 Logo e-Infrastructure



European Data

Grant agreement number: RI-283304

:

|  |  |
| --- | --- |
| Author(s) | Adam Carter (EPCC), Nagham Salman (BSC), |
| Status | Draft |
| Version | v.0.5 |
| Date | 2012-08-13 |

|  |
| --- |
| Abstract**:**  This will be become the training needs analysis deliverable. |

|  |  |
| --- | --- |
| Document identifier: EUDAT-DEL-- | |
| Deliverable lead |  |
| Related work package |  |
| Author(s) |  |
| Contributor(s) |  |
| Due date of deliverable | xx/xx/20xx |
| Actual submission date | xx/xx/20xx |
| Reviewed by |  |
| Approved by |  |
| Dissemination level |  |
| Website | [www.eudat.eu](http://www.eudat.eu) |
| Call | FP7-INFRA-2011-1.2.2 |
| Project number | 283304 |
| Instrument | CP-CSA |
| Start date of project | 01/10/2011 |
| Duration | 36 months |

Table of Content

*Disclaimer:* The content of the document herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While the information contained in the document is believed to be accurate, the author(s) or any other participant in the EUDAT Consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the EUDAT Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the EUDAT Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

[Executive Summary 4](#_Toc311020170)

[1. Introduction 5](#_Toc311020171)

[Annex A. AnnEX 6](#_Toc311020172)

[Annex B. Glossary 7](#_Toc311020173)

list of Figures

No table of figures entries found.

list of Tables

**No table of figures entries found.**

Executive Summary

<Summarize the objectives and key results of the deliverable – 1 page >

# Introduction

<Provide a brief description of the following:

* Objectives of the work related to the project as a whole;
* Purpose of the document;
* Structure of the document (what is in the different sections/chapters); >

<This section will also include a brief description of the approach used to gather the data included herein, with a justification for why this approach was taken.>

# ApproAch to COLLECTING TRAINING NEEDS

The approach used to collect the training needs from the project’s communities was primarily one of direct, iterative communication with representatives from the various communities. This approach was chosen over, for example, a widely-distributed questionnaire for several reasons. Firstly, the communities in EUDAT are particularly diverse, and in some areas, what it means to be a member of a given community is not strictly defined. This leaves us in a different position, for example, from PRACE, who were able to explore training needs in their projects by sending out a questionnaire to all users of PRACE machines, a more uniform and well-defined group. Members of the EUDAT communities include end-user researchers and people involved in providing infrastructure, and they include data producers, data consumers and data processors, and all of the various combinations of these.

Furthermore, the communities are at various stages of maturity. Some are more established than others, and not all communities are in a position to *know* at this stage, what their training requirements might be in detail. By discussing ideas and options with the various communities we aimed to capture their needs, and simultaneously make them aware of issues related to big data that may be an issue for them in the future.

It was therefore decided that the value of results that could be obtained from a mass questionnaire was limited at this stage of the project, as it would be very difficult to obtain quantitative results that were both statistically significant, and sufficiently detailed and wide in scope as to be useful in shaping the project’s training strategy.

EUDAT also benefits from having members of the various communities actively working in the project. We therefore sought to make the most of these connections, and to explore the communities’ needs by talking directly to key members of the communities.

Our approach was to identify a set of initial questions, which were posed to every community, and then to drill down in follow-up interviews and email correspondence with experts within the various communities. The initial questions are included in [Appendix 1?].

# Needs of the Communities

In this section we will describe the training needs of the various communities.

The descriptions are necessarily generalisations, and are based on discussions with members of the communities, as described above. These communities vary in terms of their diversity, and their membership (some for example, consider themselves to be “networks of networks”) and so clearly those training needs identified are not going to be relevant for all members of the community. Indeed, there are experts within the various communities who are well-placed to help deliver or develop the training that will be take place as part of the EUDAT project.

We consider each community in turn, before making comparisons in Section 4.

## CLARIN

CLARIN represents the languages and linguistics community.

