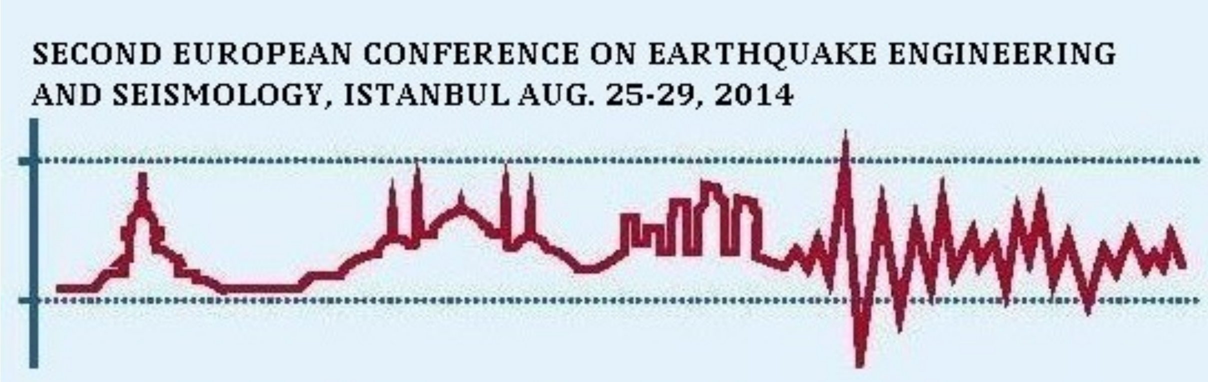

EPOS E-INFRASTRUCTURE AND SEISMOLGY

Daniele BAILO[1], Keith G JEFFERY[2] and EPOS WG7

The European Plate Observing System (EPOS) is integrating geo-scientific information concerning solid Earth data in Europe. We are approaching the end of the PP (Preparatory Project) phase and in October 2014 we will finalize the design of the IT architecture by validating and testing the prototype of the new e-infrastructure. The key aspects of EPOS concern the homogeneous access by end-users to services and multidisciplinary data (which are heterogeneous in terms of data-types, formats and involved communities) collected by monitoring infrastructures and experimental facilities as well as access to software and processing/visualization tools.
Following the creation of an inventory of relevant research infrastructures (RIDE) participating to the EPOS integration plan, the size and the pan-European dimension of the new distributed infrastructure became evident. The EPOS e-infrastructure architecture has been systematically developed through collected primary (user) and secondary (interoperation with other systems) requirements and through three distinct design refinement phases (named Strawman, Woodman and Ironman architecture) with the specification – and developed confirmatory prototypes – becoming more precise and progressively moving from paper to implemented system.
The EPOS architecture is based on Integrated Core Services (ICS), which provide access to thematic (domain-specific) services (called Thematic Core Services - TCS) integrating national research infrastructures. The key aspect is the new metadata catalogue. In one dimension it is described in 3 levels, Jeffery et al (2013):
(1) *discovery metadata* using well-known and commonly used standards such as DC (Dublin Core) to enable users (via an intelligent user interface) to search for objects within the EPOS environment relevant to their needs;
(2) *contextual metadata* providing the context of the object described in the catalog to enable a user or the system to determine the relevance of the discovered object(s) to their requirement – the context includes projects, funding, organisations involved, persons involved, related publications, facilities, equipment and others, and utilises CERIF (Common European Research Information Format) standard (see www.eurocris.org);
(3) *detailed metadata* which is specific to a domain or to a particular object and includes the schema describing the object to processing software.
The other dimension of the metadata concerns the described objects. These are classified into users, services (including software), data and resources (computing, data storage, instruments and scientific equipment).
Alternative architectures have been considered, as for instance the GEOSS driven brokering architecture (Nativi et. al, 2013), widely used in Earth observations. This technique involves writing software to interconvert between any two-nodes datasets. Given *n* nodes this implies writing $n*(n-1)$ convertors.
EPOS Working Group 7 (e-infrastructures and virtual community), which deals with the design and

[1] INGV- Istituto Nazionale Geofisica e Vulcanologia, via di Vigna Murata 605, 00143, Rome, Italy (daniele.bailo@ingv.it)
[2] Keith Jeffery Consultant, Shrivenham, United Kingdom (keith.jeffery@keithgjefferyconsultants.co.uk)

implementation of the EPOS prototype, chose to use an approach which on the one hands ensure interoperability with brokering and other architectures, and on the other hand guarantees a robust handling of metadata and endows the system with an extreme flexibility and sustainability. It is called the Metadata Catalogue approach. With the use of the catalogue the EPOS system can:

1. interoperate with software, services, users, organisations, facilities, equipment etc. as well as datasets;
2. avoid to write n•(n-1) software convertors and generate as much as possible, through the information contained in the catalogue only *n* convertors. This is a huge saving – especially in maintenance as the datasets (or other node resources) evolve. We are working on (semi-) automation of convertor generation by metadata mapping – this is leading-edge computer science research;
3. make large use of contextual metadata which enable a user or a machine to: (i) improve discovery of resources at nodes, (ii) improve precision and recall in search, (iii) drive the systems for identification, authentication, authorisation, security and privacy recording the relevant attributes of the node resources and of the user, (iv) manage provenance and long-term digital preservation.

The linkage between the Integrated Services, which provide the integration of data and services, with the diverse Thematic Services Nodes is provided by means of a compatibility layer, which includes the aforementioned metadata catalogue. This layer provides 'connectors' to make local data, software and services available through the EPOS Integrated Services layer.

In this work we present the EPOS metadata model and discuss synergies with seismological most used standards, in particular with FDSN standards for web services and data access.

In conclusion, we believe the EPOS e-infrastructure architecture is fit for purpose including long-term sustainability and pan-European access to data and services.

REFERENCES

Jeffery, K., Asserson, A., Houssos, N., & Jörg, B. (2013). A 3-Layer Model for Metadata. *Proc. Int'l Conf. on Dublin Core and Metadata Applications 2013*, 3–5.

Nativi, S., Craglia, M., & Pearlman, J. (2013). Earth science infrastructures interoperability: the brokering approach. … *Topics in Applied Earth* …, *6*(3), 1118–1129.