Compilation of Science Project Outputs Information

The EOSC Future project is co-funded by the European Union Horizon Programme call INFRAEOSC-03-2020, Grant Agreement number 101017536

March 2024
Version 1.0

Compilation of Science Project Outputs Information

Lead by LifeWatch ERIC

Dissemination Level of the Document

Public

Abstract

This document contains a compilation of information about the ten EOSC Future Science Projects, including brief descriptions and information about their impact and innovation. Dissemination outputs are also listed for many of the Science Projects.

Version History

Version	Date	Authors/Contributors	Description
V1.0	27/03/2024	Christos Arvanitidis	Compilation of SP individual documents

Copyright Notice

Table of Contents

1 Introduction

Ten Science Projects were conducted in EOSC Future demonstrating cross-disciplinary research collaboration with the support of EOSC. The EOSC Future exploitable results relating to the Science Projects lie in the success of the approach adopted. This success is demonstrated in the many scientific publications, presentations and other documents the Science Projects produced.

This document contains a compilation of information about the Science Projects, including brief descriptions of each, information about their impact and innovation, and a list of dissemination outputs for many of them.

2 NIS-IMPACT Science Project no.1 - ENVRI-FAIR Climate Change

2.1 Description

Name	NIS-IMPACT - Climate Change Impact on Biodiversity and Ecosystems in Europe - Assessing the impact of non-indigenous invasive species (NIS) in European ecosystems
Description	The Objectives of this Science Project (SP) are: (a) to integrate data from different scientific disciplines in the marine subdomain into an analytical framework in order to advance our knowledge on the impact of NIS on European marine biodiversity and ecosystems; (b) to connect the analytical framework and federate access to relevant data infrastructures at the EOSC portal in order to mobilise and empower a larger community of researchers and potential data providers; (c) to demonstrate and promote the benefits and potential of web-based science using EOSC. The SP contributes to the estimation of the impacts of the invasive species on the European Biodiversity and Ecosystems. This topic is important for European Green Deal and the new European Biodiversity Strategy. The SP is also linked with the socio-economic issues because of the NIS implications to the local ecosystems and their services, and their societal goods and services. Since many of the above impacts may be of local scale, they may alter common practices in circular economies. The above are attempted through: (a) combining different sources of data and information; (b) Using a dual workflow to analyse the data; (c) integrating its resources with core EOSC services and potentially horizontal services available; (c) engaging the relevant scientific communities. In SP1, a workflow has been developed to call up datasets from three large global databases, to analyse results from them. The workflow has 2 parallel processes, dealing with sequence data and conventional observational data respectively. The two types of data can now be compared and analysed together thanks to the orchestration provided by the workflow.
URL	https://eoscfuture.eu/data/climate-change-impact-on-biodiversity-and-ecosystems-in-europe/
WPs and Tasks involved	WP6, T6.3 SP1 partners: LifeWatch ERIC, MARIS, SeaDataNet, OGS, NOC-BODC, EMBRC ERIC, HCMR, UGOT, VLIZ, EuroArgo, IFREMER; ENVRI Cluster

Partners	Partners: LifeWatch ERIC, MARIS, SeaDataNet, OGS, NOC-BODC, EMBRC ERIC, HCMR, UGOT, VLIZ, EuroArgo, IFREMER; ENVRI Cluster

2.2 Impact and Innovation

Innovation	This SP has a truly interdisciplinary and cross-domain character and its innovative aspect is that it brings synthetic knowledge on the topic of NIS, by engaging three global databases and datasets from the field (species occurrences, metabarcoding, environmental and socio-economic) , as well as a suite of services for their analysis, never attempted before. It encompasses many research groups scattered in several ERICs and RIs but it's also open to scientists from legal entities which are not members of the EOSC Future Consortium. Based on previous experience, mobilises relevant communities in the context of projects such as ASSEMBLE Plus, core projects of many ERICs (e.g. the Internal Joint Initiative (IJI) of LifeWatch ERIC), etc. With the marine NIS impact assessment SP, EOSC Future aims to: 1) offer a prototype which makes available a suite of web services of multidisciplinary and cross-domain origin, properly orchestrated in the form of a workflow; 2) engage a vast scientific community, once again both multidisciplinary and cross-domain, working on the impacts of major drivers on the marine biodiversity and ecosystems in the EOSC ecosystem. Through ENVRI-FAIR, EOSC Life, and the association with Blue-Cloud and other relevant partners, as well as within the INFRAEOSC03 framework, new services will be added to the EOSC portal, with a proper integration so that it will be possible to be used in combination with both the core services of EOSC (vertically) and to each other (horizontally). Finally, this SP will bring together scientists making basic research on biodiversity and ecosystems, computer engineers, including software and web developers, in order to create a FAIR-compliant virtual research environment (VRE) to achieve both the scientific goals and the community engaged. SP1 has developed a prototype which makes available a suite of web services of multidisciplinary and cross-domain origin, properly orchestrated in the form of a workflow to create a FAIR-compliant virtual research environment (VRE).
Exploitation Plan	The target groups of this project are: ▪ **Researchers/engineers on Invasive Alien Species** (Scientific communities Internal Joint Initiative LifeWatch ERIC+ others): they use the result in order to get new knowledge that can inform environmental managers and

	policy makers to create and take measures to eliminate or mitigate the effects of the invasion; the message they take out of the result is knowledge deriving from different disciplines and domains, that is the synthetic one, and that they have a cloud environment to do their research in collaboration with communities from different disciplines and domains and train their students and associates; the best channels to deliver messages are publications, internet, and social media shared by researchers; they will freely access the result, guided by different sources of information (mostly internet) or colleagues. ▪ **Academicians and students**: they use the result in order to get new knowledge that can inform environmental managers and policy makers to create and take measures to eliminate or mitigate the effects of the invasion and also to use it as subject to teach (be taught); the message they take out of the result is knowledge deriving from different disciplines and domains, that is the synthetic one, and that they have a cloud environment to do their research in collaboration with communities from different disciplines and domains and train their students and associates; the best channels to deliver messages are publications, training and learning channels, internet, and social media shared by researchers; they will freely access the result, guided by different sources of information (mostly internet) or colleagues. ▪ **BON participants/managers**: they use the result in order to get new knowledge that can inform environmental managers and policy makers to create and take measures to eliminate or mitigate the effects of the invasion; the message they take out of the result is knowledge deriving from different disciplines and domains, that is the synthetic one; the best channels to deliver messages are publications, internet, and social media shared by their colleagues; they will freely access the result, guided by different sources of information (mostly internet) or colleagues. ▪ **Environmental Agencies**: they use the result in order to get new knowledge to be informed and take measures to eliminate or mitigate the effects of the invasion; the message they take out of the result is knowledge deriving from different disciplines and domains, that is the synthetic one; the best channels to deliver messages are reports, publications, internet, and social media shared by their colleagues; they will freely access the result, guided by different sources of information (mostly internet) or colleagues. ▪ **Stakeholders** (e.g. management authorities, tour operators, planners, developers, navy): they use the result

	in order to get new knowledge to be informed and to create measures to eliminate or mitigate the effects of the invasion; the message they take out of the result is knowledge deriving from different disciplines and domains, that is the synthetic one; the best channels to deliver messages are reports, publications, internet, and social media shared by their colleagues; they will freely access the result, guided by different sources of information (mostly internet) or colleagues. ▪ **Policy Makers**: they use the result in order to get new knowledge to be informed and to create measures to eliminate or mitigate the effects of the invasion; the message they take out of the result is knowledge deriving from different disciplines and domains, that is the synthetic one; the best channels to deliver messages are reports, publications, internet, and social media shared by their colleagues; they will freely access the result, guided by different sources of information (mostly internet) or colleagues. ▪ **Citizens**: they use the result in order to get new knowledge to be informed and to create activities in order to be actively involved with the environmental issues; the message they take out of the result is knowledge deriving from different disciplines and domains, that is the synthetic one, and that they have a cloud space to collaborate with their colleagues; the best channels to deliver messages are reports, publications, internet, and social media shared by their colleagues; they will freely access the result, guided by different sources of information (mostly internet) or colleagues.
Geographical Market	The geographical scale at which the result has been originally developed is the European seas. However, it can easily scaled up to the global scale.

