

Context and overview (RI Requirements)

Motivation and methods used. Relationship with prior and contemporary work. Structure and contents of the following pages.

Timing

The requirements gathering was undertaken as Task 5.1 of work package 5, of the [ENVRiplus project](#). This began in June 2015 and is scheduled to complete by April 2016. However, we encourage the project members and others to continue refining the understanding, precision and clarity.

Scope and motivation

These requirements cover all aspects of the Information and Communication Technologies (ICT) that are needed to provide the facilities and capabilities required by researchers using the environmental Research Infrastructures (RIs). We particularly focus on requirements for handling *data* throughout its lifetime, i.e. from its initial capture or generation, to its eventual use and when appropriate, archival preservation and curation. We seek to facilitate domain-crossing applications; that is those which involve more than one RI and to find economies from the use of common e-Infrastructure subsystems in multiple RI contexts. The long-term goal is to increase the power and capabilities of the combined research facilities, many of then European Strategic Research Infrastructures (ESFRI), improve their stability and to facilitate the work of each role undertaken by specialist and researchers.

Structuring requirements gathering

The requirements were gathered under a general heading, containing issues and topics that pervade many parts of an e-Infrastructure, then for six specific stages of the data lifecycle, and then for the community support needed to help all users of the e-Infrastructure use it successfully. This leads to the grouping of the requirements under the following headings:

1. **General Requirements:** issues such as the rate, scale and diversity of data gathered, and the range of priority uses.
2. **Identification and Citation:** how to identify data at the required stages of their lifecycle, how to identify groups, subsets and snapshots, in order to cite data accurately.
3. **Curation:** collecting consistent quality controlled collections of data with adequate metadata descriptions, to enable research repeatability.
4. **Cataloguing:** support for rapidly finding relevant data and methods, to accelerate their use, transport and comprehension.
5. **Processing:** model runs and analyses that typically use data as inputs and produce data for further use or visualisation.
6. **Optimisation:** improving any aspect of the data handling and processing, such as reducing energy use, making a routine task easier to perform or reducing the number of steps that require human intervention.
7. **Provenance:** recording the origins of methods and data, so that they may be understood, assessed, and re-visited.
8. **Community support:** helping users start to use and continue to use the facilities, organising the improvement of those facilities based on user feedback, and monitoring progress relative to stakeholders' objectives.

Requirements gathering methodology

The participants undertaking the requirements gathering process in Task 5.1 filled in the following roles:

1. Task Leaders: volunteers from the community of people assigned person months in Theme 2 who focused on a particular topic amongst the ones outlined in the previous section, and shaped the questions about requirements for their topic. In many cases, they will lead subsequent tasks using the information.
2. Research Infrastructure Representative (RI REPs): representatives nominated by the ENVRiplus partner RIs who gathered and provided information about their RIs.
3. Delegates of the RI REPs: colleagues of the RI REPs who replaced them if there was a need for a particular expertise, more detail or effort.
4. Go Betweens: ICT experts, volunteers from the group assigned person months in Task 5.1, who communicated with a small number (typically 2 or 3) RI REPs on behalf of the topic leaders, and also gathered general information. While collecting requirements, they ensured that the ethical procedures, e.g. pseudonymisation, were respected.

The final set of questions agreed on by the Task Leaders is available in ActiveCollab: <https://envriplus.manageprojects.com/projects/requirements/notebooks/431/pages/82/attachments/529/download?force=1&disposition=attachment>.

The Go Betweens conducted a series of discussions with the RI REPs and their delegates, either face-to-face or by using a voice call tool, to incrementally collect their replies to the set of questions, but also more information about other relevant areas for the RI. With the approval of the RI REPs and their delegates, provided in a consent form, the Go Betweens could audio record the discussions apart from taking written notes. After each discussion, the Go Betweens summarised their findings into a report which they shared with the RI REPs or their delegates for further comments, corrections or additions. Final approved reports for each RI were made available by the Go Betweens on the appropriate notebook on ActiveCollab (<https://envriplus.manageprojects.com/projects/requirements/notebooks/470>), and later transferred into a common document on the same application (<https://envriplus.manageprojects.com/projects/requirements/notebooks/431/pages/216/comments/345/attachments/422/download>). This information is now made available through this wiki.

The requirements gathering process of Task 5.1 has received formal ethical approval by the School of Informatics of the University of Edinburgh, the documentation for which is available here: <https://envriplus.manageprojects.com/projects/requirements/notebooks/431/pages/80>

Contributing Research Infrastructures

The list of Research Infrastructures who contributed to the development of these requirements is given below. They illustrate the substantial diversity that needs to be addressed. Some are mature communities with operational production systems; these in part act as exemplars, though they still have emerging issues as the scale and diversity of gathered data increases, as the range of uses of their data diversifies, and as they seek to enable cross-domain research. Some are already establishing a group of inter-working research infrastructures with agreed common standards and research practices - particularly the group of Marine RIs. Some are focused on particular goals or instruments. Others incorporate significant internal diversity.

List of contributing RIs

<Delete those for which e have no information. Link those for which we have requirements to a starter page for that RI, that summarises, and then points to the individual topic pages>

ACTRIS

[AnaEE - Analysis and Experimentation on Ecosystems](#)

[EISCAT-3D - European Incoherent Scatter 3D Radar System](#)

ELIXIR

[EMBRC - European Marine Biological Resource Centre](#)

[EMSO - European Multidisciplinary Seafloor and water column Observatory](#)

[EPOS - European Plate Observing System](#)

[Euro-ARGO - European contribution to the international ARGO programme](#)

EUROFLEETS2

ESONET

EUROGOOS

FIXO3

IAGOS

ICOS

INTERACT

IS-ENES2

JERICO

LTER

SEADATNET

SIOS