Research e-Infrastructures and e-Infrastructures for Research

Giorgio Rossi, e-IRG workshop, Amsterdam, 9 March 2016
• The Roadmap 2016 identifies 29 Landmarks (2 new) and 21 Projects (6 new)
  – All new entries have been evaluated also for e-needs
  – All the Roadmap Projects are monitored also for e-infrastructure aspects
  – The Landscape Analysis has been carried out for all thematic areas (ENE, ENV, H&F, PSE, SCI) and for the transversal e-INFRASTRUCTURES (provided by e-IRG)

• ESFRI has recently received a specific mandate on optimization of strategy on e-Infrastructure

• The RI system is a test-bed for advanced solutions impacting science and innovation at local, pan-European and global levels
The evaluation process for new proposals followed:

Assessed both scientific merit and project maturity:

1) SWG were integrated by e-IRG experts

2) IG also considered e-needs as elements of organizational and financial plan

3) External referees and hearings

The e-infrastructure aspects, internal to the projects, external, commonalities in the field, interfaces to the e-space were analysed and evaluated as key aspects of the new ESFRI RI Projects, and new ESFRI Landmarks
The ESFRI Projects listed in Part 1 are individually described in the following pages. They were selected for scientific excellence and quality and represent strategic objectives for strengthening the European Research Infrastructure system.

In-frame projects were listed in previous editions of the ESFRI Roadmap – nine in the 2008 update, and six in the 2010 update. Five new entries and one re-mentioned project integrate the Roadmap 2016. They were selected among the 20 eligible proposals through the evaluation procedure outlined in Part 1.

The ESFRI Projects have a maximum term of “radiocity” on the Roadmap of 10 years. After that term the fully implemented projects may become Landmarks. Non-implemented projects leave the Roadmap if desired they can be re-submitted with a revised programme and will compete with other new projects.
21 ESFRI Projects

15 + 6 new

Several are Data Intensive Projects with special e-needs

- ECCSEL
- EISCAT_3D
- EPOS
- SIOS
- EMBRC
- ERINHA
- EU-OPENSSCREEN
- EuBioImaging
- CTA
- EMERIS
- AnaEE
- ISBE
- MIRRI
- ACTRIS
- EMPHASIS
- KM3NeT 2.0
- E-RIHS

Year:
- 2008
- 2010
- 2016
a complex functional architecture has been designed to enable users to discover and select the data, download or visualize them, and perform processing over distributed resources in Europe. The EPOS functional architecture is composed of three connected technical and organizational elements: NRIs, TCS, ICS.

Examples of Data Intensive RI Projects
**ACTRIS**

**Aerosols, Clouds and Trace gases Research Infrastructure**

The ACTRIS project is a coordinated infrastructure designed to high-quality standards of services, on-site, on-ground and oceanic measurements of aerosols, clouds and trace gases and their interactions with the atmosphere. It will deliver precise data, services, and results enabling: the verification of climate models; the physical, optical, and chemical properties of aerosols and clouds; the climate and human health consequences of aerosol interactions with the atmosphere. ACTRIS serves a joint community of users working on climate research, atmospheric modeling, air quality, climate change, and air quality.

**Steps for implementation**

ACTRIS is a new ESFRI project that results from two previous collaborative efforts: the atmospheric aerosol infrastructure community (AARI) and the network of European aerosol research stations (E-AERONET). ACTRIS follows these efforts, integrating them into a single infrastructure that will cover the needs of the atmospheric research community. The project will be managed by the ENEA Foundation, which will ensure the coordination and operation of the infrastructure.

**Website**

http://www.actris.eu

---

**E-EIRHS**

**European Research Infrastructure for Heritage Science**

The EU Framework Programme for Research and Innovation (Horizon Europe) will establish a large-scale Research Infrastructure for Heritage Science (E-EIRHS) in order to support research, innovation, and education in the heritage sciences. The infrastructure will provide access to high-quality resources and services for the study of cultural and natural heritage, enabling researchers, educators, and practitioners to collaborate and innovate on a wide range of topics related to the protection and preservation of cultural and natural heritage. The infrastructure will be developed within the framework of the Ministerial Conference on the Future of Europe, with a focus on the cultural and natural heritage of Europe, and will be operational by 2025.
The ESFRI Landmarks listed in Part 1 are individually described in the following pages. These are former ESFRI Projects that have reached the implementation stage and are now established as major elements of competitiveness of the European Research Area.