3 Dashboard on the State of the Environment Science Project no.2 - ENVRI-FAIR

3.1 Description

Name	Dashboard on the State of the Environment
Description	To showcase the benefits of an integration platform that supports scientific workflows, the ENVRIs develop the "Dashboard for the State of the Environment" as a cross-discipline service to address with scientific facts the environmental concerns. The project brings together three scientific domains (Biodiversity, Atmosphere, Ocean) that each have set up analytical workflows to provide environmental indicators in real-time, allowing the users to visualise the "State of the Environment" by interacting with the new service interface.
URL	Dashboard for the state of the environment: https://env-dashboard.eoscfuture.eu
WPs and Tasks involved	WP6 Task 6.2 - Integration of EOSC-Core Services into European Research Practice Task 6.3 - Demonstrating EOSC Value through Cross-domain Research Test Science Projects
Partners	Leader of SP2 is ICOS-ERIC In SP2 a number of organisations, belonging to the ENVRI cluster, are involved. This way, at least six Research Infrastructures (RIs) are involved: SeaDataNet, represented by MARIS, IFREMER, NOC-BODC, and OGS; Euro-Argo, represented by IFREMER; ICOS represented by ICOS; ACTRIS represented by NILU; ELTER represented by CEH; and IAGOS represented by FZJ. Thereby, IFREMER, OGS, and NOC-BODC are involved in EOSC-FUTURE as subcontractors to MARIS, while EMSO, represented by EMSO-ERIC, is involved in EOSC-FUTURE as a LTP to ICOS. All these legal entities are beneficiaries to EOSC-FUTURE and are involved in different roles for the development and delivery of the SP2 results. Several beneficiaries are involved in providing scientific expertise for deriving environmental indicators from the data collections that are managed in the RIs. Other beneficiaries are involved in implementing the technical components underpinning the workflows from base data to indicators and making the indicators available for end-users by means of a dashboard.

3.2 Impact and Innovation

Innovation	Environmental indicators from several RIs can be accessed in one place thanks to the creation of the Dashboard, which allows the users to combine environmental data from different sources to capture the State of the Environment. The Dashboard can serve different user groups, giving easily understandable information to the public but also more advanced services to scientists from relevant fields, academia, industry, and so on. The Dashboard acts as a trustworthy source of information for users because it has been created from EOSC providers, which are long-standing and well-known service providers in the research environment. The Dashboard is scalable, in that more RIs, from different disciplines, can also provide data to it in future.
Sustainability Plan	The Dashboard content is managed by the participating RIs; thus, they bring the responsibility to support the services that bring the environmental indicators to the Dashboard. The Dashboard itself is developed and supported by EMSO ERIC using EOSC core services and EGI infrastructure, thus the sustainability of the service depends on the agreement between EMSO ERIC and EGI, as well as on the sustainability of the EOSC core services.

Exploitation Plan	Scientists (working for policymakers, industry, in research, in education, ...) • access and use of environmental data - e.g. for Green Deal or UN climate goals • a wide range of environmental (currently - and in future possibly wider) data is available through a single interface • ENVRI and EOSC engagement channels • Scientists would find and access the dashboard through the EOSC portal/marketplace and, without logging in, can access data through the dashboard. The dashboard is integrated with EOSC AAI, so users can log in and authenticate, and if they have authorisation they can download environmental indicators from the dashboard, to dig more deeply for information via individual RIs behind the Dashboard.
Geographical Market	European focus but with global application

3.3 Dissemination artefacts

Event title	Event Organiser	IP Owner	Audience	Comments (type of presentation etc)	Material (link)	Date (completed)	License
EOSC Future Workshop	EOSC Future (WP6)	Angeliki Adamaki	EOSC Future partners	Short presentation of the SP, main objectives, integration roadmap	ppt	September 2022	CC-BY
EOSC Future Workshop	EOSC Future	Raul Bardaji & Tjerk Krijger	EOSC Future partners	Short presentation of workflows behind the dashboards and behind the indicators - live demo of the dashboard frontend	ppt	September 2022	CC-BY
ENVRI-Hub Workshop	ENVRI-FAIR	Raul Bardaji	ENVRI-FAIR partners	Mini webinar with live demo of the Dashboard	ppt	October 2022	CC-BY
EOSC Symposium	EOSC	Angeliki Adamaki Tjerk Krijger		Poster	pdf	November 2022	CC-BY
EOSC Symposium	EOSC	Angeliki Adamaki Tjerk Krijger Raul Bardaji	EOSC related audience	Talk during a session where 3 SPs were demonstrated ~20min each	ppt	November 2022	CC-BY
ENVRI week	ENVRI-FAIR	Angeliki Adamaki	ENVRIs community	Invited talk at the session	ppt	February 2023	CC-BY

Event title	Event Organiser	IP Owner	Audience	Comments (type of presentation etc)	Material (link)	Date (completed)	License
				ENVRI-EOSC Future during the annual meeting of the ENVRI cluster			
Invited talk - seminar at INGV Bologna	INGV Bologna	Angeliki Adamaki	INGV staff	Seminar - presentation of ENVRI-Hub and ENVRI Dashboard	ppt	March 2023	CC-BY
EGU23 Session ESSI2.9 here	EGU - ENVRI and friends	Angeliki Adamaki	Geoscience people	Poster	Abstract Poster pdf Poster ppt (only for download-google slides doesn't support it)	April 2023	CC-BY
EGU23	EGU	Angeliki Adamaki	Geoscience people	Booth talk/video/live demo	Video1 How to find and use the Dashboard after logging in (ENVRI-Hub) Video2 How to find and use the Dashboard services without login-also find API swagger (ENVRI-Hub)	April 2023	CC-BY
ENVRI Policy event	ENVRI-FAIR	Angeliki Adamaki	ENVRIs Policy makers	Demonstration	Slides before the demo	June 2023	CC-BY
EOSC Symposium	EOSC	Angeliki Adamaki	EOSC related audience	Talk - EOSC Future Use Cases - ENVRI Dashboard for Policy Makers	ppt	September 2023	CC-BY
Pan-European digital assets supporting	EOSC Future	Tjerk Krijger Dick Schaap		Webinar	ppt	December 5, 2022	CC-BY

Event title	Event Organiser	IP Owner	Audience	Comments (type of presentation etc)	Material (link)	Date (completed)	License
research communities		Peter Thijsse					
EGI2023	EGI	Tjerk Krijger Dick Schaap Peter Thijsse		Talk-Beacon – an innovative data lake system accessible via a REST API designed for fast access to large scale environmental data	Abstract	June 2023	CC-BY
EGI2023	EGI	Tjerk Krijger Dick Schaap Peter Thijsse		Talk - EOSC-FUTURE – ENVRI-FAIR Science Project Environmental Indicators – Ocean workflow	Abstract	June 2023	CC-BY
OSFAIR	OpenAIRE	Tjerk Krijger Dick Schaap Peter Thijsse	Researchers	NAVIGATING DATA LAKES	session	September 2023	CC-BY

Event title	Event Organiser	IP Owner	Audience	Comments (type of presentation etc)	Material (link)	License
ENVRI-Hub Website	Online demo	ENVRIs	Hub users	Landing page with info and links	https://envri-hub.envri.eu/sddetail/85	CC-BY
EOSC Future Website	EOSC Future	SP2 partners	EOSC users	Data in Action-Information page	https://eoscfuture.eu/data/state-of-the-environment/	CC-BY
Science Clusters Website		ENVRIs	Science Clusters		https://science-clusters.eu/use-cases/dashboard-state-environment	CC-BY
EOSC in Practice Stories	EOSC Future	Angeliki Adamaki	EOSC	A collection of EOSC in Practice Stories demonstrating the impact of the EOSC, showcasing its collaborative essence	https://eosc-portal.eu/dashboard-state-environment	CC-BY

				and transformative capabilities across disciplines.		

4 COVID-19 Science Project no.3 EOSC-LIFE Covid-19 Metadata

4.1 Description

Name	COVID-19 - metadata findability and interoperability in EOSC
Description	The 3 main results: • Inventory of metadata schemas applied across infrastructures and domains taking existing ontologies, classifications, controlled vocabularies, workflows, provenance into consideration. • Development of a framework for a metadata model, characterising the research approach and workflow across research infrastructures. • Application of the framework to use cases linked to COVID-19 (ECRIN & BBMRI: interoperability between the ECRIN MDR and the BBMRI Directory), ECRIN & CESSDA: interoperability between the ECRIN MDR and the CESSDA Data Catalogue, CLARIN: ontology of COVID-19 related topics from parliamentary data and social media).
URL	The Science Project does not aim to provide new software or tools. It deals with metadata interoperability aspects within the Life Sciences (LS) cluster and the Social Sciences and Humanities (SSH) cluster and with cross-cluster applications. To limit the scope of the activity COVID-19 has been selected as a use case. Some URLs to project results in the form of publications in Zenodo or in scientific journals: • Protocol of a qualitative study to characterise the contextual metadata and workflows in selected research infrastructures: https://doi.org/10.5281/zenodo.7025319 • Interview guide for a qualitative study to characterise the contextual metadata and workflows in selected research infrastructures: https://doi.org/10.5281/zenodo.7025502 • Mpox Knowledge Graph: a comprehensive representation embedding chemical entities and associated biology of Mpox: https://doi.org/10.1093/bioadv/vbad045 • Improving COVID-19 metadata findability and interoperability in the European Open Science Cloud: https://doi.org/10.3897/arphapreprints.e107873 • Interdisciplinary Approach to FAIR Metadata: Enabling Collaborative Science in the European Open Science Cloud: https://zenodo.org/records/8315477. The project outputs in Zenodo have been disseminated with a Creative Commons Attribution 4.0 International license. Currently awaiting feedback from editors and reviewers on 2 publications in peer review journals. At the current stage since no acceptance information is available we cannot provide URLs or make final statements on IP.
WPs and Tasks involved	WP6 Integration of Community Services and Products into EOSC Task 6.3 Demonstrating EOSC Value through Cross-domain Research Science Projects
Partners	The main result owners are the 6 RIs represented in the Science Project: • ECRIN (European Clinical Research Infrastructure Network) (LS cluster) • BBMRI-ERIC (Biobanking and BioMolecular resources Research Infrastructure – European Research Infrastructure Consortium) (LS cluster)

