



Project information

Project full title	Research Infrastructure Access in NANoscience & nanotechnology
Project acronym	RIANA
Grant agreement no.	101130652
Instrument	Research and Innovation Action (RIA)
Duration	01/03/2024 – 29/02/2028
Website	riana-project.eu

Deliverable information

Deliverable no.	D1.2
Deliverable title	Data Management Plan
Deliverable responsible	DESY
Related Work-Package/Task	WP1
Type (e.g. Report; other)	DMP
Author(s)	Tom Minniberger, Michael Stuckelberger
Author(s) affiliation	DESY
Dissemination level	Public
Document Version	1
Date	31. July 2024
Download page	https://zenodo.org/records/13287645

Document information

Version no.	Date	Author(s)	Comment
0.1	06.06.2024	Tom Minniberger	Draft
0.2	29.07.2024	Michael Stuckelberger	Draft
0.3	05.08.2024	Tom Minniberger	Revision
0.4	09.08.2024	Michael Stuckelberger	Finalization





Table of Content

Project information	1
Deliverable information	1
Document information	1
Abstract	3
Description	3
Data summary	3
Making data findable.....	4
Making data openly accessible.....	5
Making data interoperable.....	5
Making data re-useable.....	5
Allocation of resources.....	5
Data security.....	5
Ethical aspects	6



Abstract

This Data Management Plan (DMP) outlines the data management life cycle for the data to be collected, processed, and generated by the RIANA project. It is intended to be a living document that will be updated as needed upon project progress when significant changes occur. This DMP will specify the implementation of the FAIR (findable, accessible, interoperable, and reusable) data principles and ensure compliance of data formats and related metadata with the requirements of the European Open Science Cloud (EOSC).

Description

One of the objectives of RIANA is to promote FAIR data within Analytical Research Infrastructures. This includes harmonizing and standardizing the use of Data Management Plans (DMPs) for experiments conducted at RIANA infrastructures. Our approach to managing project data serves as a model of FAIRness and transparency. Therefore, we intend to publish all major deliverables in the RIANA community on Zenodo (<https://zenodo.org/>) and through the CORDIS platform provided by the European Commission, ensuring access to all documents with a unique Digital Object Identifier (DOI).

Survey and personal data will be handled in compliance with the General Data Protection Regulation (GDPR), with oversight from an independent ethics advisor to ensure adherence.

This document is heavily inspired by the DMP document created for the ReMade-at-ARI project (<https://remade-project.eu/>). It follows the EC – Horizon Europe DMP template provided in the EC portal: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/template/report/data-management-plan_he_en.docx

Data summary

RIANA offers services to scientists who create or process scientific data. As a result, the project will produce various outputs such as documents (including texts, presentations, and videos), websites, demonstrators, and source code. It will also involve generating simulated data, reuse existing data sets, and generate new scientific data.

Data exchange within the RIANA consortium (meeting agendas and minutes, training material, dissemination material, etc.) is realised via DESY Sync&Share (cloud service). In order to exchange internal data securely, an environment for all RIANA partners was created in DESY Sync&Share (hereinafter referred to as RIANA Cloud). The RIANA Cloud is only accessible to the project partners. With the support of Helmholtz Federated IT Services (HIFIS), we have enabled secure yet simple access to the RIANA Cloud. Thanks to the technical solution of the Helmholtz Authentication and Authorisation Infrastructure (AAI), RIANA partners can log in to the RIANA cloud using their existing logins from their home institutes or with existing access from ORCID, GitHub, or Google accounts. In this way, we were able to provide secure, individualized, and easy access without the scientists having to create another new account.

Table 1: Generated and collected data in RIANA

Data type	Storage place	Expected size	Contents
Text documents	RIANA cloud RIANA webpage Scientific journals Zenodo	100s Megabytes	Deliverables Meeting minutes Documentation Publications
Videos	RIANA cloud RIANA webpage Youtube Zenodo	100s Gigabytes	Training material Dissemination material Conference contributions
Source code	GitHub Zenodo RIANA infrastructures	100s Megabytes	Extension & integration of existing catalogues and services
Experiment datasets	RIANA infrastructures	10s Terabytes	Data from measurements

All newly generated scientific data will be handled in line with the Data Management Plans (DMPs) of the partners.

For software, most of the developments within RIANA will build upon existing software, and we will ensure compliance with the existing licenses for these software tools.

