Milestone 8.4: Internal evaluation of the potential use of e-infrastructure portfolio by ACTRIS

Authors: ACTRIS Data Centre

Cathrine Lund Myhre (NILU, lead, & Head of DVAS), Patrice Henry (CNRS Deputy), Markus Fiebig (NILU, Head of In-Situ), Lucia Mona (CNR, Head of ARES), Ewan O'Connor (FMI, Head of CLU), Cathy Boonne (CNRS, Head of GRES), Benedicte Picquet Varrault (CNRS, Head of ASC)

Work package no	WP8
Milestone no.	MS8.4
Lead beneficiary	NILU
Deliverable type	X R (Document, report) ☐ DEC (Websites, patent filings, videos, etc.) ☐ OTHER: please specify .....
Dissemination level	X PU (public) ☐ CO (confidential, only for members of the Consortium, incl. Commission)
Estimated delivery date	M12
Actual delivery date	29/12/2021
Version	Final
Reviewed by	Eija Juurola
Accepted by	Eija Juurola

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115

Comments	

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115

Contents

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115

1 Background and purpose this document

ACTRIS is working together with other environmental RIs (ENVRI) in the framework of ENVRI-FAIR, other RI cluster activities and in the ESFRI framework. As a data-intensive RI, ACTRIS needs to define its links and role with the external e-infrastructure services, especially in the framework of EOSC.

This milestone is reporting and documenting the internal evaluation of the potential use of e-infrastructure portfolio by ACTRIS and evaluate the present level of engagement in service portfolios offered by several e-infrastructures linking to EOSC.

2 List of potential e-infrastructure services for ACTRIS

This section list the relevant and available e-infrastructure services offered, and available. These are numerous and cover many aspects and needs. Relevant links are included for further description of the service.

EUDAT Collaborative Data Infrastructure

- **B2ACCESS**
  Authorisation and authentication proxy for user identification and community-defined access control enforcement.

- **B2DROP**
  Low-barrier storage environment which allows users to synchronise their active data across different desktops and to easily share this data with peers.

- **B2FIND**
  Interdisciplinary discovery portal for research output that allows free term search, results may be narrowed down using several facets, including spatial and temporal search options.

- **B2HANDLE**
  Distributed service for storing, managing and accessing persistent identifiers (PIDs) and essential metadata (PID records) as well as managing PID namespaces. Includes ePIC PIDs.

- **B2SAFE**
  High-availability service which allows community and departmental repositories to implement data management policies on their research data across multiple administrative domains in a trustworthy manner.

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115

- **B2SHARE**
  Data sharing service. for researchers, scientific communities and citizen scientists to store, publish and share research data in a FAIR way.

- **B2STAGE**
  Service to transfer research data sets between EUDAT storage resources and high-performance computing (HPC) workspaces.

2.1 EGI - advanced computing for research

- **Compute: Cloud Compute**
  Run virtual machines on demand with complete control over computing resources

- **Compute: Cloud Container Compute**
  Run Docker containers in a lightweight virtualised environment

- **Compute: High-Throughput Compute**
  Execute thousands of computational tasks to analyse large datasets

- **Compute: Workload Manager**
  Manage computing workloads in an efficient way

- **Storage and Data: Online Storage**
  Store, share and access your files and their metadata on a global scale

- **Storage and Data: Data Transfer**
  Transfer large sets of data from one place to another

- **Storage and Data: DataHub**
  Access key scientific datasets in a scalable way

- **Security: Check-in**
  Login with your own credentials.

- **Applications: Applications on Demand**
  Share online applications for your data and compute-intensive research.

- **Applications: Notebooks**
  Create interactive documents with live code, visualisations and text.

- **Training: FitSM Training**
  Learn how to manage IT services with a pragmatic and lightweight standard.

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115

- **Training: ISO 27001 Training**
  Learn how to manage and secure information assets.

- **Training: Training Infrastructure**
  Dedicated computing and storage for training and education.

3 Planned and intended use of e-infrastructure services in each DC unit

This section presents the planned and used external e-infrastructure for ACTRIS DC, by December 2021.

3.1 In-Situ

The needs of ACTRIS DC In Situ in terms of e-infrastructure services are limited. This reflects In Situ's position of domain knowledge provider at the interface of the scientific data provider community and data management, equipping data with rich and informative metadata as required for data FAIRness. Of the services offered, In Situ will use now or potentially in the future:

- **B2HANDLE / ePIC**
  ACTRIS, and so In Situ, will follow the ENVRI-FAIR recommendation of using ePIC PIDs, e.g. for identification of data pre-products and instruments. EPIC PIDs are in practice identical with the B2HANDLE service.

- **Applications: Notebooks**
  ACTRIS will want to offer data analysis services in Virtual Research Environments, e.g. Notebooks, at some point in the future. These can be hosted by ACTRIS, but might also be hosted by e-infrastructures connected to EOSC.