	• EATRIS (European Infrastructure for Translational Medicine) (LS cluster) • EU-OPENSCREEN (LS cluster) • CESSDA (Consortium of European Social Science Data Archives) (SSH cluster) (represented by the linked third part TAU FSD) • CLARIN (Common Language Resources and Technology Infrastructure) (SSH cluster) (represented by the linked third party University of Antwerp)
Documentation	The main outputs of the SP being scientific documents on metadata interoperability through the prism of COVID-19 research and in the context of EOSC, there is no need for "building", "configuring" or "operating". What will be needed is "extending" by validating the generalisability of the project outputs beyond the SP partners and via use cases other than the COVID-19 one. All available documentation has been provided in the section entitled URL. Documentation in the form of peer review journals will become available but at the current stage 2 paper submissions are awaiting acceptance decisions.

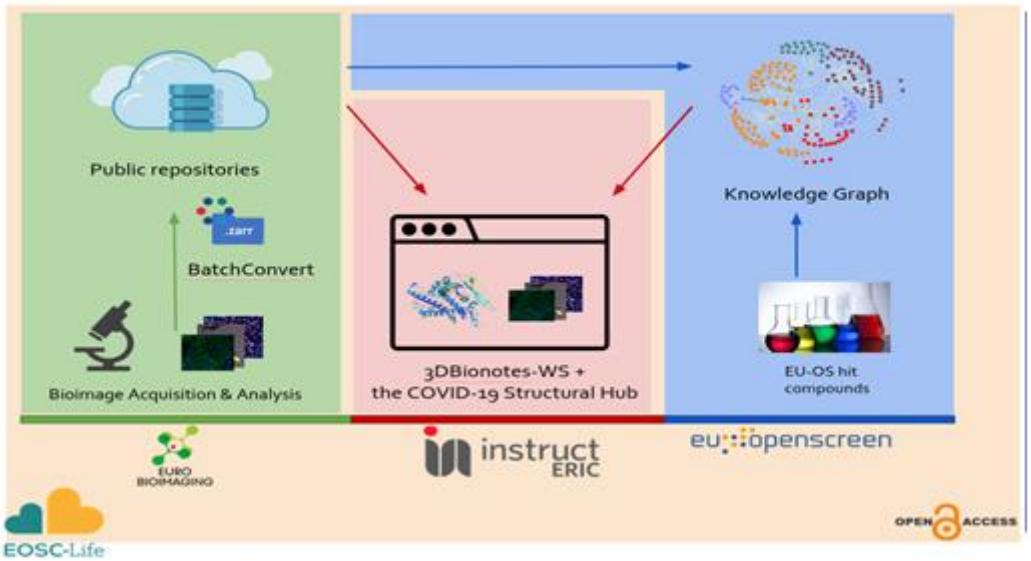
4.2 Impact and Innovation

Innovation	The COVID-19 pandemic has generated a huge variety of research, activities, studies and policies across both the LS and the SSH: examples include genomic sequencing, assays of immune response, clinical trials, population health analyses, exploring vaccine hesitancy, investigating the role of social media, public debate and economic analyses of the impact of public policy issues (e.g., lockdown measures, imposed face masking). Potential insights from combining the data and conclusions from these different forms of research are, however, made more difficult by the lack of a common metadata framework with which to describe them. Even within clusters (e.g., SSHOC, EOSC-Life), the metadata landscape is heterogeneous and numerous domain-specific standards are applied. The situation becomes even more complicated when data sharing is performed across broad disciplinary boundaries, as in this SP, which spans life sciences, social sciences, and humanities. Developing widely applicable metadata is a key part of rendering data more valuable, by allowing them to be more easily found and characterised, regardless of the discipline in which they were generated. In the context of the European Open Science Cloud (EOSC), this improves data reuse within and among scientific clusters. There are certainly metadata schemas (such as DataCite, or the Data Documentation Initiative-DDI) that can describe the concrete outputs of research, e.g., papers and datasets, but relatively little work has been done on finding a metadata schema for the research itself. This is to our knowledge the first time that the science clusters (EOSC-Life, SSHOC) look into cross-domain metadata interoperability between life sciences and social sciences and humanities. A common framework has been developed and discussed among the participating RIs. Next steps would be broadening the discussion to external experts and applying it in the context of EOSC to enable cross-domain use of relevant resources. This project focused only on the COVID-19 use case; additional examples for validation of the general applicability of the framework would be needed.
Sustainability Plan	In this Science Project there was no new development of software/tools. The aim of the project was to investigate and improve the metadata interoperability in cross-domain settings, as the one covered by the EOSC. Pre-existing and sustainable tools developed by each RI were assessed in order to explore metadata interoperability aspects: • MDR (metadata repository) for ECRIN (https://crmdr.org/) • Directory for BBMRI-ERIC (https://directory.bbmri-eric.eu) • MICHA (Minimal Information for Chemosensitivity Assays) for EATRIS (https://micha-protocol.org/)

	• ECBD (European Chemical Biology Database) for EU-OPENSCREEN (https://ecbd.eu/) • Data Catalogue for CESSDA (https://datacatalogue.cessda.eu/) • META-SHARE for CLARIN (http://metashare.elda.org/) The maintenance of these tools is responsibility of the RIs that are operating them. The continuation of progress on cross-domain metadata interoperability within the EOSC landscape is funded through FAIRCORE4EOSC[1], BY-COVID[2], EOSC4Cancer[3] where some of the EOSC Future Science Project partners will continue and expand the results of EOSC Future.
Exploitation Plan	**User groups:** • **Research infrastructures and their communities**: Especially the 6 RIs involved (ECRIN, BBMRI, EATRIS, EU-OPENSCREEN, CESSDA, CLARIN). • **Metadata and ontology experts** • **EOSC projects**: Especially FAIRCORE4EOSC, BY-COVID, EOSC4Cancer but any EOSC project looking into systems interoperable cross-domain. • **EOSC Association Task Forces**: Especially the "Semantic interoperability" task force. Use objective: Further research on metadata, ontologies and cross domain interoperability. Especially in the interdisciplinary context of the EOSC. Messages to deliver: Cross-disciplinary research would accelerate scientific breakthroughs, which are urgently needed in the context of a pandemic (e.g. COVID-19). Developing widely applicable metadata is a key part of rendering data more valuable, by allowing them to be more easily found and characterised, regardless of the discipline in which they were generated. This SP looked into the metadata schemas used across 6 RIs from the LS and the SSH domains and proposed a framework for a cross-domain metadata model. Within both the LS and SSH there is a need for better metadata standardisation and provision. This Science Project looked into the COVID-19 use case to identify similarities and differences with the aim of developing a framework for a metadata model that can describe cross-domain research. Channels used: Scientific publications, posters/presentations in relevant conferences Access: The results of the Science Project are published in FAIR repositories like Zenodo, which assigns DOI and attaches discovery metadata to the uploaded files. For the most innovative parts of the work we have opted for open access publications; journal acceptance decisions and peer reviews are still pending.
Geographical Market	The RIs involved in the project operate on a European level. The results are of interest to metadata and ontology experts in the life sciences and social sciences and humanities on an international level.