Most of the Landmarks were first identified in ESFRI Projects in the Roadmaps 2006 and 2008. Two Landmarks were selected among the 20 eligible proposals through the evaluation procedure outlined in Part 1 recognizing that their implementation is underway.

The ESFRI Landmarks need continuous support for successful completion, operation and upgrade in line with the optimal management and maximum return on investment criteria. Periodic review of the Landmarks will be carried out by ESFRI in order to verify the continuous fulfillment of the reference role in their respective domains.
29 ESFRI Landmarks

Several are leaders in e-infrastructure development

<table>
<thead>
<tr>
<th>ENE</th>
<th>ENV</th>
<th>H&amp;F</th>
<th>PSE</th>
<th>SCI</th>
<th>e-RI</th>
</tr>
</thead>
<tbody>
<tr>
<td>JHR</td>
<td>EMSO</td>
<td>BBMRI ERIC</td>
<td>E-ELT</td>
<td>CESSDA</td>
<td>PRACE</td>
</tr>
<tr>
<td>EURO-ARGO ERIC</td>
<td>EATRIS ERIC</td>
<td>ELI</td>
<td>CLARIN ERIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IAGOS</td>
<td>ECRIN ERIC</td>
<td>European Spallation Source ERIC</td>
<td>DARIAH ERIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICOS ERIC</td>
<td>INFRAFRONTIER</td>
<td>Eu-XFEL</td>
<td>ESS ERIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LifeWatch</td>
<td>INSTRUCT</td>
<td>FAIR</td>
<td>SHARE ERIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELIXIR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **2006**: ESRF-EBS, EMFL, SPIRAL2
- **2008**: ESRF-EBS, EMFL, HL-LHC
- **2016**: ESRF-EBS, EMFL, HL-LHC
Examples of Data Intensive Landmarks

HPC sustainable strategy

PRACE Partnership for Advanced Computing in Europe

One of the highlights of the Partnership for Advanced Computing in Europe (PRACE) is its aim to provide a pan-European infrastructure enabling access to state-of-the-art High-Performance Computing (HPC) facilities to large-scale, addressing user needs in terms of performance, cost-efficiency, and resilience.

Activity

PRACE is an international collaboration between six European countries: Austria, Belgium, France, Germany, Italy, and Spain. It provides access to high-performance computing resources for scientific research in various fields such as physics, chemistry, biology, medicine, engineering, and environmental research.

PRACE-AISBL

The top level of the European High Performance Computing ecosystem

PRACE is a pan-European initiative that aims to support large-scale scientific research by providing access to high-performance computing resources across Europe. It involves a network of national research centers that provide access to supercomputers and high-performance computing services.

HPC sustainable strategy

The concept of sustainable High-Performance Computing (HPC) is increasingly important as the demand for computing power continues to grow. Sustainable HPC involves minimizing the environmental impact of computational tasks by optimizing resource usage, reducing energy consumption, and promoting the efficient use of computational infrastructure.

Examples of Data Intensive Landmarks

The use of data-intensive applications in various fields such as genomics, climate modeling, and personalized medicine has been significant in recent years. These applications require large amounts of data and complex algorithms to analyze and process the information effectively.

The top level of the European High Performance Computing ecosystem

PRACE is a pan-European initiative that aims to support large-scale scientific research by providing access to high-performance computing resources across Europe. It involves a network of national research centers that provide access to supercomputers and high-performance computing services.

