E-RIHS IP

European Research Infrastructure for Heritage Science

IMPLEMENTATION Phase

CALL: HORIZON-INFRA-2021-DEV-02 | TYPE OF ACTION: CSA | GA n. 101079148

D1.2 E-RIHS IP Project DMP

Lead Authors: Lucilla Spini and Vania Virgili

With contributions from: Laura Benassi and Joseph Padfield

ABSTRACT

The Data Management Plan for E-RIHS IP pertains Deliverable D1.2, within WP1, and in particular is an output of Task 1.3 "Commitment to responsible science" under the leadership of CNR. The document addresses the different types of data and related metadata to be generated or re-used by the project, the application of open science practices and the FAIR data principles, as well as data security and ethical considerations. Building on the E-RIHS PP outputs and the E-RIHS Data Strategy. The E-RIHS IP Data Management Plan is to be considered as working and living document to be monitored, reviewed and updated throughout the project.

DOCUMENT INFORMATION

Project number	101079148	Acronym	E-RIHS IP
Full title	European Research Infrastructure for Heritage Science Implementation Phase		
Project url	www.e-rihs.eu		
Document url			
EU Project Officer	Emiliano CAROZZA		

Deliverable	Number	D1.2	Title	E-RIHS IP Project DMP
Work Package	Number	1	Title	Project Management and Coordination

Deliverable nature	Report	
Dissemination level	☒ Public ☐ Confidential ☐ Restricted	
Contractual delivery date	(02/06/2023)	
Actual delivery date	(28/02/2023)	
Status	Version 0.6	☐ Draft ☒ Final

Authors (Partner)	Lucilla Spini (CNR-ISPC) and Vania Virgili (CNR-ISPC)	
Responsible Author	Name: Vania Virgili	Partner: CNR-ISPC
First Author	Name: Lucilla Spini	Partner: CNR-ISPC

Total number of pages	12
Keywords	Data Management Plan

Version Log			
Issue Date	**Rev. no.**	**Author**	**Change**
11/04/2023	0.1	Lucilla Spini	Original Draft
13/04/2023	0.2	Vania Virgili	Revision of document
26/04/2023	0.3	Laura Benassi	Revision of document
03//05/2023	0.4	Joseph Padfield	Revised version
31//05/2023	0.5	Lucilla spini	Revised version
02/06/2023	0.6	Vania Virgili	Final version

DISCLAIMER: This document reflects only the author(s)'s view and the Agency and the European Commission are not responsible for any use that may be made of the information it contains.

TABLE OF CONTENTS

ABBREVIATIONS

Abbreviations	Expansion
CC	Creative Commons
CH	Cultural Heritage
DARIAH	Digital Research Infrastructure for the Arts and Humanities
DiSSco	Distributed System of Scientific Collections
DMP	Data Management Plan
DOI	Digital Object Identifier
EC	European Commission
E-RIHS IP	European Research Infrastructure for Heritage Science Implementation Phase
EOSC	European Open Science Cloud
ESFRI	European Strategy Forum on Research Infrastructures
FAIR	Findable, Accessible, Interoperable, and Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
iGA	interim General Assembly
HLEG	High Level Expert Group
HS	Heritage Science
NG-UCL	The National Gallery – University College London
ORE	Open Research Europe
PID	Persistent Identifier
RI	Research Infrastructure
RDA	Research Data Alliance
SC	Steering Committee
SSHOC	Social Sciences and Humanities Open Cloud
T	Task
VRE	Virtual Research Environment
WP	Work Package

INTRODUCTION

The European Research Infrastructure for Heritage Science Implementation Phase (E-RIHS IP) Project aims at enabling E-RIHS ERIC (E-RIHS) – a pan-European distributed Research Infrastructure (RI) for Heritage Science (HS) – to start its operation phase and position itself as a reference for excellent research for HS. The E-RIHS IP is set on the innovative vision of fostering an interdisciplinary domain where hard sciences, engineering and humanities join together their capacities – with equal importance and roles – in the interdisciplinary (and often transdisciplinary) scientific field of Heritage Science (HS).

While the E-RIHS IP is not expected to deliver new research data *per se*, it places particular attention on data management practices related to data and meta-data pertaining to the integration of different communities in the cultural heritage sectors encompassing scientists, humanities scholars, as well as practitioners. This document needs to take into account the management of data and meta-data of the E-RIHS IP, while taking into account the ensuing development of a specific E-RIHS ERIC Data Management Plan.