5 Life Imaging Data in EOSC Science Project no.4 (EOSC-Life Imaging Data)

5.1 Description

Name	Open Imaging Data in EOSC (Covid-19 as Demonstrator)
Description	Leveraging the COVID-19 pandemic as a driving force, three different partners, INSTRUCT-ERIC, EU-OpenScreen and EuroBioImaging, coming from the same EOSC Cluster "EOSC-Life" took advantage of the connection that EOSC enabled to start working on bringing together different parts of the imaging community and combining their resources to solve some of the issues faced during the pandemic. As a result, within this Science Project three different and totally open-access infrastructures have been developed and interconnected among them for the researchers. One of them is the open-access tool BatchConvert, coming from Euro-BioImaging which is a command line tool that triggers automated, parallelised workflows for high-performance image data transfer from / to remote storage and conversion into standardized, cloud-optimized file formats. Therefore, the tool allows for a quick and efficient submission of bioimage data to public repositories. Also, EU-OpenScreen has been working on collecting data from fragment-based drug discovery studies and bioimaging assays from public repositories on SARS-CoV-2 to create an open-access COVID-19 Knowledge Graph and MonkeyPox Knowledge Graph. These Knowledge Graphs are mainly a representation of compounds and biological entities to facilitate the understanding of the molecular basis of the diseases and the landscape of available compounds to treat them. Finally, INSTRUCT-ERIC has concentrated its efforts on multidimensional image data in the field of structural biology by extending 3DBionotes-WS and the COVID-19 Structural Hub, two open-access web platforms that facilitate user-friendly access to protein models and their annotations. More precisely, information coming from bioimaging public repositories along with quality-related data have been integrated into the interfaces of 3DBionotes-WS and the COVID-19 Structural Hub  *Figure 1: The scientific case in a nutshell for Open Imaging Data Sharing in EOSC*
URL	3DBionotes-WS • Official Webpage URL: https://3dbionotes.cnb.csic.es/ws

	• EOSC Marketplace URL: https://marketplace.eosc-portal.eu/services/eosc.cnb-csic.3dbionotes-ws COVID-19 Structural Hub • Official Webpage URL: https://3dbionotes.cnb.csic.es/ws/covid19 • EOSC Marketplace URL: https://marketplace.eosc-portal.eu/services/eosc.cnb-csic.covid-19_structural_hub BatchConvert: • EOSC Marketplace URL: https://search.marketplace.eosc-portal.eu/search/all?q=batchconvert • WorkflowHub URL: https://workflowhub.eu/workflows/453 • GitHub Repository URL: https://github.com/Euro-BioImaging/BatchConvert • Anaconda URL: https://anaconda.org/Euro-BioImaging/batchconvert KG Workflows: • Mpox KG GitHub URL: https://github.com/Fraunhofer-ITMP/mpox-kg • Mpox KG EOSC URL: https://explore.eosc-portal.eu/search/software?pid=10.5281%2Fzenodo.7351393 • COVID19 KG GitHub URL: https://github.com/Fraunhofer-ITMP/BY-COVID-KG • COVID19 KG EOSC URL: https://explore.eosc-portal.eu/search/software?pid=10.5281%2Fzenodo.7351221 • COVID19 NMR KG GitHub URL: https://github.com/Fraunhofer-ITMP/COVID-NMR-KG
WPs and Tasks involved	WP6 - Integration of Community Services and Products into EOSC T6.3 - Demonstrating EOSC Value through Cross-domain Research Science Projects
Partners	Spanish National Center for Biotechnology (CNB-CSIC), Instruct-ERIC EU-OPENSCREEN Euro-BioImaging

5.2 Impact and Innovation

Innovation	Users are now able to analyse and interact with imaging data, structural data, and quality-related data through the user-friendly interfaces of 3DBionotes-WS and the COVID-19 Structural Hub. These platforms provide a unique graphical environment, enhancing users' understanding of multidimensional image data in the context of structural biology. Users can now convert batches of bioimage data from diverse proprietary file formats into either of the standard formats, OME-TIFF or OME-Zarr, in a parallelised fashion. Remote input/output options are supported, enabling the user to directly deposit the converted data in an s3 bucket or a BioStudies user space (from where the data can be fetched by the BioImage Archive team for publishing). Execution on Slurm clusters is also supported. Similarly, the reproducibility of KG workflow enables quick generation of KGs that represent important biological and chemical entities, assays, biological processes and pathways of proteins. Also, the resources are harmonized to each other such that they can be accessed easily within each other's resources.
Sustainability Plan	• having clear ownership (legal or IP rights over the result) – Euro-BioImaging has the legal rights over the 100 % of the BatchConvert package. – CNB-CSIC/INSTRUCT-ERIC has the legal rights over the 100 % of 3DBionotes-WS and COVDI-19 Structural Hub. – EU-OpenScreen has the legal rights over the 100% of KG resources developed for this SP. • access (availability and accessibility of the result to relevant stakeholders);

	– BatchConvert is an open-source software package, which is hosted at and can be freely installed from: https://github.com/Euro-BioImaging/BatchConvert. BatchConvert is also available as an Anaconda package, which is hosted at: https://anaconda.org/Euro-BioImaging/batchconvert. – 3DBionotes-WS and the COVID-19 Structural Hub are open-access web platforms, designed to foster easy access to a wealth of information within the field of structural biology. Users can readily access these platforms without the necessity to log in. Interested individuals can access them through their respective official websites (3DBionotes-WS: https://3dbionotes.cnb.csic.es/ws, COVID-19 Structural Hub: https://3dbionotes.cnb.csic.es/ws/covid19) or via the following links available on the EOSC Marketplace (3DBionotes-WS: https://marketplace.eosc-portal.eu/services/eosc.cnb-csic.3dbionotes-ws, COVID-19 Structural Hub: https://marketplace.eosc-portal.eu/services/eosc.cnb-csic.covid-19_structural_hub) – KG resources (BY-COVID, Mpox and COVID-NMR) are open source workflows developed for representing chemotype-phenotype of diseases with FAIR data principles. Users can reuse the source codes directly from github. ▪ Mpox KG GitHub URL: https://github.com/Fraunhofer-ITMP/mpox-kg ▪ COVID19 KG GitHub URL: https://github.com/Fraunhofer-ITMP/BY-COVID-KG ▪ COVID19 NMR KG GitHub URL: https://github.com/Fraunhofer-ITMP/COVID-NMR-KG • maintenance and support (activities preserving the functionality, reliability, usability of the result); – Maintenance and further development of BatchConvert will be fully conducted by Euro-BioImaging. Support will be provided in terms of bug-fixing, feature update and new feature development, depending on the feedback from users. – Maintenance of 3dbionotes-WS and the COVID-19 Structural Hub will be conducted by CNB-CSIC/INSTRUCT-ERIC. New updates will be based on internal plans and user feedback, through different channels, all managed internally by EOSC Helpdesk. New features on both platforms will be periodically shared in journals and webinars to engage community. A detailed help page has been created to facilitate the effective and efficient use of the tools. – The KG resources will undergo new updates and improvements and therefore will continue to grow and evolve. • costs (sustainable funding models to cover expenses); – Costs of maintenance of BatchConvert will be settled internally by Euro-BioImaging. – To ensure the sustained operation and development of 3DBionotes-WS and the COVID-19 Structural Hub, CNB-CSIC/INSTRUCT-ERIC plan to pursue continued grant funding from relevant European entities aligned with our project's goals, allowing us to cover necessary expenses while maintaining free access to our tools. – The cost to cover expenses for further improvement and development of KG resources will be done by seeking fundings from relevant EU projects. Eg. BY-COVID. • expected returns (benefits or value that stakeholders anticipate deriving from the result). – We envision that BatchConvert will be especially useful for the Euro-BioImaging nodes, which include research facilities and centres focusing on bioimaging. Potential users of the software include academic researchers and data scientists working with multi-dimensional image data. The tool enables high-throughput conversion of proprietary image file formats into the standard open formats, which are recognised by most image visualisation and analysis software. This serves to unify and harmonise bioimaging data, facilitate data analysis and

	contribute to the implementation of the FAIR principles in general. The tool is easy to install and designed to support execution on local or remote machines as well as on HPC clusters, thus allowing the user to flexibly execute medium-to-large-scale jobs. The tool also supports remote input-output options, automatically handling the data transfer between the execution environment and remote storage such as private s3 buckets and BioImage Archive. These features can be exploited by facilities with access to such compute centres and/or remote storage. – Our main goal of developing the KG resource is to allow users to generate disease-specific KGs with a minimal input of disease name/id. As compared to existing workflows, the final version will be an automated workflow which is user-friendly and faster than previous workflows.
Exploitation Plan	BatchConvert addresses mostly academic research facilities and centres who conduct bioimaging research. The tool will, in particular, serve to increase the accessibility and interoperability of bioimage data and facilitate data visualisation and analysis. Therefore, the tool is suitable to be used for further research. BatchConvert also supports standardisation activities in the bioimaging domains, through its function of high-throughput conversion of proprietary file formats to standard bioimaging file formats. The KG resource (when fully developed) can be used in any other disease (beyond infectious diseases) because of its generic nature. This will be launched as KGG (Knowledge Graph Generator) tool and will come with a user friendly/interactive interface. The entities in KG will be represented with ontologies and controlled vocabularies to align with FAIR data standards, hence will enable standardisation activities. Furthermore, the KG can be exported to many standard file formats for visualisation and analysis. 3DBionotes-WS, by nature, is an integrative set of tools whose main goal is to bring together many different data sources and make them available to its users in the most intuitive way possible, through its 3D interface. Moreover, we are increasing the API in order to make easier for third party tools to get access to all data collected in 3DBionotes-WS and the COVID-19 Structural Hub. Researchers, Citizen scientists • Objective: provide a landscape of available compounds/assays to find ways to treat COVID19 • Main messages: a new open source application has been developed, providing a single place to access and interact with multiple sources of Covid-19 data. • Channels: EOSC website, portal and marketplace • How to access and use: the Open Imaging Data Infrastructure is available in the EOSC Marketplace
Geographical Market	European focus but with global application

