

101137192- AVITHRAPID

Antiviral Therapeutics for Rapid Response Against Pandemic Infectious
Diseases

WP7 FAIR DATA MANAGEMENT

D7.3 Resource dissemination into community

Lead contributor	1 – FRAUNHOFER
Other contributors	All consortium members

Data	Description
Due date	31-10-2025
Delivery date	27-10-2025
Delivery type	R
Dissemination level	PU

1 Document History

Version	Date	Description
V0.1	07-10-2025	First Draft – Leonie von Berlin
V0.2	22-10-2025	Revision – Maria Kuzikov
V0.3	25-10-2025	Revision – Björn Windshügel
V1.0	27-10-2025	Submitted version – Leonie von Berlin

2 Table of Contents

1	Document History.....	1
2	Table of Contents	1
3	Introduction.....	2
4	Scope of the deliverable	3
5	Approach.....	3

6	Results	3
6.1	Data resources and workflows	3
6.2	Data dissemination into community	5
6.3	Accordance with the FAIR principles	7
7	Conclusion and Outlook	8
8	Glossary of Terms	9

3 Introduction

[AVITHRAPID](#) (Antiviral Therapeutics for Rapid Response Against Pandemic Infectious Diseases) is a Horizon Europe research initiative that aims to strengthen Europe’s preparedness for future viral outbreaks by advancing the development of broad-spectrum antiviral therapeutics. The project focuses on repurposing and optimising small-molecule compounds with demonstrated bioactivity against multiple viral targets. By combining *in silico* modelling, biochemical and cell-based assays, and pre-clinical validation, AVITHRAPID seeks to establish a pipeline of early-stage antiviral drug candidates that can be rapidly mobilised in the event of emerging infectious diseases with pandemic potential. Together, these objectives position AVITHRAPID as a central contributor to Europe’s efforts to improve resilience against viral threats.

Given the scale and interdisciplinarity of the project, robust data management is indispensable. Data are generated across multiple experimental platforms and computational environments, each employing different formats, software tools, and metadata conventions. Without harmonised and transparent data handling, critical information could remain siloed, reducing reproducibility and long-term impact. To ensure that the project’s scientific outputs remain reusable and interoperable—both within the consortium and by the wider research community—AVITHRAPID has established a dedicated work package (WP) focused on FAIR Data Management.

WP7, FAIR Data Management, ensures that data generated within AVITHRAPID are managed according to the FAIR principles—Findable, Accessible, Interoperable, and Reusable. WP7 provides the practical and technical foundation for data capture, documentation, storage, processing, sharing, and long-term accessibility across the consortium. Operationally, WP7 defines data management guidelines, selects and implements tools that meet operational and FAIR standards, establishes infrastructure to support data aggregation and exchange, and ensures interoperability and use of common standards when collaborating inside and outside the consortium. These activities are executed by WP7.

The activities described in this deliverable build directly on this foundation. Deliverable 7.3 (“Resource dissemination into community”) documents the resulting workflows and resources and explains how these are disseminated.

4 Scope of the deliverable

Deliverable (D) 7.3, “Resource dissemination into community,” describes the workflows and resources developed under WP7 to enable structured data collection, processing, analysis, and modelisation within AVITHRAPID. It outlines how these workflows support FAIR data management and how the resulting outputs are made accessible to the wider scientific community.

The deliverable serves two main purposes. First, it documents the practical implementation of WP7 activities, presenting the tools and workflows that operationalise the project’s FAIR data management strategy. Second, it demonstrates how these resources are being disseminated to ensure transparency, reproducibility, and reuse.

5 Approach

The development of this deliverable builds on the broader FAIR data management activities conducted under WP7. The information presented here was compiled through an iterative process that combined documentation review, partner interaction, and practical implementation.

A central mechanism for gathering input was the Data Champion Group, which serves as the main interface between WP7 and the scientific work packages. Through regular exchanges every month, this group gained insights into the partners’ data management needs, practices, and challenges, enabling WP7 to design workflows and dissemination strategies that address real project requirements.

Another key element in this process was the AVITHRAPID Data Management Plan (DMP), which both guided and evolved alongside the implementation of FAIR data management. The DMP defines the overall framework for data handling across the consortium but is continuously updated to reflect new workflows, tools, and solutions implemented. The project’s data catalogue further complemented this process by providing a structured overview of available datasets and their characteristics.

Together, these inputs and interactions formed the foundation for identifying and harmonising data workflows, aligning them with FAIR principles, and developing the dissemination approaches described in this deliverable.