CLARIN describes itself as “a pan-European community of institutions emerged from an open and bottom-up driven formation process”. This community comprises *partners* in the EU FP7-funded project of the same name (which concluded in June, 2011), and a consortium of *members*, which are units from institutions throughout Europe. CLARIN is now a legal entity[[1]](#footnote-1).

CLARIN’s goal is to “establish an integrated and interoperable research infrastructure of language resources and its technology. It aims at lifting the current fragmentation, offering a stable, persistent, accessible and extendable infrastructure and therefore enabling eHumanities.”

### Initial Feedback

At present, amongst CLARIN’s members, data is typically scattered on computers and systems belonging to individual researchers and units. CLARIN is not using any large scale data infrastructure at present. The data handled by CLARIN is very diverse.

It was felt that training needed to be provided to both end-users and people who will be teaching. In particular “developers” need training on these matters.

CLARIN’s current systems are considered to be fairly “immature” and there is a need for consulting on what types of service are available and most suitable for them.

CLARIN needs training from very basic things for end-users through to specialised training for their system developers. They need help in defining their overall system landscape, how they should build their architecture.

The need for training is deemed to be pressing, preferably starting this year, and see MPI-PL as a site which would be a good site for training.

The table below indicates those subject areas that are deemed to be the most valuable subjects for training.

|  |  |
| --- | --- |
| Primary | Secondary |
| Data staging & movement | RDBMs |
| Workflows/Dataflows | AAI principles (SAML, certificates, etc.) and Shibboleth |
| Persistent Object Identifiers (Handles, DOIs, URNs, etc.) and APIs |  |
| Policy Rules and iRODS |  |
| Repository Systems (Fedora, D-Space, etc.) |  |
| Metadata Principles and Components (Principles, Standards, OAI-PMH, Aggregation and Indexing, Semantic Mapping, Portals, etc.) (“harvesting”) |  |
| Workflow Orchestration and Execution (Principles, Standards, Frameworks, etc.) |  |

### Follow-up based on initial feedback

## EPOS

The European Plate Observing System (EPOS) is the integrated solid Earth Sciences research infrastructure approved by the European Strategy Forum on Research Infrastructures (ESFRI) and included in the ESFRI Roadmap in December 2008. EPOS is a long-term integration plan of national existing RIs[1].

### Initial Feedback

The needs expressed by EPOS were expressed within the wider context of the EUDAT project. In particular, it was noted that the core services provided by EUDAT will mostly be hidden from the end-user, and EPOS therefore felt that the training (needs) should be targeting those services that allow users to explore the CDI.

EPOS identified several subjects that were considered to be the most valuable in terms of candidates for training, and these are shown in [table].

*Data staging & movement* was considered to be very relevant because, “An EPOS user would like very much to be able to move the data set he/she want to analyze to a HPC center in order to perform analysis otherwise impossible on his/her own desktop or even a small cluster locally”. *Data-intensive computing* is considered important for the same reasons.

*Streaming data* is important in the seismology community in particular due to the growing use of real-time analysis of streams of data which are acquired continuously, or in response to real-world events. The real-time analysis is becoming ever more relevant given that information is now circulating so fast in the population at large, facilitated by social networks.

*Workflows/Dataflows* are considered to be very important to users. Even if the underlying services which would allow such workflows are largely hidden from the user, the user would need to learn to use tools to build and describe their workflows/dataflows.

## ENES

ENES is a group of partners who have agreed “to create a European Network for Earth System Modelling -ENES- with the purpose of working together and cooperating towards the development of a European network for Earth system modelling. These institutions include university departments, research centres, meteorological services, computer centres and industrial partners.”

### Initial Feedback

The data currently being used by the users of ENES, typically comprises lots of (often large) files, with metadata catalogs, in a federated network of servers. The data is stored on hard disk, accessible on-line, with core data backed up between nodes. ENES’s data is very large, and distributed. Some users of ENES use infrastructure provided by the Earth System Grid [ref?]. The volume of data is becoming too large to be processed and analysed with current tools.