In this context the E-RIHS IP Data Management Plan (DMP) is developed taking into account the following issues:

1. The specific characteristics of the project facilitating the implementation of a pan-European distributed infrastructure enabling research in the interdisciplinary domain of Heritage Science (HS).
2. The E-RIHS IP commitment to the promotion of Open Science Practices, FAIR Data Principles, and ongoing developments in Persistent Identifiers (PIDs).
3. The good practices developed in E-RIHS PP and IPERION HS, as well as existing tools (e.g., the EOSC Alignment Toolkit for data FAIRification).
4. The recommendations from ESFRI and HLEG (see 2.5/ESFRI and 8/HLEG) to foster E-RIHS ERIC readiness for EOSC and FAIR data.
5. The different types of ensuing data and meta-data (e.g., personal data) and the specificities of the identified internal and external stakeholders (e.g., data/meta-data developers and/or users) of the E-RIHS IP.
6. The EC template for developing Data Management Plans.

The plan has been designed as the venue to document and record the agreed data management procedures and protocols used by the project partners to organize, preserve, and share the data and meta-data produced during the project and to make them as open and FAIR as possible, but also as closed as necessary. It encompasses the various plans to manage and archive, where appropriate, data and meta-data related to institutions and individual researchers (and other stakeholders) engaged in the E-RIHS community as well as other outputs (e.g., project presentations, procedural work protocols, communications, discussions and the project public presentations in social media and the project website). The E-RIHS IP DMP also outlines how the work of the project will be transferred, upon completion, to the E-RIHS ERIC for long-term management. This DMP is developed as a working and living document, which will be periodically monitored, reviewed and updated in order to take into account the latest developments in data/meta-data management throughout the different phases of the project.

It is important to note that this document is the DMP for the E-RIHS IP Project, and not the one foreseen for ERIHS ERIC to be developed under the leadership of NG-UCL.

1. DATA SUMMARY

E-RIHS IP DMP is relevant to different WPs which relies on the ability to collect, process and re-purpose information gathered. Given that this phase regards the Implementation Phase towards the launch of the E-RIHS ERIC, the data and meta-data to be collected and curated pertains the procedures towards the establishment and consolidation of the Research Infrastructure. In this context, the data and meta-data will be collected within the context of the operationalisation process, and not from purposively-designed *ad hoc* research.

Hence, the data and/or meta-data will be collected within the framework of the following actions:

1. Identification and aggregation of all prior reports (ensuing from e.g., E-RIHS PP, IPERION HS).
2. Enlargement of the membership (T4.2 and WP6).
3. Establishment of the governance structure (T1.1 and WP2).
4. Finalisation of the distributed architecture (WP2).
5. Development of ICT and data management solutions, including possible open access to data (T5.1 and T6.2).
6. Development of access policies and user strategies (T5.2, T5.3, T5.4, T5.5 and T6.2).
7. Staff-related issues (T2.3 and T3.1) and procurement-related issues (T3.2).

In addition to newly generated data and meta-data, the E-RIHS IP builds on existing data and meta-data developed by the E-RIHS PP, as well as data and meta-data ensuing from the utilization of IPERION HS services (e.g. ARCHLAB, MOLAB, and FIXLAB) by the HS community; and thus, it maintains contacts with IPERION HS, at the operational level, with periodic meetings and sharing of data, protocols and best-practices, for the exploitation of knowledge and tools from IPERION HS actions that have a direct bearing on E-RIHS IP (e.g., central support activities, #HSAcademy, principles and prototype studies for DIGILAB, and the Strategic Development Agenda).

The data and meta-data can be grouped as followed:

1. **EXTERNAL DATA: (i) background data** from completed project activities within E-RIHS PP and other related projects (e.g., IPERION HS); and **(ii) newly developed data** from *other* concurrent projects and activities. The data will comprise all the reports and outputs identified; and in case the background and/or newly developed data is not public, reference will be the made to the sources.