Other e-infrastructure services will not be used by In Situ because:

- **B2ACCESS**
  ACTRIS will use an authentication solution coordinated with the ENVRI community.

- **B2DROP, B2SAFE, B2SHARE, B2STAGE**
  Data storage and archive capacities aren't a major issue for In Situ since data volume isn't a limiting factor, and can be provided with less overhead by ACTRIS itself.

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115

- **B2FIND**
  ACTRIS will focus on other data discovery portals which are of larger relevance for the atmospheric domain and offer more domain specific services.

- **Other EGI services**
  Since data volume isn't a critical issue for In Situ's services, the same applies to computation resources. These are provided more efficiently and with less overhead by the RI itself.

3.2 ARES

ACTRIS ARES DC unit developed through the years the expertise and technological means at host institutions (CNR-IT and CNRS-FR) required for the scientific data management. Therefore, the need for external e-infrastructure is limited for this unit.

Among the different service listed above, ARES can be interested in using:

- **Applications: Notebooks**

If implemented, such applications are likely to be developed by ACTRIS as a whole, or even across RIs rather than at Unit-level. ARES provision in this sense will probably be dedicated to specific services and features. More in general ARES is also interested in general to the provision of Virtual Research Environments

In the past, the **B2FIND** services was used in the past by ARES for providing an easy access to some of EARLINET datasets. In the future it will be evaluated if this will be still done.

As stated for other unit, the other services will not be used as similar solutions will be coordinated at the ENVRI-level or are more efficiently performed by the unit/RI itself. For PID managment, ARES implemented a local Handle server with direct interfsce with global Handle registry without intermediate steps.

Similarly, ARES unit can benefit of the necessary data storage and computational resources provided by the host institutions. Therefore, EGI services are not required for the ARES unit, being much more efficient to make use of internal resources and expertise.

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115

3.3 CLU

External e-infrastructure needs for the ACTRIS DC CLU unit are also limited as the unit, and host institution, have the required expertise and experience in scientific data delivery. From the services offered, CLU uses

- **B2HANDLE / ePIC**
  The CLU unit follows the ENVRI-FAIR recommendation of using ePIC PIDs for all data objects and uses the B2HANDLE service provided through EUDAT by SURFsara.nl (https://userinfo.surfsara.nl/systems/epic-pid). The CLU unit will also use ePIC PIDs for instruments.

The CLU unit may also participate in creating services that use the following

- **Applications: Notebooks**
  If implemented, such applications are likely to be developed by ACTRIS as a whole, or even across RIs rather than at Unit-level. The CLU unit is also interested in the provision of Virtual Research Environments but any CLU provision will likely be dedicated to specific services such as spectral analysis.

As stated for the In Situ unit, the other services will not be used as similar solutions will be coordinated at the ENVRI-level, or are more efficiently performed by the unit/RI itself.

EGI services are not required for the CLU unit as the necessary data storage and computational resources are provided by the host insitution for CLU, FMI. Using the in-house resources and expertise is much more efficient for CLU since, as noted above, FMI has extensive experience in providing the relevant resources.

3.4 GRES

The GRES DC has limited needs in terms of e-infrastructure services because as the others DC, we have in our structure the required expertise and experience in scientific data management. As regards data we can describe it using rich metadata following fairness principles. However, among the e-infrastructure services proposed, we will use :

- **B2HANDLE / ePIC**
  The GRES unit follows the ENVRI-FAIR recommendation of using ePIC PIDs for all data objects : data from level0 to level 1, and for instruments.

Other e-infrastructure services will not be used by GRES DC. Similar in-house solutions can be used and performed by the GRES unit.

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115

In the same way, the EGI services are not required for GRES unit. Within AERIS/CNRS, our host, we can benefit of data storage and computational resources and also relevant experiences in providing such resources.

3.5 ASC

As GRES DC, ASC DC has limited needs in terms of e-infrastructure services. We have in our structure the required expertise and experience in scientific data management. As regards data we have already described it using rich metadata following fairness principles. However, among the e-infrastructure services proposed, we will use :

- **B2HANDLE / ePIC**
  The ASC unit follows the ENVRI-FAIR recommendation of using ePIC PIDs for all data objects : data from level0 to level 1, and for instruments.

Other e-infrastructure services will not be used by ASC DC. Similar in-house solutions can be used and performed by the ASC unit.

In the same way, the EGI services are not required for ASC unit. AS GRES DC, within AERIS/CNRS, our host, we can benefit of data storage and computational resources and also relevant experiences in providing such resources.

ACTRIS IMP (www.actris.eu) is supported by the European Commission under the Horizon 2020 – Research and Innovation Framework Programme, H2020-INFRADEV-2019-2, Grant Agreement number: 871115