Making data findable

If a deliverable is deemed to have a broader scope and be of interest beyond the RIANA Consortium, the Executive Board decides about making it publicly available. Approved deliverables will be uploaded to Zenodo, assigned a DOI, and made searchable via the Zenodo platform.

Where feasible, experimental datasets will adhere to metadata standards consistent with PaNOSC and ExPaNDS, as outlined in the proposal. When these datasets are made openly accessible, they should be referenced using persistent identifiers generated by the facilities. Each facility will be encouraged to provide DOIs for the data they produce to ensure easy traceability.

To enable detailed tracking of data produced within RIANA, versioning will be applied as follows:

- The data set versions will follow the data policy of host infrastructures.
- Source code will use the Zenodo or GitHub release management if compatible with the policies of an existing project.

The combined use of GitHub and Zenodo to strictly version the different releases of software is a good practice that will be promoted at the different facilities of RIANA.

Making data openly accessible

All public project deliverables will be uploaded to Zenodo where RIANA has created a community site. Only a few deliverables will not be public because they contain sensitive data (Collaboration Agreement, ethics and human resources allocation). Deliverables containing sensitive information, intended only for the consortium and the Commission services, will be hosted exclusively on the RIANA cloud and by the Commission services with access through ascertained identity using credentials.

For the project documents, only widely used or open source software will be needed for access. Source codes and git repositories will be accessible via a web browser. Open datasets will be accessible through GitHub, Zenodo, and the repositories of the research infrastructures.

There is no restriction on use for all public RIANA deliverables, which will be under Creative Commons licenses.

Making data interoperable

RIANA provides services for scientists who generate or process scientific data. It will therefore generate documents (text, presentation, video), websites, demonstrators and source code. It will also generate simulated or re-use existing data sets and generate new scientific data that shall be disseminated through scientific publications that are encouraged to be open access. The project deliverables are shared and can be re-used by other related projects for efficient collaboration and good practices.

Making data re-useable

The use of Creative Commons licenses (CC) BY 4.0 for all public project deliverables will be implemented. Concerning the source code for software development, we encourage to choose a Creative Common license.

Depending on their origin, datasets will be made available as soon as possible according to the partner facility data policy. Public deliverables will be made available as soon as they are published on the EU portal. There is no restriction on the re-use of RIANA's public deliverables after the end of the project.

Allocation of resources

Zenodo is hosted by CERN and paid for by the EU. GitHub is a public resource. Human resources dedicated to FAIR in RIANA are included in its Project Grant.

WP leaders are responsible for the data management in their work packages. WP1 is responsible for the general compliance of the data management with respect to this DMP.

Data security

Data security for all documents uploaded to Zenodo naturally relies on Zenodo's security measures. Documents for long-term storage will be in PDF/A format (en.wikipedia.org/wiki/PDF/A). Data security for confidential deliverables relies on DESY's nextcloud infrastructure.

For the development phases, data security relies on each partner facility's IT infrastructure.

Ethical aspects

The handling of the Protection of Personal Data (POPD) requirements in RIANA will strictly follow GDPR. An external ethics advisor will monitor compliance with these regulations and issue corresponding reports twice during the project running time.

Quantitative and qualitative data collection (interviews & surveys) are made in the framework of RIANA. Measures are taken to prevent conclusions being drawn about individual participants. These include commonly employed methods of anonymization, pseudonymization, and aggregation of data in accordance with the guidelines of the German Research Foundation (DFG) and the ethical codes of German and European learned societies (Deutsche Gesellschaft für Soziologie, Deutsche Vereinigung für Politikwissenschaft, European Consortium for Political Research).

The same guidelines of good scientific practice stipulate that the raw data of publications should be kept in their anonymized form for 10 years to ensure the reproducibility of results unless there are compelling (in particular ethical) reasons not to do so.

The informed consent of the participants in these surveys and interviews is a prerequisite for the use and publication of the research data. This consent is requested at the time of collection and can be revoked later. The data collection and analysis will be carried out in consultation with the external ethics advisor and internal DEI Officer. Personal data will be deleted after anonymisation or aggregation.

Other personal data, like the project's "who's who" document or the list of attendees to training, workshops and conference events, will be stored in the RIANA cloud with access restricted to the project's contributors.

Consent will be obtained before any public dissemination.