2. **PROJECT DATA**: Data (and meta-data) from procedural work concerning the implementation of the WPs. The following main categories are identified:
   a. **PROCEDURAL DATA**: Data (meta-data) related to implementation procedures and governance structure (e.g., progress reports, procedural protocol, minutes, recommendations and decisions). Some of this type of data which is not made public (e.g., board nominations including personal data, see below), reference will be made to the sources.
   b. **INSTITUTIONAL DATA**: Data (and meta-data) from members and other potential member institutions. The E-RIHS IP will collect data from the members and potential members (e.g., via surveys[1] and *ad hoc* consultations) within the enlargement and consolidation of the membership. The data will include *inter alia*, (i) staff-related and procurement-related

---

[1] For surveys, E-RIHS IP plans to use the EUSurvey Platform (https://ec.europa.eu/eusurvey/)

practices; (ii) gender equality policies and practices, and (ii) procedures related to the application of open science and FAIR principles. The institutional data will be used for the implementation project, but also for developing the E-RIHS ERIC.

c. **PERSONAL DATA**: Data (and meta-data) from individuals. The E-RIHS IP is expected to collect personal data via surveys and questionnaires and/or within the organization of stakeholder events. The processing and management of personal data and the protection of privacy in recorded information (audio and/or visual) and electronic communications, as well as in public communication networks in full compliance with General Data Protection Regulation (GDPR).

d. **COMMUNICATION AND OUTREACH DATA**: Data (and meta-data) from outreach and communication activities related to the E-RIHS IP. This type of data includes open access scientific publications, grey literature publications, policy briefs and reports, project presentations, videos, images, project website (incl. analytics and downloads) and social media (incl. presences / likes). For further information, please refer to "E-RIHS IP Dissemination, Exploitation and Communication - D.6.1" (Benassi *et al*., 2023).

It is to be noted that a gender lens will be applied in the generation, management and re-use of data, metadata and other E-RIHS IP outputs. Whenever possible gender-disaggregated data are to be collected and analysed.

Concerning "data utility", the data and meta-data ensuing from the E-RIHS IP could be useful to both internal and external stakeholders, namely:

- *Internal Stakeholders*: the members and potential members of E-RIHS IP will benefit from the collected and curated data as they will be useful towards improving their readiness for active participation in E-RIHS ERIC, including adopting best practices in open science and FAIR data practices as well as in ethics and security.
- *External Stakeholders*: future users of E-RIHS ERIC will benefit from the data and meta-data in order to ensure compliance with RI procedures.   In addition, other consortia developing and establishing RIs could find the data and meta-data useful to identify best practices.

2. FAIR DATA

Building on the commitment to the practices of Open Science and FAIR principles by the E-RIHS PP (also taking into account the recommendations of the Commission HLEG of June 2020), the E-RIHS IP will continue to adhere to Open Science Practices and to apply the FAIR Principles to the data and meta-data ensuing from the Project. In this context, given the interdisciplinary nature of the E-RIHS community, particular attention will be placed on the different needs and approaches of the humanities scholars; hence, E-RIHS IP will follow the guidance provided by the Parthenos project[2]. Furthermore, through its engagement in the Digital Research Infrastructure for the Arts and Humanities (DARIAH, https://www.dariah.eu/), E-RIHS IP will also follow the development of DARIAH principles as complementary to the FAIR Principles within the framework of increasing digitalisation and application of advanced technologies within the CH sector.

---

[2] https://www.parthenos-project.eu/

Findability: Making data findable, including provisions for metadata

The E-RIHS IP will foster the uptake of Persistent Identifiers (PIDs) within the E-RIHS IP community and will ensure PID for e.g., digital objects – including data – individuals, activities, events, as appropriate. A DOI will be assigned to all outputs of the project (e.g., reports, presentations, and datasets) through Zenodo[3] or other mechanism (e.g., publishing in scientific journals). E-RIHS IP will also explore the possibility to assign PID to conferences and/or other events (see e.g., Frankel *et al*., 2022).

In addition, researchers engaged in the E-RIHS IP will be asked to foster a comprehensive utilization of PIDs, (also by registering for an ORCID at individual and appropriate ID such as ROR at the institutional level, if not already done); this will be conducive to further exploit PIDs in the E-RIHS ERIC.

E-RIHS IP project will ensure that data and meta-data (e.g., keywords) are findable/readable by both humans and machine so as to enable automatic discovery of datasets, services and outputs from the project.

Accessibility: Making data accessible

Following the practices implemented in E-RIHS PP, all outputs of ERIHS IP will be released as open access documents. Accessibility is ensured through a robust data life cycle within the E-RIHS IP project towards ultimately making it open access to the wider community through trusted repository (e.g., through Zenodo[4]).