TYPE: distributed
COORDINATING ENTITY: PRACE-AISBL
PARTICIPANTS: AT, BE, DE, FR, IT, NL
TIMEFRAME
- Preparation phase: 2010–2011
- Construction phase: 2012–2013
- Operations start: 2014
- Legal status: AISBL, 2010
ESTIMATED COSTS
- Capital value: 300 M€
- Operation: 120 M€/year
HEADCOUNT
- High-level
- Volunteers
INTERNET
http://www.prace-ri.eu/
Examples of Data Intensive RI Projects:

Detector Front End: efficient trigger and DAQ systems

Offline: World Wide LHC Computing Grid

Skills retention. (www.hepsoftwarefoundation.org)

Cost/Sustainability: 100 MCHF/year (in-kind WLCG MoU)

Model: a hybrid model with a mix of public and private (commercially procured) resources. (Test: HNSciCloud EC funded PCP project at the 5% level of today’s WLCG.)

Open access: Data from LHC experiments (year 2010) already made publicly available
Examples of Data Intensive RI Projects:

ESRF EBS will lead to higher yield of thick-data (e.g. CDI images)

Real Time Analysis: The IT infrastructure needs to enable assessing data quality in quasi real-time and first data reduction and analysis synchronously with the experiments

On-line data visualisation: first analysis will allow to store only useful data

Offline: Advanced data analysis services and reduction to transportable formats

Storage: Data will be stored on-site in a High Performing Computing environment providing local and remote access for typically up to one year, and on robotic tape media for longer-term access and archival

Broader community: strong collaboration among all European light and neutron sources

“cloud infrastructure” and “data analysis as a Service” : PaNDaaS

ESRF UPGRADES
Phase II: Extremely Brilliant Source

The European Synchrotron Radiation Facility (ESRF) is the world-leading source of synchrotron X-rays operating 24/7 since 1994, with an annual user number of 5,000, an annual throughput of 3,000 experiments and 17,000 users per year. ESRF has 50 users and 11 research teams, and it is a leader in crystallography, materials science, lanthanide, rare earths, and a key component of the ERA. The ESRF reached its Upgrade Programme in 2008, and has completed the final phase with EUR and structural biology, moving the beamline design and development to high performance with rapid microsyringe and high pressure experiments.

The ESRF is the new planar mirror upgrade project (~400 HRE, 2015-2023), formed on rebuilding the ESRF storage ring by upgrading an all-new, 18-sector, superconducting lattice design, well deliver improved source performance and brilliance (~400 HRE). The ESRF project also included the consideration of four new states of the art beamlines, a scientific laboratory programme with multiple research projects, and a data management and analysis strategy. A data management upgrade is planned for a future mirror upgrade (going through the “Facility programme” expected in the Collaborative Access Projects). Once the high brilliance of the ESRF, methods developed at the ESRF for beamline design, such as virtual prototypes, will be used in the new experimental infrastructure, thus expanding the capabilities for beamline design and scientific research in Europe.

Activity
The ESRF started operations in 1989 and construction was completed in 1990. Today, more than 500 scientific staff access all disciplines of natural science use the ESRF and their work generates ~500,000 publications annually. ESRF has delivered almost ~250,000 instrument-hours (~1,300,000 instrument-hours per week). Approximately 15% of the beams at the ESRF is protected by a large external scientific excellence based access and 26% for prospective researchers. Any non-compliance of all projects submitted to the ESRF Levine Innovation and Technology development, and in the period where beams is available, the science will be planned for a “warm” access with the top project in the “hot” beamline.

A programme of continuous design, upgrade and replacement of facilities has been implemented since the beginning. The ESRF provides scientific support to users and controls the necessary research and development work on the storage ring, including, among others, Nova Phase in December, 2009, 2010 and 2011, the ESRF has created, together with the ESRF and ESRF, a new “Importance of Excellence” that has been used to the users of the ESRF, such as the Institute for Molecular Biology, the University for Molecular Biology and the Molecular Research Institute.