It was felt that subject-area meetings would provide the best location for training activities in terms of outreach and acceptance amongst community members.

ENES reported that an approach which “taught the teachers” would be better than one which taught the users, because users in this community typically have a “very light background in technology”.

## LifeWatch

## VPH

The wider VPH community comprises participants in the various projects funded within the EU Seventh Framework programme which address the objective ICT-2007.5.3: Virtual Physiological Human. Within EUDAT, the VPH community is represented by the VPH Network of Excellence.

VPH projects study different physiological systems (in the human body). These systems vary in scale from the cell membrane up to whole organs, thus multi-scaling analysis and simulation is important.

### Initial Feedback

The data in the VPH community are diverse and include imaging data, simulation data, genetic data. A variety of formats are used including, for example, DICOM, JPEG and XLS. At present data is typically stored on hard disk and in private online systems. Important aspects of VPH’s data are the need for security (particularly as some data describes identifiable individuals) and the fact that the data is very large and distributed. At present, there is little use of large-scale data infrastructure.

It was felt that the best place to deliver training would be at existing VPH meetings. Users could be trained, with the expectation that they would then go on to train other users. It was stated that training could consist of courses delivered during the EUDAT project, but that recorded lectures posted on YouTube would also be valuable.

### Questions to be Addressed by a VPH-EUDAT Training Course

The needs of the VPH community are broad, and the details differ from project to project. In this section, we consider the needs of two VPH projects, which were deemed to be representative of the wider community. These projects, p-medicine and VPH share both stand to benefit greatly from EUDAT infrastructure.

VPH indicated that their community would benefit from a community-specific VPH-EUDAT training course, and provided a fairly detailed set of points that they would wish to see addressed in such a course. Their suggestions were as follows:

A VPH-EUDAT training course should be based around a number of practical use cases drawn from p-medicine and VPH share. In VPH there are two types of users: Clinicians who do not care where or how the data is stored but are interested in a GUI that would allow them to search for data sets according to various criteria. The second type is software developers who need to understand the technical interfaces and EUDAT functional and security mechanisms in order to enable their tools to interoperate and use EUDAT functionalities such as uploading, downloading, replication, authenticating and archiving.

The training course should cover the basic aspects of the EUDAT infrastructure functionalities such as storage, archiving and the software tools that you use to upload, replicate, download and manage data.

A list of questions to be addressed is included in Table 1.

|  |
| --- |
| How can end users of EUDAT add/upload patient data to EUDAT resources? Via the web, a desktop application, a dedicated Web Services (or as virtual machine)? |
| How could a user grant access to the shared data set to other EUDAT users from the same VPH community? |
| How could a user recover/distinguish versions of data sets? How could PIDs be used? |
| How could a user replicate data from one EUDAT storage resource to another EUDAT storage resource near an HPC facility? |
| How can VPH users hook up their tools to EUDAT services? |
| How to track patient data sets copies? Who copied a specific data set from EUDAT? Where was it copied to? (Auditing transfer and replication operation is important for these communities because they are bound to contracts on where to store patient data.) |
| What if a VPH user requires extra storage space? Simulation results are usually quiet big in size. |
| How to authenticate to EUDAT infrastructure? What options are in common with the credentials used within VPH community? E.g. EDUROAM, X509 Digital Certificates, and SHIBBOLETH. |
| How to back up data on EUDAT? Archiving? Restoring data from archive? |

Table 1: Questions to be Addressed by a VPH-EUDAT Training Course

The following points should also be addressed:

* VPH would like to replicate data to HPC centres - both EUDAT partner sites and other PRACE centres - to be able to support simulations in a more efficient way. It would be useful to show how replication of large volume of data can be done in a reliable way. In many instances, one does not know beforehand where CPU time will be granted, so quick transport of data between the HPC and Data Centres would be necessary. Non-optimal choices can result in lots of data movement.
* Illustration on how to acquire credential to use EUDAT resources and assign permissions on data sets to share with other users.
* In particular, how to propagate access rights with the replicas within EUDAT, since data centres may have different ACL solutions.