Within the Project, services such as MS Office (including Teams) and a Virtual Research Environment (VRE) in D4Science (see: https://services.d4science.org/group/e-rihs_ip) are exploited for collaborative work and to ensure accessibility of the documents within the project. E-RIHS IP will continue to use the existing E-RIHS Zenodo community (see: https://www.zenodo.org/communities/e-rihs/?page=1&size=20) established for the E-RIHS PP.

Any scientific contributions or other outputs produced by the project could be published via the Open Research Europe (ORE) platform (https://open-research-europe.ec.europa.eu/), where appropriate.

The E-RIHS IP community is adopting and promoting a collaborative, open, sustainable, responsive and participatory research methodology and is speeding up the adoption of open science practices to improve the robustness, validity, reliability and marketability of its work. In this context, an Openaire "gateway" dedicated to Heritage Science (see https://heritage-science.openaire.eu/) is being developed with the aim to collect and, raise visibility of, the research outputs of the Heritage Science community. In fact, E-RIHS IP will implement the EU Open Access strategy and contribute to shaping how the E-RIHS ERIC will fit within the EOSC ecosystem by sharing its accumulated knowledge, tools and data with other RIs (e.g., DARIAH and DiSSCo), in collaboration with EOSC related projects (e.g., SSHOC, RDA working groups, EGI-ACE) and beyond. It follows the paradigm that the data will be "as open as possible, as closed as necessary".

Upon completion of the E-RIHS IP, discussion on the incorporation of data and meta-data in the E-RIHS ERIC will be held: further details will be incorporated into this and other documents of E-RHIS IP (e.g., DMP E-RIHS ERIC).

---

[3] Please note that this is only for E-RIHS IP: further information will be added here as appropriate. In WP 5.1.: DataCite (https://datacite.org/become.html) is currently explored for PIDs development and management.   A further DMP for E-RHIS ERIC is currently being developed.

[4] Kindly note the ongoing discussion on proposing an approach in the Zenodo metadata schema for Cultural Heritage datasets (Bucciero *et al*., 2023).

Interoperability: Making data interoperable

All data and meta-data within E-RHIS IP will be developed and agreed within standard format. All the choices and decisions will be documented and shared in open and accessible platforms. E-RIHS IP aims at attaining interoperability that is both human and machine-readable: formats will be chosen as appropriate to achieve this.   Such process will be taken into account also towards the development of the planned user-friendly Catalogue of Services for the ensuing E-RIHS ERIC.

It is important to recall here that that the E-RIHS IP WP 5.1 address the issue of data and meta-data vocabularies and ongoing work in the E-RIHS community.   Concerning E-RIHS Heritage Science Interoperability, see current literature (e.g., Padfield *et al*., 2022).

Re-usability: Increasing data re-use

The E-RIHS IP will work towards optimising data re-use, by exploiting an accessible, searchable, open, rich, and well-described meta-data. In this context, both data and meta-data will be released with a defined and documented re-use licence.   To ensure that the data produced in "as open as possible and as closed as necessary" manner.

If the actual data are embargoed for an agreed period of time or if certain data-sets will not be made public (where appropriate or legally required), these issues shall be clearly documented in the meta-data which – all the meta-data produced by E-RHIS IP – may be licensed under an open Creative Commons licence34, or an agreed equivalent licence (preferably CC-0 or CC-BY).

In this context, building on E-RIHS PP practices, the E-RIHS IP will continue its engagement within the governance of DARIAH towards the development of the DARIAH Heritage Data Reuse Charter (https://datacharter.hypotheses.org/), the Lexicographic Data Seal of Compliance (see Romary et al., 2021), mechanisms to which heritage science practitioners should adhere as well as the adoption of above-mentioned the DARIAH Principles (principles of trustworthiness, interoperability, stewardship, citability, reciprocity and openness) which are designed as complementary principles to the FAIR Principles.

To ensure the long-term sustainability of the E-RIHS IP datasets, the Project will select a repository which have a well documented sustainability plan (e.g., Zenodo).

3. OTHER RESEARCH OUTPUTS

As mentioned above in the Introduction, all the E-RIHS IP deliverables and outputs (e.g., project presentations, procedural work protocols, communications, discussions and the project public presentations in social media and the project website) will adhere to Open Science Practices, be allocated a PID, follow the FAIR principles to the extent possible, and will also abide to the GDPR.