Impact
The new ESRF and the future ESRF impact on scientific research in a partner country, in addition to the new ESRF (2012-2018), the ESRF has six of the global infrastructure for high-end research. Science and research capacities by the ESRF in member and associate states help several Fellows-and-related benefits. The engineering challenges of the ESRF will boost innovation in other areas, such as magnets and detector technology, laser-based applications, control systems, computer science, and data analysis technology. progress in mechanical and high power electrical technology. This will enable ESRF to continue to support new facilities as a driving force in the ERA.
**LifeWatch**

**e-infrastructure for Biodiversity and Ecosystem Research**

**Description**

LifeWatch is an infra-structure for biodiversity and ecosystem research (DARIAH-ERIC). It is a distributed platform advancing biodiversity research and to address the big environmental challenges and support knowledge-based strategic solutions to environmental processes. The vision is achieved by performing access to a network of data sets, services and tools enabling the construction and operation of Virtual Environmental Research Environments.

**Activity**

LifeWatch is an infra-structured distributed e-research, composed by three federated observatories (Observatories Centers, Coordinated Observatories, Observatories Centers) and the Netherlands Centre for Information and Observation Centres. The Observatory Core and the E-infrastructure Technical Platform will work in close cooperation in the development and management of the day-to-day infrastructure technologies, administration, legal and financial issues. These include, among others, infrastructure, management, procurement and RP runtime, and the formal agreements with all the external data and e-services suppliers, and the interim legal agreements (AGS) with local, regional, national and international entities, including decision makers and environmental managers. Also, they will coordinate and manage the E-infrastructure distributed observatories, maintenance and implementation operators, including coordination of the core and management of F-r e-services delivered by the Service Centre, the Virtual Laboratories and Observatories centers and the Inter-nodal E-infrastructure services.

The Service Centre will provide the infrastructure with the LifeWatch Scientific Community, identify the needs of the different areas of science and the need to include distinct levels of services, from the development of the observatories to the service delivery. These will be expected to include the services provided by the LifeWatch Research Infrastructure, including the collection, discovery, visualization, and retrieval of data and applications for services, grid, and creation and delivery of scientific software. The Science Centre will coordinate new data resources, incorporate users, services and data to generate large complex groupings of data. It also plays the role of the information services for the access and consumption of the research infrastructure in the context of the wider European-wide network of Virtual Competency Centers (VCCs). DARIAH-ERIC is discipline, multi-institutional and international and aimed at the provision, use, and exploitation of a LifeWatch Research Infrastructure to facilitate the technology transfer to the Arts and Humanities research, and to the Research’s Information Equalizer. The project aims at research and education work in the digital humanities and the digital equalizer is the key concept to share knowledge and to promote sharing of already digitized content, in digital humanities and social sciences research.

**ESTIMATED COSTS**

- Capital: €41k-M
- Operational: €0-M

**HEADQUARTERS**

- Status: EJ
- Coordinates: BSC-ES-JUL

**WEBSITE**

http://www.lifewatch.eu

---

**DARIAH-ERIC**

**Digital Research Infrastructure for the Arts and Humanities**

**Description**

The Digital Research Infrastructure for the Arts and Humanities (DARIAH) aims to enhance and support digital-culture-related research and teaching across the humanities and arts. DARIAH is a network of people, expertise, information, communication technology, methods, and technologies that variously contribute to the development, manipulation, and operation of information in support of ICT-related research and teaching applications. DARIAH research is driven by the needs of arts and humanities professionals, by following national and international developments, and by the need to make digital research services available to the arts and humanities researchers. DARIAH brings together innovation-driven use of digital humanities and information science, and research in the digital humanities. It promotes, provides access to and disseminates research that stems from these collaborations and ensures that best practices, methodological and technical standards are shared and followed.

DARIAH-EU has established a European Research Infrastructure Consortium (ERIC) in August 2014 to indicate the longer-term sustainability for the European Arts and Humanities research community at the unit level.

**Activity**

DARIAH-EU integrates national node-to-node and national e-infrastructures in the Arts, and humanities, using digital technologies, in an inter-operable way through a European-wide network of Virtual Competency Centers (VCCs). DARIAH is a discipline, multi-institutional and international and aimed at the provision, use, and exploitation of a LifeWatch Research Infrastructure to facilitate the technology transfer to the Arts and Humanities research, and to the Research’s Information Equalizer. The project aims at research and education work in the digital humanities and the digital equalizer is the key concept to share knowledge and to promote sharing of already digitized content, in digital humanities and social sciences research. DARIAH-ERIC further aims to provide a specific work in the area of digital humanities, and to develop digital humanities in a new collaboration with the Open Humanities and Big Data.

