European Plate Observing System (EPOS): A FAIR Research Infrastructure

Keith G. Jeffery\textsuperscript{1}, Daniele Bailo\textsuperscript{2}, Kuvvet Atakan\textsuperscript{3} and Matt Harrison\textsuperscript{4}

\textsuperscript{1} Keith G. Jeffery Consultants, Farringdon, UK
\textsuperscript{2} INGV, Rome, Italy
\textsuperscript{3} GEO, University of Bergen, Norway
\textsuperscript{4} BGS, Keyworth, UK

GEOProcessing-2020 - Valencia, Spain, 21-25 November 2020
A researcher wishes to understand the relationships between a volcanic eruption and its effect on human and animal life, health, water supply, power supply, transport, agriculture....

This requires interoperation across geosciences, environmental sciences, and onward to health sciences, civil sciences, agricultural science.....

But the relevant assets are hard to Find, Access, Interoperate and Re-use (FAIR)

Location, description, rights, formats, language.......
Common problem: Heterogeneous Metadata

- Asset descriptions in many formats, languages
- Same asset described differently multiple times
- Many assets not described adequately (or at all)
- Asset descriptions with different and more-or-less formal syntax
- Asset descriptions with different and more-or-less formal semantics

2 approaches:

- Broker between any pair of asset metadata descriptions means $n(n-1)$ brokers

- A canonical rich metadata format and convert to it from each asset description means $n$ brokers (convertors)
FAIR Principles

Findable:
F1. (meta)data are assigned a **globally unique and eternally persistent identifier**.
F2. data are described with **rich metadata**.
F3. (meta)data are **registered or indexed in a searchable resource**.
F4. metadata **specify** the data identifier.

Accessible:
A1. (meta)data are retrievable by their identifier using a **standardized communications protocol**.
   A1.1 the **protocol** is open, free, and universally implementable.
   A1.2 the **protocol** allows for an authentication and authorization procedure, where necessary.
A2. **metadata are accessible**, even when the data are no longer available.

Interoperable:
I1. (meta)data use a **formal, accessible, shared, and broadly applicable language** for knowledge representation.
I2. (meta)data use **vocabularies that follow FAIR principles**.
I3. (meta)data include **qualified references** to other (meta)data. **metadata are accessible**, even when the data are no longer available.

Re-usable:
R1. meta(data) have a **plurality of accurate and relevant attributes**.
   R1.1. (meta)data are released with a **clear and accessible data usage license**.
   R1.2. (meta)data are associated with their **provenance**.
   R1.3. (meta)data **meet domain-relevant community standards**.
Fair Guiding principles

Findable:
F1. (meta)data are assigned a globally unique and eternally persistent identifier.
F2. data are described with rich metadata.
F3. (meta)data are registered or indexed in a searchable resource.
F4. metadata specify the data identifier.

Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol.
   A1.1 the protocol is open, free, and universally implementable.
   A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
A2. metadata are accessible, even when the data are no longer available.

Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles.
I3. (meta)data include qualified references to other (meta)data. metadata are accessible, even when the data are no longer available.

Re-usable:
R1. (meta)data have a plurality of accurate and relevant attributes.
   R1.1. (meta)data are released with a clear and accessible data usage license.
   R1.2. (meta)data are associated with their provenance.
   R1.3. (meta)data meet domain-relevant community standards.
FAIRness Maturity and Implementation

(Bailo et al., 2020)
Metadata catalogs must have

• **Compliance with FAIR principles**
  • Work on how to achieve FAIRness from GO-FAIR, FAIR’sFAIR ... and domain-specific projects e.g. FAIR4Health, ENVRIFAIR ....
  • Indicators of FAIRness developed by RDA FAIR Data Maturity WG

• **Sufficient information for intended use**
  • Discovery, contextualisation (relevance, quality, permissions), action

• **Formal syntax and declared semantics**
  • for autonomic processing

• **Referential and functional integrity**
  • For reliable processing
FAIR and SERVICES

- **FAIR originally designed for DATA not SERVICES**
  - Problem: Implicit download
    - compare taking out a library book
    - Increasingly impractical (size, network latency)
  - Problem: processing associated with the data
    - Locality of data relative to computing resources, processing software and user
    - Resources required (computing, sensor networks, lab equipment)
    - Legalistics (permissions, security, privacy, liability)

- **FAIR for SERVICES**
  - If FAIR is applied to services which provide the data
    - Services intrinsically offer resources required
    - Services may be moved to the data (lower network latency)
    - Services may reduce the data (data management or analytics) for network transport
  - But...Metadata for services is different from metadata for data
Key Characteristics Required for Implementation

**Identity:** Resolvable Universal Unique Persistent Identifier (RUUPID)

Identity is not the same as address (e.g. URL) but can resolve to a URL

**Rich metadata:** “plurality of attributes”

Sufficient to use the metadata to discover the relevant asset (using the resolvable RUUPID or address in the metadata record).

For automation: formal syntax and declared semantics.
For representativity: referential and functional integrity

**Vocabulary:** formal accessible, shared for knowledge representation

Involved in qualified references (as role)

**Licence:** Need to represent as authorisation assertions to control access

**Provenance:**

Used for contextualisation (relevance, quality) – needs to be integrated in catalog
Related to logs and audit
Related to curation (versions, partitions)
EPOS: European Plate Observing System
Research Infrastructures and e-science for data and observations on geo-hazards and geo-resources
European Tectonic Plate covers a considerable geographical area.

EPOS RI provides monitoring of the European Plate through distributed sensor-networks and laboratories.
What is EPOS?

EPOS is a long-term plan for the integration of research infrastructures for solid Earth Science in Europe

EPOS integrates the existing (and future) advanced European facilities into a single, distributed, sustainable infrastructure taking full advantage of new e-science opportunities.

Several PetaBytes of solid Earth Science data will be available.

Several thousands of users expected to access the infrastructure.
Main elements of the EPOS Architecture, the Integrated Core Services Central Hub (ICS-C) and the Executive and Coordination Office (ECO) belong to the EPOS-ERIC legal subject.
EPOS Integrated Core Services (ICS-C)

- Building new **Integrated Core Services** to foster **innovation** and **excellence** in science
ICS-C Portal

http://ics-c.epos-ip.org/
ICS-C Data Portal

There is now a functional FAIR EPOS ICS-C Portal (ICS) with 281 services representing 186 different types of data and data products.

https://www.ics-c.epos-eu.org/
EPOS ICS-C Architecture

ICS-C Provides:
- Search/find/access
- Select/download
- Workspace with selected data
- Work-flows for advanced visualization and processing in ICS-D
- Microservices oriented
- New services easy to plug
- ICS-D ready
- Debugged
- TCS connectors
EPOS - FAIR implementation (1)

Compliant with A1, I3

Three layer metadata model

- Discovery (DC) and (CKAN, eGMS)
- Contextual (CERIF metadata model)
- Detailed (community specific)

- Map & match only contextualized metadata
- Pointers to detailed metadata
- Export metadata in any standard
EPOS - FAIR implementation (2)

Compliant with F2, F3, F4, R1.3

Rich metadata model

- Supports several concepts
- Superset of many metadata standards
- Referential integrity
- Formal syntax, declared semantics
EPOS - FAIR implementation (3)

Compliant with A1.1

Multi layer approach

Web-API based (standard protocols)
EPOS - FAIR implementation (4)

Compliant with A1.1, A1.2
References for FAIR-EPOS


Thank you for your attention!