Planet Earth

is the laboratory of environmental sciences

www.envri.eu

ENVRI

the ENVRI Reference Model

exists to illustrate common characteristics of environmental research infrastructures in order to provide a common language and understanding, promote technology and solution sharing and improve interoperability.

MOTIVATION

Independent development of research infrastructures leads to unnecessary replication of technologies and solutions; the lack of standard definitions makes it difficult to relate experiences in one infrastructure with those of others.

The ENVRI Reference Model (**ENVRI-RM**)[1] uses the **Open Distributed Processing**[2] in order to model the "archetypical" environmental research infrastructure. The use of the ENVRI-RM to illustrate common characteristics of existing (planned) **ESFRI**[3] project infrastructures from a number of different perspectives will provide a common language and understanding, promote technology and solution sharing between infrastructures, and improve interoperability between implemented systems.

OPEN DISTRIBUTED PROCESSING

Open Distributed Processing (ODP), an ISO/IEC standard, provides a means to model a distributed system from the perspective of five different viewpoints:

Enterprise — concerning the "business" (in ENVRI, actually the scientific) viewpoint of roles, communities and policies.

Information — concerning the main information objects passed between actors within a system and their characteristics.

Computational — concerning the main functions and interfaces to be distributed across the system.

Engineering — concerning the actual physical distribution of resources in the implemented system.

Technology — concerning the technologies deployed within the system and the standards to which technologies must adhere.

A core idea is that different aspects of a system's design may be better described from different perspectives, pland certain subtleties of design may only become evident when considering the correspondence between different viewpoints. ODP provides an opportunity to explore these perspectives within a single framework.

COMMON SUBSYSTEMS

Analysis of existing infrastructures identifies a common set of functional requirements which can be distributed across five subsystems, each with their own domain of responsibility, each of which every environmental science research infrastructure must either implement itself or delegate to a client infrastructure:

Data Acquisition — for collecting raw data from registered data sources to be made accessible by the infrastructure.

Data Curation — for storing, managing and ensuring access to all persistent data-sets produced within the infrastructure.

Data Access — for enabling discovery and retrieval of scientic data subject to authorisation.

Data Processing — providing a toolbox of services for performing a variety of data processing tasks.

Community Support — encompassing the interaction between an infrastructure and its users.

As shown in Figure 1, amongst the five subsystems can be identified seven points-of-reference wherein interfaces between subsystems can be implemented. Depending on the distribution of resources in an implemented infrastructure, some of these reference points may not be present in the infrastructure due to the conflation of subsystems. Alternatively, certain infrastructures may choose to delegate all or part of a subsystem to one or more client infrastructures (for example delegating data acquisition and curation to individual national research centres whilst providing a common access interface and data processing facilities). In such a scenario certain reference points take on particular importance as they define the interface between federated resources. The ENVRI-RM will define the interfaces existing on subsystem boundaries and specify the function groups and objects needed to collaborate across such boundaries.

By initially focusing on this minimal model, it then becomes practical to produce a partial specification of the ENVRI-RM which nonetheless reflects the final shape of the ENVRI-RM without the need for signicant refactoring. Further development of the ENVRI-RM will focus on designated priority areas based on feedback from the contributing ESFRI project representatives.
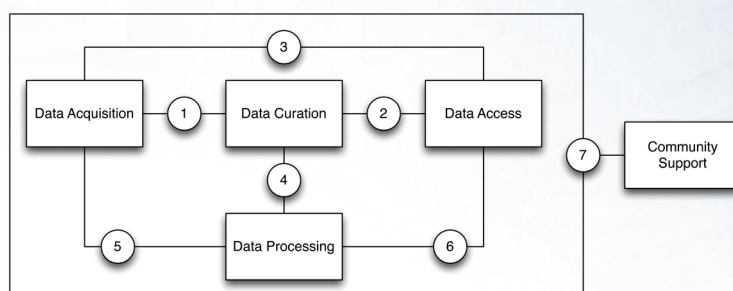

Figure 1

The five subsystems of the ENVRI-RM with the seven reference points (numbered 1-7)

COMMON REQUIRED FUNCTIONALITIES

Analysis of the common requirements of the six ESFRI project infrastructures afiliated with the ENVRI project (being EISCAT-3D, EMSO, EPOS, Euro-Argo, ICOS and LifeWatch) has resulted in the identication of a number of common functionalities.