As underlined in the Grant Agreement (GA), it is to be recalled that all E-RIHS IP deliverables will be public by default. However, E-RIHS IP will abide by the decisions of the E-RIHS iGA about the dissemination level or the embargo of deliverables that concern the strategic interests of the ERIC.

4. ALLOCATION OF RESOURCES

Responsibility for ensuring that E-RIHS IP data and meta-data is compliant with the FAIR Principles will rests primarily with the Data Protection Officer of the Project Coordinating Institution, namely National Research

Council (CNR).[5] The Steering Committee of E-RIHS IP will be consulted in case of this DPM have to be revised and updated according to project specifications and development.

5. DATA SECURITY

Management of data and meta-data will remain the responsibility of E-RIHS IP participating content providers. E-RIHS IP will follow best-practice to ensure the integrity of such objects are secure from accidental misuse or malicious tampering and will ensure that personal data is protected according to EU regulations including the General Data Protection Regulation (GDPR) (see below). Towards preventing misusing of data and metadata, the E-RIHS IP will implement:

- Two-factor authentication for infrastructure administration tasks.
- Application of PID for identification of providers and users of data and meta-data.
- Ensure any stored passwords are hashed with modern cryptographic algorithms.
- Frequently update IT resources to ensure that they run the latest security patches and that they will allow for components to be rapidly and reliably redeployed in the event of a security breach.

6. ETHICS

E-RIHS IP community is fully committed to the principles of responsible science that includes research ethics, data management compliance with the open science and FAIR principles, gender equality and non-discrimination and the implementation of GDPR requirements. The operationalisation of the E-RIHS IP project is not expected to raise any particular ethical issues; however, E-RIHS IP Steering Committee will always control and guarantee that all collected and processed personal data will be treated following the ethics provisions set out in the Grant Agreement, and notably the highest ethical standards and the applicable international, EU and national laws. Approval from local/national authorities in charge of data protection will be secured, if requested. Personal data that will be collected on questionnaires for user and provider evaluation and event's organization, the processing of personal data and the protection of privacy in recorded information (audio and/or visual) and electronic communications, as well as in public communication networks will fully comply with General Data Protection Regulation (GDPR), Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. Finally, E-RIHS IP fully complies with the "do not significant harm" principle as per Art. 17 of Regulation (EU) No 2020/852. A "Privacy and Cookies policy" - already in place in the framework of the E-RIHS community[6] - will be fine-tuned accordingly.

---

[5] See: Nomina del Responsabile della Protezione dei Dati personali (RDP) ai sensi dell'art. 37 del Regolamento UE 2016/679. Retrieved from   https://www.cnr.it/sites/default/files/public/Provv_103_2020_nomina_Rpd.pdf (accessed on 5/5/2023).

[6] Reference is made to https://www.iperionhs.eu/wp-content/uploads/2021/03/IPERION-HS-E-RIHS_privacy-policy-e-cookies_2021.01.11.pdf   (accessed on 5/5/2023).

If considered appropriate, an ethical independent advisory committee will be appointed to monitor emerging issues.

REFERENCES

Benassi, Laura, Striova, Jana, Castillejo, Marta, Carmona, Paula, & Padfield, Joseph. (2023). E-RIHS IP Dissemination, Exploitation and Communication (1.0). *Zenodo.* https://doi.org/10.5281/zenodo.7849261

Bucciero, A., Demetrescu, E., Fanini, B., Chirivì A. & Taurino, F. (2023). An approach to extend the metadata schema of Zenodo for Cultural Heritage datasets. *Research Ideas and Outcomes 9:* e93859. DOI: 10.3897/rio.9.e93859

E-RIHS ERIC Data Management Plan (in preparation)

Franken, J, Birukou, A, Eckert, K, Fahl, W, Hauschke, C and Lange, C. (2022). Persistent Identification for Conferences. *Data Science Journal*, 21: 11, pp. 1–18. DOI: https://doi.org/10.5334/dsj-2022-011

Padfield, Joseph. (2021). *D6.1 Data Management Plan for IPERION HS* (1.0). Zenodo. DOI: https://doi.org/10.5281/zenodo.4541267

Padfield, J., Sotiropoulou, S., Schmidle, W., and W. Fremout (2022). E-RIHS Heritage Science Interoperability. *Zenodo*: 18 October 2022. DOI: https://doi.org/10.5281/zenodo.7221156