A potential demonstration of data flow was also described, and is included in Table 3

.

Table 2: An illustration of the diversity of data that a VPH user might want to upload to EUDAT

Patient data come in different formats. In particular the following major data types are dealt with in many VPH projects:

* imaging data (fMRI, etc) in the standardized DICOM format
  + DICOM header includes specs for imaging data organization
  + DICOM header includes typical metadata (patientID, doctor, time, etc.)
  + overall data organization is stored in PACS server
* microphotos in the JPEG format
  + limited metadata (patientID, etc) is in the filename
* treatment data in XLS files in defined structure (but not verified)
  + patientID and some other MD is stored in the XLS file
* genetic data
  + all relevant incl. gen snippets in XLS files
  + patientID and some other MD is stored in XLS as well
* simulation model data
  + yet simulations only per patient, i.e. patientID sufficient to indicate relation
  + parameter sets are provided to simulation
  + parameter sets include some molecular dynamic (MD) simulation output is widely textual contained in a small database

The training needs to show how will EUDAT storage be able to store this type of heterogeneous data which require different systems (such as relational database, file server, etc…), are all linked and used as a feed for MD simulations?

Table 3: Example of Possible Demonstration for a VPH-EUDAT Training Course

P-medicine/VPH SHARE cover:

* Relational database which stores all textual and metadata in different tables.
* DICOM metadata is stored in PACS Server and transferred to PACS Server related tables in the database. The imaging data are stored in files.
* The MD from MicroPhotos is stored in the database, while the images are stored as files.
* The treatment data is extracted from spreadsheets and also put into tables.
* The same holds for genetic data.
* Some simulation information is extracted to be part of the IMENSE database, parameters sets and result files are stored in filesystem.

All metadata describing objects is thus included in database tables. Also some data objects are stored as database tables.

One the key issue to demonstrate in a training session is how to preserve metadata and how to replicate relevant data within EUDAT (since databases are dynamic). How EUDAT data centres can store different databases for different communities to deal with metadata?

### Target Audience

Most of the above issues are technical issues relevant to software developers in VPH and scientists. Clinicians tend not to care about the technicalities and usually all the above features are provided behind the scenes. VPH believes that targeting the training towards software developers within VPH can be very useful. They in turn can train clinicians on relevant aspects of data management to an adequate level of detail.

### Format of Courses and Training Materials

Previously, VPH have had courses delivered over a 1 full day (including short talks and lab tutorial) using pre-built use cases on USB sticks (including credentials and tested scenarios) in a group 10-15 people. This has proved to be very helpful and productive because the instructor can closely help the end-users.

EUDAT can aim to link up with HPC training courses that are provided on regular basis (e.g. PRACE, HECToR and DEISA). Since many VPH project involve the usage of both HPC and storage.

## Other Contributing Community 1

## Other Contributing Community 2

# COMMON THEMES & Community Specific Requirements

In this section we will extract common themes from the communities’ needs and deter

## Common Themes

## Community Specific Requirements

# Prioritised List of Planned Training

REFERENCES

1. EPOS Website, online at <http://www.epos-eu.org/>, accessed 2012-07-26
2. *Welcome to the European Network for Earth System Modelling* from the ENES Website, online at <https://verc.enes.org/community/about-enes>, accessed 2012-07-26
3. initIAL QUESTIONS

The initial questions posed to the communities are listed below. As described in [ref:section], these questions were discussed during informal interviews with representatives from the communities, and so not all questions were responded to by every community. They are included here so that they can be referred to in the main text of this document without having to repeat, for example, the list of possible subject areas.

In the questions below, “you” should be interpreted as meaning the community represented by the interviewee.