**Impact**

DARIAH-ERIC has a significant impact on the inter-disciplinary research environment. The cross-disciplinary nature of the project is supported by a network of virtual research centers, and collaboration between researchers across Europe. DARIAH-ERIC also aims to enhance and support digital culture-related research and teaching across the humanities and arts.

**ESTIMATED COSTS**

- Capital: €41k-M
- Operational: €0-M
LANDSCAPE ANALYSIS

The Landscape Analysis provides the current context, in each domain, of the operational national and international research infrastructures open to European scientists and technology developers through peer-review of competitive science proposals. It represents an impression of the European RI ecosystem. This responds to the invitation by the Competitiveness Council to broaden the view of ESFRI beyond the Roadmap list of projects. It has been produced by the five Strategy Working Groups (SWGs) of ESFRI that are composed of well-recognized scientists and are coordinated by a member of a permanent expert of the ESFRI Forum. The e-infrastructure landscape, transversal to all domains, has been elaborated by the e-Infrastructure Reflection Group (e-IRG). The Landscape Analysis is a key ingredient of the new ESFRI evaluation methodology as it supports the understanding of the impact of new projects. It does not represent in any way the view or prioritization of ESFRI or of any Member State for commitments or future investments. ESFRI in no case acts as an advocate of specific potential future projects. ESFRI and its Member States have taken note of it.

P and L highlighted marks in the text are toggles for direct access to the relevant cards of Part 2.
THE HEALTH CHALLENGE: data collection and stewardship for precision medicine faces the pan-European infrastructures, including ethical, legal and social implications.

ICT for health is indispensable for proper storage and smart retrieval of these data and knowledge. Data for precision medicine are not only generated in the laboratory and clinic.

Citizens and patients are increasingly users of social media and app technologies to share information about their own health and lifestyle: citizen-networking can serve to accelerate clinical research.
THE FOOD CHALLENGE:

The Ris on the Roadmap are the starting point to face the 100% increase in food demand by 2050. The ESFRI Project AnaEE on experimental manipulation of managed and unmanaged ecosystems, EMBRC on marine biological resources, ICOS ERIC on high precision monitoring of greenhouse gas fluxes, MIRRI on microorganisms-oriented services to biotechnology and food production, Euro-BioImaging on integrating imaging technologies and crop phenotyping), ECRIN ERIC on clinical trials and nutritional trials, INSTRUCT on the use of structural biology to support plant and animal sciences, and ELIXIR on life sciences large-scale data and knowledge management (applied to agriculture and bioindustries), must be linked, nd also to other multidisciplinary Ris.
interconnections of Health and Food with other domains.
A section has been contributed by e-IRG addressing to overall landscape and the key issues

- HPC, big data and ultra-fast internet access are enabling technologies for all European economy sectors and society.

- Europe’s competitiveness depends critically on strengthening its HPC technology, Big Data and Applications both in terms of capacity and original capabilities.

ESFRI will play a role in addressing some of the challenges of HPC/Data/Network as RIs of diverse scientific areas and with different kinds of data and use:

- Investments in “e-RI” may need pooling resources (competences, financial)

- HPC and HTC must bridge between science, industry and the public sector.

- Medium term sustainability (10-years) is the common minimum goal for RI data
- HPC infrastructure for research and innovation should extend to all Member States to help the development of strategic sectors;

- Storage infrastructure for research data accessible by diverse thematic software;

- Data infrastructure for data processing, open access, curation, certification and sustainable preservation strategies;

- Fast and Secured transport of data via European data networks.

- Training of Data Practitioners for research, development, industry and services.
ESFRI mandate on e-RI

Working Group on investment strategies in e-infrastructures - ToR

The Competiveness Council of May 28-29 2015 adopted conclusions on: “open, data-intensive and networked research” and “invited ESFRI to explore mechanisms for better coordination of Member States' investment strategies in e-infrastructures, covering also HPC, distributed computing, scientific data and networks.”

The Working Group Mandate