects. Seismological services are in general developed in coordination with the US, but differ in being essentially distributed. Below we make an effort to identify and classify the European initiatives.

### **Integrating data archives and computational resources**

EIDA created within the NERIES ([www.neries-eu.org](http://www.neries-eu.org)) and NERA ([www.nera-eu.org](http://www.nera-eu.org)) projects requires us to address optimal data access, exchange and coherency issues. EUDAT ([www.eudat.eu](http://www.eudat.eu)), an IT infrastructure project, is addressing this in collaboration with INGV and ODC. VERCE (Virtual Earthquake and Seismology Research Community in Europe e-science environment) addresses the issue how to integrate optimally large data archives and HPC resources. This includes developments on “work flows”, “gateways”, “moving big data sets across the Internet”, “optimal DB structure”, ... . Three PRACE computer centers (CINECA, LRZ and (indirectly) EPCC) participate and connections with the QUEST EC project ([www.quest-itn.org](http://www.quest-itn.org)) and the US CIG initiative (<http://geodynamics.org/cig>). Within the Dutch IT project COMMIT the MonetDB is investigated as an effective tool in the future to realize rapid data discovery.



**Figure 3.** Current on-going developments within Category I RI and services. The encircled area indicates the area of current concern on how to realize optimal seismological services as funded through different projects, which the community aims to coordinate to create a coherent seismological thematic service in Europe.

### **Web services and the broker concept**

The NERIES project developed the first European web services for waveform data in Europe. IRIS-DMC and UNAVCO are following two concepts within The EarthCube initiative (<http://earthcube.ning.com/page/concept-awards>); the web services and the broker concepts. Within the FDSN four standard web services with regard to data type 0 and 1 have been agreed on. IRIS-DMC has an extensive web service library and plans to extend this. The four standard web services will be implemented in Europe at individual ORFEUS-EIDA nodes within the COOPEUS project. The IRIS-DMC promoted concept for distributed archives is the Federated Archives, in which a broker leads the user to the associated archives. The European ArLink ([www.seiscomp3.org/wiki/doc/applications/arlink](http://www.seiscomp3.org/wiki/doc/applications/arlink)), a related service concept developed at GFZ in Germany, has been implemented at European seismological data centers.

### **Data provenance and referencing**

Seismological data archives are considering more advanced data provenance tools, such as identifiers (DOI and PID) and Quality Control characteristics. EUDAT ([www.eudat.eu](http://www.eudat.eu)) is testing the implementation of PID's at INGV and possibly ODC. COOPEUS ([www.coopeus.eu](http://www.coopeus.eu)) is making an inventory of the identifiers in projects like EPOS and actively discussing implemented DOI's at GFZ for mobile data. IRIS-DMC is considering implementing identifiers, but, as far as we know, they don't have concrete implementation plans. The implications for implementing identifiers at seismological data centers could potentially be large, thus delaying developments.

### ***Data quality control and metadata***

Data QC is of great concern to the user community. For data types 0 and 1, the waveform data and its direct derivatives this issue is addressed in a number of tools. However, the user seems to require metadata describing the data quality. The NERA project is preparing an overview of existing tools, on-going initiatives and plans within the European seismological RI's. IRIS-DMC launched its MUSTANG project to address this and create metadata (web) services.

Standard global XML descriptions for seismological metadata are being promoted. The seismological community (FDSN) agreed so far on two: station information (stationXML), earthquake information (QuakeML), ... . For data models and its derivatives standardized metadata schemas are being proposed and implemented: like for examples tomographic models, earth models and its associated synthetic seismograms.

ENVRI ([www.envri.eu](http://www.envri.eu)), an ESFRI environmental coordination project, addresses a comprehensive metadata reference model. Within EPOS seismology, we follow these discussions.

### ***Data handling software with significant usage impact***

Within the seismological community there are a few data management packages that may have significant impact on how data will be delivered. For data level 0, 1 these are: SeisComP3, SAC, ObsPy, ... for data level 1 - 3 these are: OpenSHA, ...

## **3. Conclusion and discussion**

A few major research infrastructure projects lead European seismological RI developments: NERIES (2006 - 2010), NERA (2010 - 2014), EUDAT (2011 - 2015?), VERCE (2011 - 2015), ... . A number of research projects like SHARE (2009 - 2013), REAKT (2010 - 2014), MATRIX (?), ..., QUEST (2010 - 2013), ... are relevant, primarily in defining the user requirements, but also data products.

A comprehensive overview of national projects is difficult to compile, but it is clear that many countries are consolidating their monitoring, archival, and computational resources. This implies that different national solutions, more or less coordinated on an international level, are being implemented.

The European community faces a challenge in keeping the different projects coordinated and taking advantage of each other developments and mistakes. The awareness and willingness to coordinate, also with expert IT projects has in later years improved significantly. Earth science and IT collaborative efforts are in the US coordinated within the EarthCube framework created through NSF. This different approach and the different time lines poses significant challenges in synchronizing European and US seismological RI and services developments. Within this aspect COOPEUS can play an important role.

### **References:**

- Lay, T., ed. 2009. Seismological Grand Challenges in Understanding Earth's Dynamic Systems. Report to the National Science Foundation, IRIS Consortium, 76 pp.
- Bell, G. T. Hey and A Szalay, 2010. Beyond the data deluge. Science 10.1126/science.1170411
- Wood, J. et al, 2010. [Riding the wave: How Europe can gain from the rising tide of scientific data](#); Final report of the High Level Expert Group on Scientific Data. October 2010, 40 pp.
- NRC. Improved Seismic Monitoring: improved decision making: assessing the value of reduced uncertainty. National Academic Press, 2006, 183 pp
- Crawford, W., F. Tilmann, A.M. Brisbourne and CHEOPS, 2012. Integrating Mobile European Plate Observing Systems: seismology. 5 pp
- Brisbourne, A. and J. Clinton, 2012. Strategy document on the Euro Array concept. NERA report D2.2, 12 pp.