Example 4: Using the Reference Model as design reference (EPOS)

Descriptions of the Example

Although it looks similar, this example provides a difference perspective on the usage of the Reference Model to that of Example 3: Using the Reference Model in documentation (EMSO).

The ENVRI Reference Model is characterised by being both an ontology and a model. While Example 3: Using the Reference Model in documentation (EMSO) demonstrates how to make use of its ontological framework in documentation, in this example, we exploit its representation as a model, which enables structural thinking and is more useful in the construction of an infrastructure and the organisation of design activities.

The European Plate Observing System (EPOS) is the European integrated solid earth sciences research infrastructure; a long-term plan to integrate existing national research infrastructures for seismology, volcanology, geodesy and other solid earth sciences. One of EPOS' goals is to provide the technical and legal framework by which to automate discovery and access to datasets and services provided by existing national (and trans-national) research institutions and monitoring networks throughout Europe. Another goal is to provide a standard set of core services by which researchers and other interested parties can interact with the federated infrastructure independently of the any particular data centre or national infrastructure. By providing such a common service interface and federation of resources, EPOS will be able to provide greater access to data recorded by existing and future monitoring networks, laboratory experiments and computational simulations, and foster greater cross-disciplinary research collaborations.

EPOS was included in the European Strategy Forum on Research Infrastructures (ESFRI) Roadmap in December 2008 and is currently in its Preparatory Phase; EPOS is scheduled to enter its Construction and Operational Phase in 2015.

EPOS is an infrastructure that intends to integrate several existing infrastructures which in the past have generally been constructed on a national scale only. There already exist established data centres with established working practices and monitoring networks. The challenge for EPOS is to provide a lightweight service layer that can be placed over these existing established infrastructures whilst disguising the underlying heterogeneity of components; this challenge is at least partially mitigated by the existence of certain protocols and data formats that are already standard in some parts of (for example) seismology, and a general drive within EPOS to further extend standardisation throughout its constituent institutions -- though it is not clear how extensively this level of standardisation will apply to all of the (currently highly disparate) earth sciences covered by EPOS' remit.

It is intended that ENVRI contribute in some way to the design of the EPOS Core Services, whether by the production of useful tools (via ENVRI WP4) or by the application of the ENVRI Reference Model (ENVRI-RM) for infrastructure layout and design (via ENVRI WP3). Focusing on the latter, ENVRI-RM should be able to simplify the design problem by breaking it down into well-defined subsystems of components specified from different complementary viewpoints (principally 💡 **Science Viewpoint**, 💡 **Information Viewpoint**and 💡 **Computational Viewpoint**).

How to Use the Reference Model

Following the guidence of the 💡 **ENVRI Common Subsystem** the EPOS design issues can be broken down as follows:

Acquisition

Data acquisition is performed by EPOS' constituent 'client' infrastructures; existing monitoring networks and laboratories, collected by data centres and presented for discovery and access to the EPOS integration layer. Many of these client systems operate in real-time (for example the continuous data streams produced by seismograph networks), requiring concurrently active data curation facilities (storage, persistent identification and metadata assignment).

Curation

Data is principally curated within existing data centres that publish their datasets according to some agreed protocol. These data centres have their own data collection policies, but EPOS intends to promote the adoption of common metadata in order to ease interoperation, based on a three-level model consisting of discovery metadata (using extended qualified Dublin Core) which is derived from contextual metadata (using CERIF, the Common European Research Information Format), which points to detailed metadata (domain-specific and associated with a particular service or resource). EPOS will also provide a global persistent identification mechanism for continuous data streams and discrete datasets (the latter possibly using the mechanisms produced by the EUDAT project).

Curation, Brokering and Processing

Given global persistent identification and metadata, as well as the use where possible of standard data formats, it is intended that tools be produced to search over and extract specific datasets from different sites based on geospatial (and other) requirements. This along with tools for modelling, processing, data mining and visualisation form the data-oriented integration layer of the EPOS Core Services. These sit atop the 'thematic layer' of the Services, which divide services by domain and forms (for example seismology, volcanology and geodesy as well as satellite data, hazard maps, geomagnetic observatories and rock physics laboratories).

Because EPOS is making a concerted effort to integrate data standards and services, the resultant infrastructure should be less reliant on the brokering model than otherwise expected; the homogenisation of resources means that it will not be so necessary to maintain interfaces between heterogeneous resources required to be interoperable.

EPOS also intends to provide access for researchers to high-performance computation facilities as provided by such infrastructure projects as PRACE.

Use

EPOS intends to provide training facilities to its research demographic; it is as yet unclear if EPOS intends to provide any kind of 'social' aspect to its core services (annotation of datasets, record of individual researchers' interactions with the infrastructure, etc.). It is a goal however of EPOS to promote best practices and reward participation, as well as to increase the visibility of research results produced using EPOS services. This implies that community support will become an increasingly important aspect of the EPOS infrastructure as the basic integration challenge it faces becomes solved.

Summary

Like EPOS, ESFRI Environmental Research Infrastructures are characterised as large-scale distributed complex systems involving numbers of organisations across different European countries. Design and implementations become large collaborative activities subject to change and are evolving, which bring significant challenges. Considering the difficulty of ensuring efficiency and productivity, it is not only what to do but how to do it that is important. We observe no approach is currently in use to assist the organisation of the design activities.

The ENVRI Reference Model captures common requirements of a collection of representative environmental research infrastructures, providing a projection of Europe-wide requirements they have, which in potential can be served as a technology roadmap to position and orchestrate collaborations in design and developments. It provides well-defined subsystems of components specified from different complementary viewpoints (Science, Information and Computation), which can help break down the complexity and simplify the design problems, enabling designers to deliver a practical architecture that leads to concrete implementations. It offers a descriptive framework for specifying uniform distributed systems, allowing designers from different organisations to carry out design activities in parallel.

EPOS/ENVRI modelling

EPOS/ENVRI Modelling.