1. What kind of data?
2. What sort of training?
3. General questions on the perception of needs
4. Subject of the Training
   * What aspects of “data” would you benefit from training in?
     + Data staging & movement
     + Data storage
     + Data-intensive computing
     + Streaming data
     + Workflows/Dataflows
     + Datamining
     + Data labelling / re-use (metadata, etc., formats)
     + Data semantics
     + RDBMSs
     + Data hardware (disks, SSDs, controllers, networks)
     + Distributed data (incl., e.g., distributed joins)
     + AAI principles (SAML, certificates, etc.) and Shibboleth
     + Persistent Object Identifiers (Handles, DOIs, URNs, etc.) and APIs
     + Researcher IDs and ORCID
     + Policy Rules and iRODS
     + Repository Systems (Fedora, D-Space, etc.)
     + Metadata Principles and Components (Principles, Standards, OAI-PMH, Aggregation and Indexing, Semantic Mapping, Portals, etc.)
     + Workflow Orchestration and Execution (Principles, Standards, Frameworks, etc.)
5. Other Comments
6. Glossary

<This glossary should be modified to reflect the content of the document>

|  |  |
| --- | --- |
| AAA | Authentication, Authorisation and Accounting |
| AAI | Authentication and Authorization Infrastructure |
| ADMIRE | **ADMIRE** (Advanced Data Mining and Integration Research for Europe) is a project co-funded by EU within the FP7 |
| APA | Alliance for Permanent Access |
| APARSEN | APARSEN (Alliance for Permanent Access to the Records of Science in Europe Network) is a project co-funded by EU within the FP7 |
| API | Application programming interface |
| ARC | Advanced Resource Connector |
| Aurora Borealis | An ESFRI project in the Environmental Sciences domain. |
| BBMRI | Biobanking and Biomolecular Resources Research Infrastructure. An ESFRI project in the Biological and Medical Sciences domain. |
| BMS | Biological and medical sciences |
| CASPAR | CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval) is an Integrated Project co-funded by the EU within the FP6 |
| CDI | EUDAT Collaborative Data Infrastructure |
| CDO | Climate Data Operators |
| CERT | Computer Emergency Response Team |
| CESSDA | Preparatory phase project for a major upgrade of the Council of European Social Science Data Archives. An ESFRI project in the Social Sciences and Humanities domain. |
| CIT | Community Integration Toolkit |
| CLARIN | Common Language Resources and technology Initiative. An ESFRI project in the Social Sciences and Humanities domain. |
| CMS | Content Management System |
| CODATA | International Council for Science: Committee on Data for Science and Technology |
| COPAL | Heavy Payload Long endurance Tropospheric Aircraft. An ESFRI project in the Environmental Sciences domain. |
| CSMD | Core Scientific Metadata Model |
| CTA | Cherenkov Telescope Array. An ESFRI project in the Physical Sciences and Engineering domain. |
| Curation | Provision of domain-dependent contextual support for permanent access to the meaning of data – including metadata, lexica, etc |
| Curation and Preservation | The process of ensuring that data can be re-used over time. |
| D4Science | Data Infrastructures Ecosystem for Science |
| DAITF | Data Access and Interoperability Task Force |
| DARIAH | Digital Research Infrastructure for the Arts and Humanities. An ESFRI project in the Social Sciences and Humanities domain. |
| DECI | DEISA Extreme Computing Initiative |
| DEISA | Distributed European infrastructure for supercomputing applications |
| DG | Directorate General |
| EB | EUDAT Executive Board |
| EBI | European Bioinformatics Institute |
| EBS | Amazon Elastic Block Store |
| EC | European Commission |
| ECCSEL | European Carbon Dioxide Capture and Storage Laboratory Infrastructure. An ESFRI project in the Energy domain. |
| ECGA | EC Grant Agreement |
| EEF | The European e-Infrastructure Forum |
| E-ELT | European Extremely Large Telescope. An ESFRI project in the Physical Sciences and Engineering domain. |