6 Results

6.1 Data resources and workflows

AVITHRAPID relies on several complementary digital resources and coordination mechanisms that together enable FAIR data management across the consortium. Experimental and analytical data are generated in partner laboratories and recorded within a shared Electronic Laboratory Notebook (ELN) platform (Figure 1). The ELN ensures structured capture of experimental information, metadata and provenance and follows a project-wide file-naming convention that links each entry to the corresponding dataset stored on the consortium server. The ELN runs on a separate secure server, while data intended for storage and sharing are exported in structured formats and transferred to the consortium’s shared storage.

For operational data management, partners use standardised data sheets, mostly in Excel format, which serve as templates for recording and transferring datasets. These sheets are uploaded to the Fraunhofer ownCloud for collaborative access. The ownCloud environment acts as the secure central platform for data storage and exchange, with files organised by work package and task. This structure ensures that raw and processed data as well as associated documentation can be found, reused and updated in a controlled way. Project members have been instructed not to store data on local computers or removable media but to rely on this central environment, which allows encrypted access and strict user-rights management in line with the project's data-security policy.

New datasets, along with their types and formats, are continuously registered in the AVITHRAPID data catalogue, which provides a structured overview of the consortium's data landscape. The catalogue serves as an entry point for identifying available data resources and also as a feedback

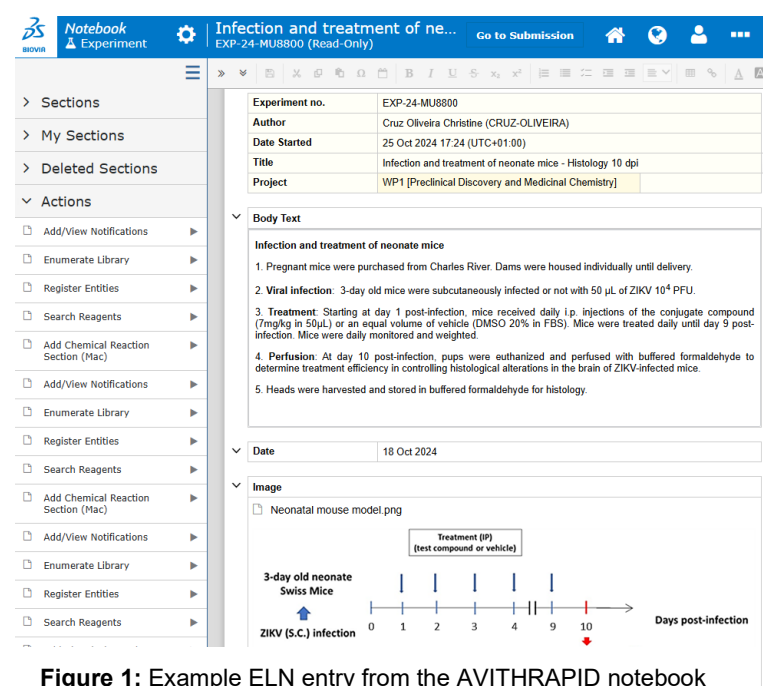


Figure 1: Example ELN entry from the AVITHRAPID notebook

mechanism for maintaining the DMP. As new data types emerge or formats evolve, they are incorporated into the catalogue and reflected in subsequent DMP updates, ensuring that the plan remains a living document aligned with project realities.

The Data Champion Group, coordinated by WP7, provides an accessible forum for partners to discuss data-related questions, exchange experiences and receive guidance on FAIR data management. The group supports partners in applying the principles described in the DMP, offers a channel for identifying common challenges and contributes to keeping the DMP and related

processes up to date.

To ensure consistent implementation across partners, project members are onboarded to the data-management tools and dissemination procedures at the start of their involvement. A set of manuals is available on our internal ownCloud (and partially on Zenodo), describing in detail how to use the ELN, how to upload datasets to the ownCloud and how to prepare and deposit publications, datasets and models in dissemination platforms like Zenodo or ModelArchive. WP7 complements these materials with practical assistance by offering one-on-one support sessions, email consultations and feedback on data documentation and deposition.

Data prepared through these processes are subsequently transferred from the internal storage environment to the project's dissemination platforms. The following chapters describe in more detail how AVITHRAPID ensures that these outputs are made accessible to the wider scientific community in line with the FAIR principles. The current data flow within AVITHRAPID is already functional but continues to evolve toward greater harmonisation.

6.2 Data dissemination into community

The overarching goal of dissemination in AVITHRAPID is to make the project's outputs openly available and reusable for the scientific community and, where relevant, for the wider public. By ensuring that data, models, and results are FAIR, the project contributes to transparency, accelerates scientific progress, and supports the rapid utilisation of research outcomes for pandemic preparedness and antiviral discovery.