6 Dark Matter (Science Project no.5)

6.1 Description

Name	Dark Matter
Description	ESCAPE has several different experiments studying dark matter in various ways. In this Science Project, ESCAPE datasets from the various experiments are brought together in the data infrastructure built in ESCAPE. Making use of services discovered through the EOSC Exchange, analysis is performed on the datasets to produce new insights on the existence of dark matter. This is the first time this has been done in such a reproducible and coherent way. In SP5, new workflows have been developed to support the end-to-end process and new algorithms have been developed to perform the combined analysis.
URL	https://escape2020.pages.in2p3.fr/virtual-environment/home/sp-dark_matter/
WPs and Tasks involved	WP6, T6.3 Also worked with T6.1 and T6.2 Partners in SP5: CERN, CNRS, INFN, FAU
Partners	CERN, CNRS, INFN, FAU

6.2 Impact and Innovation

Innovation	Ability to combine different experiment outputs together for the first time, making use of services discovered through EOSC Exchange, to allow potential new insights into Dark Matter. The main results are innovative algorithms, workflows, and scientific results.
Sustainability Plan	The main results are innovative algorithms, workflows, and scientific results. These will be maintained by the scientific community as long as there is interest in continuing the analysis. ESCAPE is now an Open Collaboration and creates the framework within which to maintain the results as a collaborative activity between the RI partners of the collaboration. All of the software, data, workflows are open access and licensed, and are being published into the EOSC Exchange and marketplace via the ESCAPE OSSR and data portals. They will all be openly available and findable, and usable according to the various licensing conditions. The actual scientific results will be published as open access papers and made findable through the marketplace.
Exploitation Plan	The main user groups for the results of this Science Project are primarily other scientists and citizen scientists. The main uses and exploitation of the results are: • Use for further research, validation of results, access to new algorithms and software; • We have demonstrated the ability to use EOSC to perform open science, and to make all the resources available to non-ESCAPE users who can repeat or adapt the analyses; • The main channels to deliver messages are: domain conferences, EOSC Symposia, web (EOSC-Future, ESCAPE, and Science Clusters), news items; • All the software, data, workflows are open access and licensed, and are being published into the EOSC Exchange and marketplace via the ESCAPE OSSR and data portals. They will all be openly available and findable, and usable according to the various licensing conditions. The actual scientific results will be published as open access papers, and also made findable through the marketplace.
Geographical Market	Global

6.3 Dissemination Artefacts

Name	Description	URL	IP owners	Licence used
Web Site	Dark Matter Science Project Summary	https://projectescape.eu/dark-matter-test-science-project	ESCAPE	CC-BY
Web Site	Dark Matter Project Description	https://escape2020.pages.in2p3.fr/virtual-environment/home/sp-dark_matter/	ESCAPE	CC-BY
Web Site	EOSC-Future Data in Action: Dark Matter	https://eoscfuture.eu/data/dark-matter/	EOSC-Future	CC-BY
Web SIte	EOSC Portal - EOSC in Practice Stories: C. Doglioni, L. Heinrich, F. Calore, The Dark Matter Test Science Project	https://eosc-portal.eu/dark-matter-test-science-project-tsp	CERN, CNRS	CC-BY
Web Site	Science Clusters Web Site - Use Cases; Maximizing Collaboration on Dark Matter and Gravitational Wave Research	https://science-clusters.eu/use-cases/collaboration-dark-matter-gravitational-wave-research	Lund University, CNRS, TUM, SNS	CC-BY
Poster at JENAS 2022; May 2022	Jared Little - **The Dark Matter Science Project in European Open Science Cloud (EOSC) Future**	https://indico.cern.ch/event/1040535/contributions/4843827/	CNRS	CC-BY
Talk at FAIR-AI workshop, June 2022	L. Heinrich, C. Doglioni - **FAIR on the European Open Science Cloud: the case of the Dark Matter Science Project in ESCAPE**	https://indico.cern.ch/event/1152431/contributions/4871314/	TUM, UNIMAN/Lund University	CC-BY
Talk at ACAT22, October 2022	Elena Gazzarrini - **The Virtual Research Environment: towards a comprehensive analysis platform**	https://indico.cern.ch/event/1106990/contributions/4991200/	CERN	CC-BY
Presentation material and video recordings at ESCAPE to the Future	Caterina Doglioni, Francesca Calore, Lukas Heinrich, **Dark Matter Science Project Overview**	https://projectescape.eu/events/escape-future	UNIMAN/Lund University , CNRS, TUM	CC-BY

Conference, October 2022				
idem	Jared Little, **ATLAS Results**	https://projectescape.eu/events/escape-future	CNRS	CC-BY
idem	Valerio Ippolito, **Darkside Plans and Results**	https://projectescape.eu/events/escape-future	INFN	CC-BY
idem	Alex Gallen, Alex Ekman, **Machine Learning Tools**	https://projectescape.eu/events/escape-future	Lund University	CC-BY
idem	P. Bhattacharjee, C. Eckner, **FermiLAT Results and CTA Plans**	https://projectescape.eu/events/escape-future	CNRS	CC-BY
idem	Caterina Doglioni, Francesca Calore, Lukas Heinrich, **Results, Consolidation and Open Challenges**	https://projectescape.eu/events/escape-future	Lund University, CNRS, TUM	CC-BY
Science Project Template, November 2022	C. Doglioni, L. Heinrich, F. Calore, I.Bird, **ESCAPE Science Project for EOSC-Future: The Dark Matter Science Project**	https://technopolisltd223.sharepoint.com/:w:/r/sites/INFRAEOSC-03Proposal/_layouts/15/Doc.aspx?sourcedoc=%7BAB2D3EED-448F-4D6E-9BA2-A41CD63EE57B%7D&file=20221201_Proposal_Template_SPs_ESCAPE_DMTSP-251122.docx&action=default&mobileredirect=true	UNIMAN/Lund University, TUM, CNRS	CC-BY
Talk at CHEP23, May 2023	Jared Little - **The Dark Matter Science Project in the European Open Science Cloud (EOSC) Future**	https://indico.jlab.org/event/459/sessions/2024/#20230508	CNRS	CC-BY
Talk at CHEP23, May 2023	Elena Gazzarrini - **The Virtual Research Environment: towards a comprehensive analysis platform**	https://indico.jlab.org/event/459/contributions/11671/	CERN	CC-BY
Webinar, June 2023	Elena Gazzarini, Jared Little - **Dark Matter: An EOSC-**	https://eoscfuture.eu/eventsfuture/dark-matter-an-eosc-	CERN, CNRS	CC-BY

	Future Science Project Webinar	future-science-project-webinar/		
Talk at ESCAPE Collaboration Meeting, July 2023	Caterina Doglioni, **EOSC Future - Dark Matter Science Project**	https://indico.in2p3.fr/event/30249/contributions/126987/	UNIMAN/Lund University	CC-BY
Talk at TAUP2023, August 2023	M. Smirnov, **Open-Science Integration of a Combined Analysis of KM3NeT and CTA into the EOSC Infrastructure**	https://indico.cern.ch/event/1199289/contributions/5450452/	FAU	CC-BY
Talk at Fysikdagarna (Swedish physics days), June 2023	C. Doglioni, **The ESCAPE Dark Matter Science Project**	https://indico.fysik.su.se/event/7762/contributions/12565/	UNIMAN/Lund University	CC-BY
EOSC Future Webinar	E. Gazzarrini, J. Little, **Dark matter: An EOSC Future Science Project webinar**	https://eoscfuture.eu/eventsfuture/dark-matter-an-eosc-future-science-project-webinar, https://indico.in2p3.fr/event/30241/	CERN, CNRS	CC-BY

7 Extreme Universe (Science Project no.6)

7.1 Description

Name	Extreme Universe
Description	The TSP-Extreme Universe includes several pilot projects for analysing extreme phenomena in the universe. One project analyses blazars, a type of Active Galactic Nuclei (AGN) that emits intense radiation, using a multi-wavelength and multi-messenger approach. Another project (MatchMaker) studies fast radio bursts to gain insight into the astrophysical processes that produce them. Gravitational wave detection is also being used to study extreme phenomena, such as black holes and neutron stars, with unprecedented detail. The Wavefier project integrates data from various sources to gain a more complete understanding of astrophysical phenomena. Lastly, the KM3NeT observatory is being used to study high-energy cosmic rays and neutrinos from point sources in the sky, including gamma-ray bursts and other transient events.
URL	https://escape2020.pages.in2p3.fr/virtual-environment/home/sp-extreme_universe/
WPs and Tasks involved	WP6, T6.3 Also worked with T6.1, T6.2 Partners in SP6: SNS-Pisa/EGO, UvA, NWO, CNRS, INFN, FAU
Partners	SNS-Pisa/EGO, UvA, NWO, CNRS, INFN, FAU