| EESI | European Exascale and Software Initiative |
| EGEE | Enabling Grids for E-sciencE |
| EGI | European Grid Initiative |
| EIDA | European Integrated Data Archive |
| e-IRG | e-Infrastructure Reflection Group |
| EISCAT\_3D | The next generation European incoherent scatter radar system. An ESFRI project in the Environmental Sciences domain. |
| ELI | Extreme Light Infrastructure. An ESFRI project in the Physical Sciences and Engineering domain. |
| ELIXIR | European Life sciences Infrastructure for Biological Information. An ESFRI project in the Biological and Medical Sciences domain. |
| EMI | European Middleware Initiative |
| EMSO | European Multidisciplinary Seafloor Observatory. An ESFRI project in the Environmental Sciences domain. |
| ENES | European Network for Earth System Modelling |
| EPIC | European Persistent Identifier Consortium |
| EPOS | European Plate Observing System. An ESFRI project in the Environmental Sciences domain. |
| ERA | European Research Area |
| ERIC | European Research Infrastructure Consortium |
| ESA | European Space Agency |
| ESF | European Science Foundation |
| ESFRI | European Strategy Forum on Research Infrastructures |
| ESG | Earth System Grid |
| ESO | European Southern Observatory |
| ESRF | An ESFRI project in the Materials and Analytical Facilities domain. |
| ESS | European Spallation Source Preparatory Project. An ESFRI project in the Materials and Analytical Facilities domain. |
| EUFORIA | EUFORIA (EU Fusion fOR Iter Applications) is a project co-funded by EU within the FP7 |
| EURO-ARGO | Global Ocean Observing Infrastructure. An ESFRI project in the Environmental Sciences domain. |
| EuroHORCS | EUROpean Heads Of Research Councils |
| EURO-VO | EURO-VO (The European Virtual Observatory) project has been funded by the EU under the FP6 and the FP7. |
| FAIR | Facility for Antiproton and Ion Research. An ESFRI project in the Physical Sciences and Engineering domain. |
| FP7 | Seventh Framework Program |
| FTE | Full Time Equivalent |
| GC | EUDAT General Council |
| GÉANT | European multi-gigabit computer network for research and education purposes. |
| GENESI-DR | GENESI-DR (Ground European Network for Earth Science Interoperations — Digital Repositories) was a two-year project co-funded by the EU under the FP7. |
| Geo-Seas | The Geo-Seas project is an Integrated Infrastructure Initiative (I3) of the Research Infrastructures programme within EU FP7. |
| GLOREA | Global European Research Area |
| HELIO | HELIO (Heliophysics Integrated Observatory) was a three-year project co-funded by the EU under the FP7. |
| HEP | High Energy Physics |
| HiPER | High Power Laser Energy Research Facility. An ESFRI project in the Energy domain. |
| HLEG | High level Expert Group |
| HPC | High Performance Computing |
| HPC-Europa2 | Pan-European research infrastructure on high performance computing for 21st century science |
| IAGOS | In Service Aircraft for a Global Observing System. An ESFRI project in the Environmental Sciences domain. |
| ICOS | Integrated Carbon Observation System. An ESFRI project in the Environmental Sciences domain. |
| ICSU | International Council for Science |
| ICT | Information and communication technologies |
| IFMIF | International Fusion Materials Irradiation Facility. An ESFRI project in the Energy domain. |
| ILL | An ESFRI project in the Materials and Analytical Facilities domain. |
| IMPACT | IMPACT (IMproving Protein Annotation through Coordination and Technology) is a three-year EU-funded project |
| IPCC | The Intergovernmental Panel on Climate Change |
| IPR | Intellectual Property Right |
| iRODS | Integrated Rule-Oriented Data System |
| ISO | International Organization for Standardization |
| JHR | Jules Horowitz Reactor. An ESFRI project in the Energy domain. |
| KM3Net | Kilometre Cube Neutrino Telescope. An ESFRI project in the Physical Sciences and Engineering domain. |
