Implementing data provision and services for solid Earth sciences: the EPOS integrated approach

Massimo Cocco\textsuperscript{1}, Beata Orlecka-Sikora\textsuperscript{2} & EPOS Team\textsuperscript{3}

\textsuperscript{1}Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy
\textsuperscript{2}Institute of Geophysics, Polish Academy of Sciences, Poland
\textsuperscript{3}EPOS Consortium, www.epos-eu.org
Access to data: a global challenge

**KEY MESSAGE TO BRING HOME:**
Induced Seismicity and Anthropogenic Hazard must be included in these integration initiatives both for innovation in science and for a science for society

- More value from reuse of data
- More and more researchers are seeing the value of sharing
- Many countries developing open research data policies
- Central Role of Research Infrastructures

More and more researchers are seeing the value of sharing

- Many countries developing open research data policies
- Central Role of Research Infrastructures
European Plate Observing System: Solid Earth Science

- Different communities involved (24 countries)
- Multidisciplinary contributions
- Community building
- Services to society
- Geo-Hazards
- Geo-Resources
- Environmental hazards (including anthropogenic hazard)

www.epos-eu.org
EPOS ARCHITECTURE

THE GOAL OF THIS PRESENTATION:
Further engage this community for participating in the EPOS integration plan by sharing data and products as well as develop services & tools for scientific users and other stakeholders.

EPOS aims to provide seamless access to pan-European data & services.
EPOS Impact & Exploitation of results

European Plate Observing System | www.epos.eu.org

fundamental SCIENCE

TCS Seismology
TCS Volcano Observations
TCS GNSS Data and Products
TCS Geological Information & Modeling

TCS Satellite Data
TCS Near Fault Observatories
TCS Geomagnetic Observations

TCS Anthropogenic Hazards
TCS Multi-scale Laboratories
TCS GETB for Low Carbon Energy

GEOHAZARDS

Surveillance

Safety

GEORESOURCES
EPOS – European Plate Observing System

EPOS will offer to diverse communities data products, tools, and services for intelligible integrated analyses

Accessible data and new e-infrastructures bring novel cross-fertilization of ideas and lead to innovative research, new discoveries & applications for society.
ANTHROPOGENIC HAZARDS RESEARCH INFRASTRUCTURE INTEGRATION

WG 10 in EPOS PP & WP14 in EPOS IP [B. Orlecka-Sikora, S. Lasocki]

MAIN END-USER

SCIENTIST

To study the response of geosphere to exploitation and exploration of geo-resources

INTERMEDIATE NEEDS

TCS Anthropogenic Hazards

A Collaborative Research Environment

Needs catalogue for the end-user

1) Observations of process
2) Advanced, standard tools for data processing
3) Environment & abilities for performing unconstrained virtual research experiments based on 1) & 2)

Creation & Innovation in research
Episode: a time-correlated collection of geophysical data representing the geophysical process, technological data representing the technological activity, which is the cause of this process and all other relevant geodata describing the environment, in which the technological activity and its result - the geophysical process, takes place.

Key barriers to cross (from ERIS v1, 2014):

**Discovery** – proper terminology

**Open & easy access** – standardized formats, proper metadata

**Understanding** – understandable formats, versioning, language, metadata, uniform documentation of all issues regarding data

**Trust** – complete documentation & explanations
Services: already existing and tested, problem-oriented, with the particular attention devoted to methods analyzing correlations between technology, geophysical response and resulting hazard.

- **Data Integration & Stewardship Service (DISS)** - data tagging, formatting, conversion & homogenization. 
  - Dr. Grzegorz Kwiatek, GFZ, Germany

- **Data Handling & Mining Service (DHMS)** - data access, mining & merging and visualization tools.
  - Jean Robert Grasso, ISTerre, France

- **Services for geo-mechanical model for impact of geo-resource production** - stress-deformation patterns generated by type of geo-resource production; THM reservoir models.
  - Stanislaw Lasocki, IGF PAS, Poland

- **Services for geophysical data analysis** - seismicity-deformation signal processing, sustained and/or emergent signals identification, Moment tensor and point source inversion; elastic properties change over time.
  - Jean Robert Grasso, ISTerre, France