7.2 Impact and Innovation

Innovation	Ability to combine different experiment outputs together for the first time, making use of services discovered through EOSC Exchange, to allow potential new insights into extreme astrophysical phenomena. The main results are innovative algorithms, workflows, and scientific results.
Sustainability Plan	The main results are innovative algorithms, workflows, and scientific results. These will be maintained by the scientific community as long as there is interest in continuing the

	analysis. ESCAPE is now an Open Collaboration and creates the framework within which to maintain the results as a collaborative activity between the RI partners of the collaboration. All of the software, data, workflows are open access and licensed, and are being published into the EOSC Exchange and marketplace via the ESCAPE OSSR and data portals. They will all be openly available and findable, and usable according to the various licensing conditions. The actual scientific results will be published as open access papers and made findable through the marketplace.
Exploitation Plan	The main user groups for the results of this Science Project are primarily other scientists and citizen scientists. The main uses and exploitation of the results are: • Use for further research, validation of results, access to new algorithms and software; • We have demonstrated the ability to use EOSC to perform open science, and to make all the resources available to non-ESCAPE users who can repeat or adapt the analyses; • The main channels to deliver messages are: domain conferences, EOSC Symposia, web (EOSC-Future, ESCAPE, and Science Clusters), news items; • All the software, data, workflows are open access and licensed, and are being published into the EOSC Exchange and marketplace via the ESCAPE OSSR and data portals. They will all be openly available and findable, and usable according to the various licensing conditions. The actual scientific results will be published as open access papers, and also made findable through the marketplace.
Geographical Market	Global

7.3 Dissemination Artefacts

Name	Description	URL	IP owners	Licence used
Web site	Extreme Universe and Gravitational Waves - Science Project	https://projectescape.eu/extreme-universe-and-gravitational-waves-science-project	ESCAPE	CC-BY
Web site	Extreme Universe Project Description	https://escape2020.pages.in2p3.fr/virtual-environment/home/sp-extreme_universe/	ESCAPE	CC-BY
Web site	EOSC-Future Data in Action: The Extreme Universe and Gravitational Waves	https://eoscfuture.eu/data/the-extreme-universe-and-gravitational-waves/	EOSC-Future	CC-BY
Web SIte	EOSC Portal - EOSC in Practice Stories: D.Vohl, Searching for bright compact needles in a radio haystack of galaxies and stars	https://eosc-portal.eu/searching-bright-compact-needles-radio-haystack-galaxies-and-stars	UvA	CC-BY
Web Site	Science Clusters Web Site - Use	https://science-clusters.eu/use-	Lund University, CNRS, TUM, SNS	CC-BY

	Cases; Maximizing Collaboration on Dark Matter and Gravitational Wave Research	cases/collaboration-dark-matter-gravitational-wave-research		
Talk at 30th International Symposium on Lepton Photon Interactions at High Energies, January 2022	Cuoco E., **Gravitational Waves and Multi-Messenger Astrophysics: perspectives with Machine Learning**	https://indico.cern.ch/event/949705/contributions/4575406	EGO/SNS	CC-BY
Presentation material and video recordings at ESCAPE to the Future Conference, October 2022	Elena Cuoco, **Extreme Universe Science Project Overview**	https://projectescape.eu/events/escape-future	EGO/SNS	CC-BY
idem	Iess A., **Wavefier: a framework for Multi-messenger astrophysics**	https://projectescape.eu/events/escape-future	EGO/SNS	CC-BY
idem	Alessandro Parisi, **Measuring the Dark Matter environments of black hole binaries with gravitational waves**	https://projectescape.eu/events/escape-future	SNS	CC-BY
idem	Mikhail Smirnov, **KM3NeT IRF for point-source analysis**	https://projectescape.eu/events/escape-future	FAU	CC-BY
idem	Dany Vohl, **Extreme Universe VRE Demo: Searching for FRB persistent radio source counterparts in dwarf galaxies using LOFAR**	https://projectescape.eu/events/escape-future	UvA	CC-BY
idem	Elena Cuoco, **Consolidation and Open Challenges**	https://projectescape.eu/events/escape-future	EGO/SNS	CC-BY
Science Project Template, November 2022	E. Cuoco, I.Bird, **ESCAPE Science Project for EOSC-Future: TheExtreme Universe Science Project**	https://technopolisltd223.sharepoint.com/:w:/r/sites/INFRAEOSC-03Proposal/_layouts/15/Doc.aspx?sourcedoc=%7B61EA12D6-ADDD-4E1C-	EGO/SNS, CNRS	CC-BY

		A5B7-F9B90D90E311%7D&file=20221201_Proposal_Template_SPs_ESCAPE_EUTSP-011222.docx&action=default&mobileredirect=true		
Webinar, May 2023	Dany Vohl, Alberto less, **Use Cases from the EOSC Community - ESCAPE Gravitational Waves**	https://eoscfuture.eu/eventsfuture/eosc-future-use-case-event/	UvA,EGO/SNS	CC-BY
Talk at ESCAPE Collaboration Meeting, July 2023	Elena Cuoco, **EOSC Future - Extreme Universe Science Project**	https://indico.in2p3.fr/event/30249/contributions/126986/	EGO/SNS	CC-BY
Talk at TAUP2023, August 2023	M. Smirnov, **Open-Science Integration of a Combined Analysis of KM3NeT and CTA into the EOSC Infrastructure**	https://indico.cern.ch/event/1199289/contributions/5450452/	FAU	CC-BY
WDF Software Website	Web site of Wavelet Detection Filter Software	https://wdfpipe.gitlab.io/	EGO	GPLv3
Wavefier Software Website	Web site of Wavefier MMA Software	https://wavefier.gitlab.io/	EGO	GPLv3

8 Tracing Bio-Structures (Science Project no.7)

8.1 Description

Name	Tracing Bio-Structures with serial crystallography
Description	The main goal of this science project is to show the added value of EOSC by employing Virtual Infrastructure for Scientific Analysis (VISA) within the realm of the emerging serial crystallography research community. The project is expected to onboard researchers from the photon and neutron community in EOSC.
URL	• https://visa.esrf.fr/ • https://github.com/ILLGrenoble/visa-api-server • https://data.esrf.fr/
WPs and Tasks involved	WP6, task 6.3
Documentation	• https://github.com/ILLGrenoble/visa-api-server • https://visa.esrf.fr/help

8.2 Impact and Innovation

Innovation	In this science project, we have showcased how the serial crystallography community can benefit from the VISA service.

	During the experiment conducted at ESRF, we employed synchrotron X-rays to capture micro-crystals (10µm) of proteins and recorded one diffraction frame per crystal. By accumulating millions of 2D diffraction frames, we reconstructed intricate 3D bio-structures. The sheer magnitude of the collected datasets amounted to an impressive 25 TB, accumulated over a 12-hour exposure period. It is noteworthy that the ESRF facilitates the collection of approximately 100 TB of data per experiment in this category. For users from smaller laboratories, who lack access to high-performance computing clusters and sufficient storage capabilities, processing such massive volumes of data becomes an immense challenge. ESRF, being a prominent research infrastructure, provides its users with the necessary computing power and storage to tackle data processing at this scale. This invaluable resource extends to the global scientific community, attracting approximately 13 research groups worldwide to leverage ESRF's serial crystallography infrastructure at ID29. However, due to limited access duration (beamtime), most of their time is allocated to raw data collection, leaving data analysis to be conducted upon their return to their respective home institutions. VISA enables remote access for data processing through a user-friendly web interface. This way, researchers can process their data efficiently without the need to physically be present at the ESRF facility. Notably, both VISA[1] and the ESRF Data portal[2] are available on the EOSC marketplace, offering external users the opportunity to request access to the ESRF HPC cluster, and storage resources, facilitating the analysis of open data available at ESRF. The true innovation of this scientific project lies in its demonstration of the EOSC's value for fostering the adoption of open-access services provided by the research infrastructures. [1] https://marketplace.eosc-portal.eu/services/eosc.esrf.visa?return_path=search%2Fservice&search_params=q%3Dvisa&source_id=9423dba2-01cf-44c4-8f14-c3af1ed4cad3&client_uid=afac3a70-7e4b-4f24-9b2e-62e99532ae87 [2] https://marketplace.eosc-portal.eu/services/eosc.esrf.ecc74ab09791c52aa238ee77ae988874?return_path=search%2Fdata-source&search_params=q%3Desrf%2520data%2520portal&source_id=d2cdc3fa-299b-4bbc-9fcf-c4de4c25b039&client_uid=afac3a70-7e4b-4f24-9b2e-62e99532ae87
Sustainability Plan	This particular science project does not yield any significant results that necessitate intellectual property (IP) protection. All the individual components utilized in this project's workflow are already open-source. The VISA service, developed at ILL, is available under the GPL v3 license, ensuring its openness and accessibility. Similarly, the open datasets available in the ESRF data portal are licensed under Creative Commons Attribution 4.0, promoting unrestricted use and sharing. The science project itself will be supported until the end of EOSC Future project i.e 30 September 2023. VISA has seamlessly integrated into ESRF IT infrastructure, thanks to its widespread adoption by ESRF users. It is now an integral part of the ESRF IT services catalogue, ensuring its continued presence. ESRF is committed to upholding the sustainability of VISA, as it has evolved into a fundamental component of our mission. ILL will assume responsibility for ongoing development and maintenance of the code. The Memorandum of Understanding (MoU) to ensure the sustainability of VISA is actively being worked on by the Research Infrastructures (RIs) within the PaN community The expected benefits encompass the successful onboarding of the photon and neutron research community to EOSC.
Exploitation Plan	The VISA service and ESRF Data portal have the potential to extend their applications beyond the realm of serial crystallography and other research infrastructures. VISA service can be adapted for use in various scientific domains, including photon and neutron infrastructures, and can be deployed in other research facilities as well. This broad applicability makes it an excellent candidate for use as a compelling use case in policy development, emphasising the significance of open-source services and data emanating from research infrastructures.