These functionalities can be partitioned amongst the five subsystems of the ENVRI-RM and presented as interfaces of each subsystem. They encompass a range of concerns, from the fundamental (e.g. data collection and storage, data discovery and access and data security) to more specific challenges (e.g. data versioning, instrument monitoring and interactive visualisation). In order to better manage the range of requirements, and in order to ensure rapid publication of incremental refinements to the ENVRI-RM, a minimal model has been identified (as highlighted in Figure 2), which describes the fundamental functionality necessary to describe a functional environmental research infrastructure.
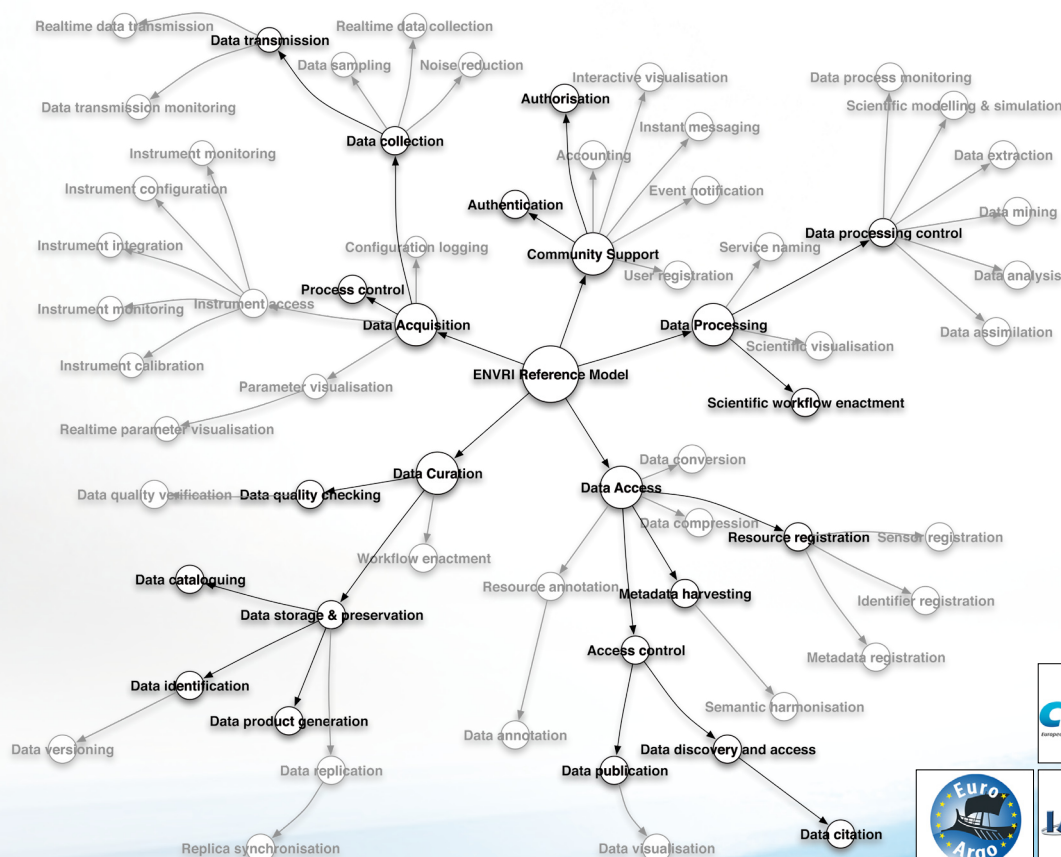
CONCLUSION

The ENVRI-RM will allow comparisons to be drawn and contrasts made between existing and future environmental research infrastructures in order to both determine the extent of each infrastructure within the scope of the model and identify where existing technologies and solutions can fill gaps in infrastructure implementations.

By providing a common conceptual framework of terms used to describe commonalities, better understanding and communication across diverse communities can be achieved.

Wasted effort spent solving old problems and re-implementing existing services may be avoided, and furthermore an additional level of interoperability attained.

Figure 2

A hierarchy of required infrastructure functions assigned to subsystems of the ENVRI-RM. The minimal model is emphasized in the context of the larger requirement set.