- **Services to rate the interactions between technology operations and the upper crust seismic deformation** - 3D cross correlation between the technology driven stress-strain changes and the local seismic-aseismic deformation through time-space.
  - Stanislaw Lasocki, IGF PAS, Poland

- **Services for quantitative probabilistic assessments of anthropogenic seismic hazard** - statistical properties of anthropogenic seismic series and their dependence on time-varying anthropogenesis; ground motion prediction equations; stationary and time-dependent PSH estimates, related to time-changeable technological factors inducing the seismic process.
  - Stanislaw Lasocki, IGF PAS, Poland

- **Simulator for Multi-hazard/multi-risk assessment in ExploRation/exploitation of GEoResources (MERGER)** - numerical estimate of the occurrence probability of chains of events or processes impacting the environment.
  - Alexander Garcia, AMRA, Italy

- **Services for outreach, dissemination & comunication** - communication strategies and bespoke tools for a comprehensive policy of dissemination inside TCS AH, to the wider scientific and technical community of Researchers and to external stakeholder.
  - Peter Styles, KeU, UK
A functional *e-research environment* with an access to HPC resources, ensuring a researcher unbounded possibility to perform in-silico experiments by providing a virtual laboratory in which he/she will be able to create his/her workspace with own processing streams.

**IT Platform Challenges:**

Data: all relevant sources collected and made available within the platform; data includes results of actions - products.

User space: the way how Data and Actions are organized into research projects by a group of researchers.

The goal is to create an IT platform that:

1. Integrates many types of data of large volumes;
2. Enables performing actions on data organized in user space;
3. Supports a natural way of doing research by taking advantage on novel technologies and make progress beyond traditional "science gateways";
4. Provides enough scalability for hundreds of users operating on 'big data' at the same time.
WP17 GEOENERGY TEST BEDS

Components

• Underground Laboratories and associated surface monitoring systems
• Geothermal, Hydrothermal, Unconventional gas,
• Underground Geological Energy Storage
• Gas, compressed air/water potential energy storage
• Radioactive Waste Storage
• Test laboratories such as tunnels and deep mines
• 4-D Subsurface Monitoring and reservoir imaging
• Sensor and data management (physical, chemical and biological)
• High end computer Modelling (fluids, faults, PT, Geochem bio-geochem)
WP 17 GEOENERGY TEST BEDS

Services

• Brokering service for access to industrial applications
• Models for de-Risking Underground Operations
• Data and Software products for monitoring systems
• New sensor systems (including commercial products)
• Transnational Access to underground laboratories
• Technical Support, Standards and Protocols
Big Data Open Data

Open Data & Services
- Accessibility (scientific use)
- Commercial use
- Dissemination to Society
- Service to society

Implications
- Metrics (use & re-use)
- Public funding
- Education & training
- Ethic Issues
# Conclusions

## EPOS

- Integrates national and transnational research infrastructures for solid Earth science
- Guarantees open access to multidisciplinary Research Infrastructures
- Creates novel e-infrastructure and integrated core services
- Fosters scientific, technological and ICT innovation
- Improves geo-hazard assessment, risk mitigation, and sustainable management of georesources

## Goals

- Seamless access to pan-European data and services
- Cross-disciplinary and transnational research
- A multidisciplinary community of users
- Successfully addressing global Grand Challenges in Earth science
- A safe and prosperous society
EPOS Data, Access, and IPR policy

Guiding principles: – open access
– licensing
– no charges

Balance: Legal risk : Openness : Traceability

Protect EPOS legally
Unrestricted use & access
Trace EPOS use & users

Licensing
IPR
Terms & Conditions
Restrictions

Data & Service Providers
Open Access deposit terms
EPOS
Open Access license
Data & Service Users

Data & Data Products
Level 0, 1, 2, 3

Tools & Software
Open : Restricted : Embargoed

mix and match as required

Anonymous : Registered : Authorized
Users

Categorization
as needed for legal aspects

www.epos-eu.org

European Plate Observing System | FP7 Preparatory Phase Project
Users & Stakeholders