	The primary aim of this science project is to provide essential support to the serial crystallography community at ESRF ID 29[1] for efficient data processing. The key focus is to highlight the successful integration of open-source services like VISA, available on EOSC, and showcase how it propels scientific advancements. The demonstration of this project's success will serve as a compelling case for deploying VISA at other beamlines within the ESRF. The subsequent objective could be to facilitate the adoption of the VISA service by other photon and neutron infrastructures to better assist their respective user communities. [1] https://www.esrf.fr/id29
Geographical Market	The global geographical region includes a broader audience, such as research infrastructures and institutes worldwide.

9 Dynamics of Biological Processes (Science Project no.8)

9.1 Description

Name	Dynamics of Biological Processes
Description	sas_helper Jupyter notebook is designed for the analysis and fitting of SAXS/SANS data. This notebook utilises existing open-source software and libraries, including the Inria library and IMP library, offering a simple syntax and commands for Python. The analysis results are visualized using the interactive plotting capabilities of matplotlib as a backend.
URL	https://replay.notebooks.egi.eu/ with https://github.com/isafiulina/sas_helper as a binder
WPs and Tasks involved	WP6 WP6.3
Partners	Institut Laue-Langevin (ILL) as a main developer of the project "Dynamics of Biological Processes" and associated Jupyter notebook for SAXS/SANS data analysis and fitting. European Grid Infrastructure (EGI) as a provider of a platform for Jupyter Notebooks and storage space. European Synchrotron Radiation Facility (ESRF) as one of the project coordinators and IT support.

9.2 Impact and Innovation

Innovation	**Empowering Users and Scientists:** The project provides users and scientists with a comprehensive introduction and a set of tools for their data analysis needs. This empowers researchers with the capability to perform in-depth analyses efficiently. **User-Friendly and Accessible Platform:** The Jupyter Notebook serves as a powerful analysis tool and a valuable platform for training students and non-expert users. Thanks to the EGI Replay platform, it consolidates existing software into a single accessible environment, eliminating the need for individual software downloads and familiarization. **Advancements in Multimodel Fitting:** A significant milestone has been achieved by creating a unique tool for multimodel fitting, expanding the platform's capabilities to work with both SAXS and SANS data. This enhances the quality and depth of data analysis. **Scalable and Adaptable:** The Jupyter Notebook architecture allows for continuous improvement and enhancement of the available tools. This ensures that users have access to the latest advancements in data analysis, making the platform future-proof and adaptable to evolving research needs. Overall, the innovative Jupyter Notebook developed in this project not only addresses existing challenges in data analysis but also opens up new possibilities for scientific advancements and wider societal impact in the field of structural biology.

Sustainability Plan	The sustainability of the results is ensured through ongoing in-kind support and manpower contributions from research infrastructures and institutions, fostering continuous development, maintenance, and user support.
Exploitation Plan	***Group 1: ILL Users and Scientists*** **Exploitation/Use Objective:** Demonstrate to expert users at Institut Laue-Langevin (ILL) that utilizing the Jupyter Notebook and its associated tools can significantly enhance their efficiency in data analysis. Highlight how the platform streamlines the analysis process, saving time and effort while delivering accurate results. **Main Messages:** Emphasize the powerful capabilities of the Jupyter Notebook and its integration with existing open data from PDB and SASBDB databases. Showcase how the platform enables advanced analysis techniques, providing researchers with comprehensive insights into their experimental data. **Best Channels to Deliver Messages:** Presentations and workshops at ILL user meetings and conferences, as well as tailored training sessions for scientists using the facility. **Access and Use of Results:** ILL users can access the Jupyter Notebook and related resources through internal channels or dedicated platforms provided by the project. The terms of use will be defined within the project, and users can approach the project coordinator or designated contacts for support and guidance. ***Group 2: ESRF Users and Scientists*** **Exploitation/Use Objective:** Illustrate to expert users at the European Synchrotron Radiation Facility (ESRF) the benefits of adopting the Jupyter Notebook for data analysis. Showcase how the platform simplifies complex analyses, allowing researchers to focus on interpreting results and advancing their scientific investigations. **Main Messages:** Highlight the unique feature of multimodel fitting in the Jupyter Notebook, demonstrating its applicability to both SAXS and SANS data. Showcase the platform's compatibility with existing data repositories, enabling seamless integration with ESRF-generated datasets. **Best Channels to Deliver Messages:** Presentations and demonstrations at ESRF user meetings, beamline workshops, and scientific conferences where ESRF users and scientists are present. **Access and Use of Results:** ESRF users can access the Jupyter Notebook through designated channels provided by the project. Information on access and usage terms will be communicated through official project communication channels. Users can approach project representatives or designated contacts for support and guidance. ***Group 3: SAS Community*** **Exploitation/Use Objective:** Engage the Small Angle Scattering (SAS) community and demonstrate the value of the Jupyter Notebook and its capabilities. Showcase how the platform bridges the gap between experts and non-experts, making SAS data analysis more approachable and accessible. **Main Messages:** Highlight the simplicity and user-friendliness of the Jupyter Notebook, focusing on its potential for both novice and experienced SAS researchers. Emphasize the collaborative and community-driven nature of the project, encouraging active participation and feedback. **Best Channels to Deliver Messages:** Presentations and workshops at SAS-related conferences, workshops, and user group meetings. Utilize online platforms and forums where the SAS community gathers. **Access and Use of Results:** The SAS community can access the Jupyter Notebook and related resources through dedicated platforms and project communication channels. Terms of access and usage will be clearly communicated, and the community can approach project representatives or designated contacts for support and collaboration. ***Group 4: Structural Biologists (Outside SAS Community)*** **Exploitation/Use Objective:** Showcase the value of Small Angle Scattering (SAS) techniques to structural biologists who are not familiar with the SAS community. Provide simple and accessible tools that have low entry barriers, making SAS data analysis approachable to this group. **Main Messages:** Communicate the benefits of SAS techniques in providing structural insights for a wide range of biological processes. Highlight how the Jupyter Notebook

	simplifies the analysis process, allowing structural biologists to gain valuable information from scattering data. **Best Channels to Deliver Messages:** Presentations and workshops at structural biology conferences, research institutes, and educational institutions where structural biologists are present. **Access and Use of Results:** Structural biologists outside the SAS community can access the Jupyter Notebook and related resources through public platforms or project communication channels. The project will aim to provide user-friendly documentation and tutorials to facilitate adoption. Users can approach project representatives or designated contacts for further assistance and collaboration.
Geographical Market	**Local Geographical Region:** The local geographical region primarily includes the organizations involved in the project, such as Institut Laue-Langevin (ILL), European Synchrotron Radiation Facility (ESRF), and the SAS community in general including large scale facilities. These organizations are actively engaged in the project's development, and the result is tailored to meet their specific needs and requirements. **Global Geographical Region:** The global geographical region includes a broader audience, such as researchers, scientists, and structural biologists worldwide. The Jupyter notebook and its tools are designed to be accessible and applicable to researchers and users across different geographical locations, regardless of their expertise in small-angle scattering techniques. Since the project aims to provide a user-friendly platform for data analysis and facilitate access to open data from databases like PDB and SASBDB, it is intended to benefit researchers and scientists globally. The platform's online accessibility and integration with existing resources make it a valuable asset for the wider international scientific community interested in structural biology and small-angle scattering analysis. In summary, the result has been developed to serve both the local organizations involved in the project and the global scientific community, aiming to facilitate and enhance small-angle scattering data analysis worldwide.

10 Climate Neutral and Smart Cities (Science Project no.9)

10.1 Description

Name	Climate Neutral and Smart Cities
Description	The project demonstrates that relevant environmental data and data on citizens' values, attitudes, behaviour, and involvement can be combined for social, political and scientific analysis. Indicators regarding environmental indices have been developed based on data from the European Environmental Agency (EEA) and Copernicus ERA5 and methods, and workflows for computing related variables have been produced. The resulting indicator variables are integrated with data for a selection of big cities in Europe for three rounds of the European Social Survey (ESS). The Python code for getting the raw data, marshalling, and processing the data is made available for reuse. In addition, a prototype application for describing the full workflow related to the data integration and variable creation has been developed. A set of papers related to the project work are produced. All deliverables resulting from the project are being published on a related web-page under the ESS namespace (ESS Labs). The ESS Labs page has been made available as a service from the EOSC Portal. The service has been integrated with the EOSC Helpdesk Core service in the form of a light integration.