| LHC | Large Hadron Collider |
| LifeWatch | E-Science and Technology Infrastructure for Biodiversity Data and Observatories. An ESFRI project in the Environmental Sciences domain. |
| LOFAR | LOw Frequency ARray |
| METAFOR | Common Metadata for Climate Modelling Digital Repositories |
| NEERI | Networking Event for European Research Infrastructures |
| NERIES | NERIES (Network of Research Infrastructures for European Seismology) was a project co-funded by the EU within the FP6. |
| NGS | Next generation sequencing |
| NREN | National Research and Education Network |
| OAIS | Open Archival Information System |
| OASIS | Organization for the Advancement of Structured Information Standards |
| ODE | Opportunities for Data Exchange |
| OGF | Open Grid Forum |
| OGS | Open Grid Service |
| OGSA | Open Grid Services Architecture |
| OpenAIRE | Open Access Infrastructure for Research in Europe |
| OSCT | Operational Security Coordination Team |
| PARADE | Partnership for Advanced Data in Europe |
| PESI | PESI (A Pan-European Species directories Infrastructure) is a three-year project co-funded by the EU under the FP7 |
| PID | Persistent Identifier |
| PLANETS | Planets (Preservation and Long-term Access through NETworked Services) was a project co-funded by the EU within the FP6 |
| PM | EUDAT Project Manager |
| PMO | EUDAT Project Management Office |
| PRACE | Partnership for Advanced Computing in Europe |
| Preservation | Provision of generic support for permanent access to ‘physical’ data – the bits and bytes – including storage, replication, provenance, etc |
| PRINS | Pan-European Research Infrastructure for Nanostructures. An ESFRI project in the Physical Sciences and Engineering domain. |
| QA | Quality assurance |
| QoS | Quality of Service |
| RAD | Rapid Application Development |
| RI | Research Infrastructure |
| SAF | EUDAT Services and Architectural Forum |
| SAML | Security Assertion Markup Language |
| SCAPE | SCAPE (SCAlable Preservation Environments) is a project co-funded by EU within the FP7 |
| SDI | European Spatial Data Infrastructures |
| SEALS | SEALS (Semantic Evaluation at Large Scale) is a project co-funded by the EU within the FP7. |
| SET-Plan | The European Strategic Energy Technology Plan |
| SHAMAN | Sustaining Heritage Access through Multivalent ArchiviNg. A Large Integrated Project co-funded by the EU within the FP7. |
| SHARE | Survey of Health, Ageing and Retirement in Europe project. An ESFRI project in the Social Sciences and Humanities domain. |
| SIAEOS | The Svalbard Integrated Arctic Earth Observing System. An ESFRI project in the Environmental Sciences domain. |
| SKA | The Square Kilometre Array. An ESFRI project in the Physical Sciences and Engineering domain. |
| SLA | Service Level Agreement |
| SME | Small and medium enterprises |
| SNIA | Storage Networking Industry Association |
| SOA | Service Oriented Architecture |
| SPIRAL2 | An ESFRI project in the Physical Sciences and Engineering domain. |
| SRB | Storage Resource Broker |
| SSH | Social Sciences and Humanities |
| STM | Science, Technology and Medicine |
| TERENA | Trans-European Research and Education Networking Association |
| TRAC | Trustworthy Repositories Audit & Certification |
| UNICORE | Uniform Interface to Computing Resources |
| VO | Virtual Organisation |
| VOMS | Virtual Organization Membership Service |
| VPH-I | Virtual Physiological Human Initiative |
| Web 2.0 | A term commonly associated with web applications that facilitate interactive information sharing, interoperability, user-centred design, and collaboration on the World Wide Web. |
| WLCG | Worldwide LHC Computing Grid |
| XFEL | The European X-Ray Laser Project. An ESFRI project in the Materials and Analytical Facilities domain. |
| XML | eXtensible Markup Language |

1. Specifically, it is a European Research Infrastructure Consortium (ERIC). [↑](#footnote-ref-1)