Stakeholders categories

- Data and service providers from the solid Earth science (SES) community
- Scientific user community (including researchers from outside SES)
- Governmental organizations
- Industry
- Other data & service providers
- General Public

Influence & interest

www.epos-eu.org
Ethic Issues

Science Communication
- Dissemination
- Education
- Training
- Information
- Knowledge

Risk Communication
- Awareness
- Preparedness
- Resilience to geo-hazards
- Risk perception
- Decisions
OECD Principles and Guidelines for Access to Research Data from Public Funding

13 principles

A – Openness
• Openness means access on equal terms for the international research community at the lowest possible cost, ....

B – Flexibility,  C – Transparency,  D – Legal conformity,  E – Protection of intellectual property,  
F – Formal responsibility,  G – Professionalism

H – Interoperability
• Technological and semantic interoperability is a key consideration in enabling and promoting international and interdisciplinary access to and use of research data. ...

I – Quality,  J – Security,  K – Efficiency,  L – Accountability

M – Sustainability
• ... taking administrative responsibility for the measures to guarantee permanent access to data that have been determined to require long-term retention.
EPOS KEYWORDS

• **Integration** of the existing national and trans-national RIs
• **Interoperability** of thematic (community) services across several multidisciplinary communities
• **Open access** to a multidisciplinary research infrastructure for promoting cross-disciplinary research
• **Acknowledgment** of the data source
• **Progress in Science** through prompt and continuous availability of high quality data and the means to process and interpret them (e.g., *explore and mine large data volumes, results easily reproducible/replicable*)
• Data infrastructures and novel core services will contribute to **information, dissemination, education** and **training**.
• **Implementation** plans, which require strategic investment in research infrastructures at national and international levels.
• **Societal** contributions, e.g., hazard assessment and risk mitigation
“Geo - energy test beds ..... continued ”

<table>
<thead>
<tr>
<th>Sustainability</th>
</tr>
</thead>
<tbody>
<tr>
<td>National energy programmes research foundation and government agencies</td>
</tr>
<tr>
<td>Products would attract support and possible co-funding from oil and gas companies, utilities and energy and environment consultancies (current lead BGS, CCS, INGV, GFZ, Almeria, Utrecht ...)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ De-Risking Underground Operations</td>
</tr>
<tr>
<td>✓ A catalyst for industry both onshore and offshore to stimulate investment and speed new technology options to commercialization, for example Geothermal, shale gas and UCG, energy storage</td>
</tr>
<tr>
<td>✓ It will thus act as a bridge from ideas to application</td>
</tr>
<tr>
<td>✓ Spin outs in a renewed European energy industry</td>
</tr>
</tbody>
</table>
Geo energy test beds

Earth scientists with the energy companies are developing infrastructures to allow the subsurface to be monitored at time scales that are consistent with our use of the subsurface, these will ...

- Increase efficiency and environmental sustainability
- Act as a catalyst to stimulate investment and speed new technology energy options to commercialisation.
- Ensure confidence amongst policy makers and industrial investors, and most of all public confidence.
- Link with the upstream energy producers and other ESFRI projects such as ECCSEL, EU radioactive waste directive, EU renewable energy storage schemes.
Comments on data sharing in EPOS

- EPOS communities feature very different levels of data organization development & maturity
- Most communities have developed in-house their own data services
- Many communities are already striving for their own data archive and services and they are afraid and in some cases difficult to share their data (e.g., why should I put resources in changing what I am doing if I can barely keep track of the services I am compelled to provide?)
- Many communities think they have already the best services (i.e., they can carry out their own research!) and they do not see why the data should be shared (or better qualified).
- Overall, it is a slow process to introduce new concepts, to adopt the same jargon and users/scientists often not yet ready
- BUT it is a positive maturation process
The EPOS chain: high gain/high-\textit{but manageable} risk

- **Access**
  - Data acquisition
  - Integrated use of data
  - Facilities

- **Understanding**
  - Processing and modelling
  - Data massive applications

- **Discovery**
  - User strategy
  - Stakeholder strategy
  - Training and education

- **Trust**
  - Industry
  - Society