URL	• ESS Labs web page: https://www.europeansocialsurvey.org/esslabs • Link to the DDI-CDI provenance description application repository https://github.com/sikt-no/ddi-cdi_process2web • Link to the repository for the code used to integrate data and create variables https://github.com/sikt-no/ess-labs-data-sp9 • Links to papers: – 'Climate Neutral and Smart Cities, combining data about city people and their physical environment' https://preprints.arphahub.com/article/107872/ – 'Climate and Air Quality Indices for the European Social Survey' https://preprints.arphahub.com/article/114675/ – 'Reuse and reproducibility: Describing Cross-Domain Research Data in the Science Project *Climate Neutral and Smart Cities*'. Link: https://preprints.arphahub.com/article/115047/ – 'DDI-CDI-Workflow description of the EOSC Future WP6 Task 3, Science Project 9 'Climate Neutral and Smart Cities''. Link: https://preprints.arphahub.com/article/116667/ – 'Climate and Air Quality Data in Attitudinal Research: A Feasibility Use-Case' Link: https://zenodo.org/records/10581886
WPs and Tasks involved	WP 6 Task 3 (Science Projects)
Partners	Partners are ESS ERIC (ESS Head Quarters (HQ); Sikt) and CESSDA ERIC (SND; ADP) from the SSHOC cluster, and IAGOS from the ENVRI community. • Technical prototypes are developed by Sikt • Indicator variables are developed by SND and Sikt • Code for data retrieval, processing and merging is developed by Sikt and SND • Scientific papers are written by: – Project Template: ESS HQ, Sikt – Data and Methods paper: SND, Sikt, ADP, IAGOS – Reuse and Reproducibility paper: Sikt – Workflow paper: Sikt – Scientific paper: ESS HQ

10.2 Impact and Innovation

Innovation	The innovation of the project is firstly related to its cross-community collaboration, involving experts from very different research fields (social surveys, air quality and climate), resulting in reliable project outcomes from the perspective of the SSHOC and ENVRI clusters. Secondly, the project will demonstrate possible use of the new metadata standard DDI Cross-Domain Integration (DDI-CDI) in two ways: • The 'Process' part of the specification is used as the basis for a provenance description prototype application developed by the project. This will be used for documenting the data management workflows of the project. • In addition the project will be a use case for DDI-CDI specification. Thirdly, a set of reusable environmental indicators have been produced that are well adapted for analysis together with social attitude data. The resulting data is available for scientific analyses. Fourthly, the code used to produce the integrated data has been made available. Fifthly, a set of research papers are produced, among those a scientific paper containing data analysis results.

Sustainability Plan	• Deliverables are freely and openly accessible from the EOSC Platform and the ESS Labs page. For IP rights, Copyrights, licensing, and contribution, see the IP Foreground section. • Ongoing access to and support of the project results as delivered will be secured. • Requirements are foreseen to be minimal: existing ESS ERIC infrastructure can be leveraged to provide ongoing access and support to the project results as delivered. Any extensions to the ESS Labs from related projects are possible in future, and funding possibilities will be explored within the frames of those projects. We foresee researchers and implementers to be the primary users: • Researchers will have continuous access to integrated data from the European Social Survey together with data from the European Environmental Agency (EEA) and Copernicus ERA5 for the selected regions and time period delivered by the project. A set of indicators have been produced that facilitates analyses of climate and air quality indices relative to the timing of the interview. • Researchers can read about the data integration and indicator production methods used and explore preliminary analyses • The provenance application tool allows users to explore the process step used to create each indicator variable, and displays the logic behind their computation, thus providing a fully transparent access to the workflow and forms basis for evaluation of the work done. • Implementers can install and use the process prototype application to document their own workflows.
Exploitation Plan	We foresee that social scientists will be the primary users in the short term, for the purpose of performing scientific research on integrated data from the environmental and social research domains. Policy makers and researchers from the environmental domains are foreseen as users for the intermediate or long term. Due to the relevance of the project's topic and that the combined data are prepared to be easy to use, we believe it could serve as a valuable resource for educational purposes as well. Other infrastructure projects may find the use of metadata standards and the prototype tools developed in the project useful for their own purposes. Users are able to access data, tools and scientific papers resulting from the project through the ESS Labs prototype application that has been made available as a service from the EOSC Portal. All deliverables from the project are openly available through the licenses described in the IPR section below in this document. The main message we would like to convey is about the achievements of this project, both when it comes to making a new, cross-domain data resource available and providing a prototype application for describing workflows and processes. We will promote this by organising webinars, writing papers and presenting the project in multi-disciplinary sessions at various conferences, presenting both research deliverables and metadata technical solutions. We are also reaching out to other infrastructures regarding reuse of our methods in their projects. One example is the Horizon-Infra-2024-TECH-01-02 application SoGreen (application number 101188188).

	The project is supporting standardisation activities as it will serve as a use-case for the DDI-CDI metadata standard, as well as for other metadata related initiatives.
Geographical Market	The project is intended for a global market

11 Access Management for Distributed RIs (ARIA) (Science Project no.10)

11.1 Description

Name	Access Management for Distributed RIs (ARIA)
Description	The solution enables data repositories that publish Social Science and Humanities study level metadata catalogues to use ARIA for processing data access request applications from Consumers (such as researchers). This can be integrated into their catalogues, so that Consumer requests for access to restricted data become seamless.
URL	https://aria.services/
WPs and Tasks involved	WP6 Task 6.2 - Integration of EOSC-Core Services into European Research Practice Task 6.3 - Demonstrating EOSC Value through Cross-domain Research Test Science Projects Leader of SP10 is CESSDA ERIC

11.2 Impact and Innovation

Innovation	Managing Consumer access requests for restricted data can be particularly challenging for growing organisations that do not (yet) have solutions in place to manage such requests. At the same time, existing solutions might not be easily adaptable to changing needs and requirements of the institution. In contrast to adoption of generic workflow solutions, such as ticketing systems, the use of the ARIA platform enables repository providers to outsource the work with all of its parts into a single platform, where the interactions with Consumers and data access request reviewers are all realised in a single place. ARIA has been built from the ground up to manage access to scientific resources (initially physical but adapted through this project to manage virtual resources as well). It has an integrated review process, reporting and feedback collection, etc. ARIA is "generic" in as far as it can be customised for specific applications - e.g., in the Consortium of Social Sciences Data Archives (CESSDA) use case a white labelled version could be created, with appropriate branding and catalogue, but utilising the same platform under the hood. An additional (optional) feature is the use of central user management through the European Open Science Cloud (EOSC) Authentication and Authorization Infrastructure (AAI), in particular LifeScience ID [8], for managing user identities directly in ARIA.
Sustainability Plan	• The service can be accessed by any Consumer of the study level metadata that uses the ARIA platform to manage data access requests. For instance, the So.Da.Net catalogue metadata is harvested by the CESSDA Data Catalogue, and aggregated for there by BASE, GoTriple, OpenAIRE, B2Find and Google Datasearch. Therefore, a Consumer could in principle find the same metadata record in any of those catalogues and be directed to the ARIA platform data access request management service in the same way from any of those locations. • There is no cost to Consumers of metadata/data at the point of use. • Instruct-ERIC offers ARIA to scientific projects and Research Infrastructures (subject to fees), but this is done as an agreement between the individual Research Infrastructure/Project and Instruct-ERIC (Research Infrastructures are facilities that provide resources and services for research communities to conduct research and foster innovation). – This is subject to timelines and Instruct's capacity at the time.

	– ARIA is not a commercial platform that anyone can sign up to with a credit card. • The ARIA platform data access request management service is supported and maintained by Instruct-ERIC. • Requests for access to study data must be managed by the publisher of the data, in each case. • Instruct-ERIC has legal rights over the ARIA platform, granted by perpetual licence. The IP rights are currently held by Oxford University for historic reasons. • Benefits to metadata aggregators (such as CESSDA ERIC): – Consumers see consistent data access request mechanisms across multiple publishers (assuming wide-spread adoption) – White label version of ARIA can be created, with appropriate branding, so it looks like a CESSDA service • Benefits to data publishers: a single, seamless workflow for managing data access requests. Various models are supported, including – Verification as researcher – Verification of successful completion of required training courses or certification – Verification of successful completion of required documents, such as e.g. ▪ Terms-of-Use ▪ Licence Agreement ▪ Non-Disclosure Agreements – Submission of a full research proposal for formal review with fields required by the data owner or data processor
Exploitation Plan	The ARIA pilot involves a single CESSDA Service Provider. There are many other CESSDA SPs that could benefit from using this solution for managing data access requests. However, as stated above, a fee is payable to Instruct-ERIC for the use of the ARIA platform to provide this service. In the same way, 3rd parties could adopt a similar solution (subject to commercial negotiations with Instruct-ERIC) as an Interoperability Guideline has been produced to explain the integration process. See https://doi.org/10.5281/zenodo.7805774.
Geographical Market	Europe