Scientific publications form a primary channel of dissemination and are published in open-access journals, making them permanently and freely accessible and citable through persistent digital object identifiers (DOIs). In addition to these journal publications, AVITHRAPID uses Zenodo as its main repository for openly sharing research outputs. In the [AVITHRAPID Zenodo community](#) (Figure 2), publications are directly uploaded to increase their findability and are complemented by associated materials such as standard operating procedures (SOPs), datasets, and posters, providing full contextual information for each result. The community currently hosts seven entries, a number that will continue to grow as further project outputs become available. All materials deposited in Zenodo receive DOIs, which guarantee long-term findability and proper citation, and make the project's outputs accessible to anyone without institutional barriers.



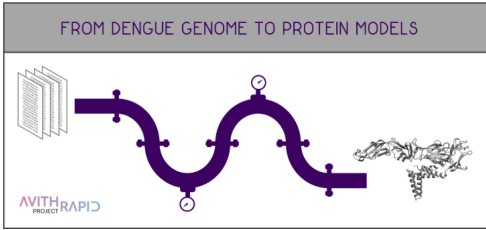
Figure 2: The AVITHRAPID Zenodo community

Beyond Zenodo, several domain-specific repositories are used to ensure data reach their appropriate scientific audiences. ChEMBL is used for viral and biochemical assay data, the Protein Data Bank (PDB) for crystallographic structures, and [ModelArchive](#) (Figure 3) for computational structure models. These repositories are well-established within their respective communities and provide rich metadata, versioning, and open-licensing frameworks, ensuring that the uploaded material can be integrated into broader scientific workflows. A project-level [GitHub organisation](#) (Figure 4) complements these repositories for hosting analysis scripts, workflow templates, and model-generation code. As an open platform with version control, GitHub enables transparency and the exact repetition of analytical steps, which strengthens reproducibility and reusability.

Modeling Dengue Virus Proteins: A Strategic Approach for Enhanced Pandemic Preparedness

DOI: (not active) 10.5452/ma-denv

Overview



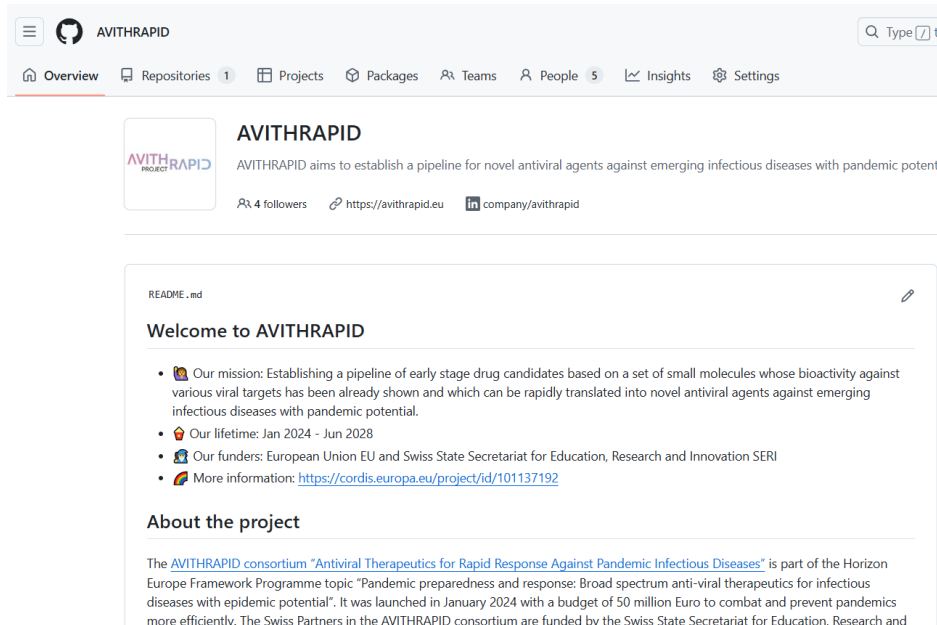
Abstract Throughout history, humanity has faced multiple viral threats, with science playing a key role in survival. Despite the development of new innovative vaccines and drugs, the constantly mutating and evolving pathogens impose continuous challenges. In addition, climatic variations significantly influence the dynamics of natural selection. Some tropical countries face major dengue outbreaks that return cyclically. However, what was once an exclusively topical problem is now spreading to several other regions of the planet: autochthonous cases of dengue have been reported in several European countries. This highlights the urgent need for advanced scientific tools and techniques to better understand these pathogens in time to prevent a new pandemic.

We believe that harnessing the rapid growth of computing technology is essential to stay ahead of the inevitable threats of infectious diseases. This project aims to establish a solid foundation for understanding the proteins of viral pathogens, aiming to improve preparedness and response capabilities.

Through the development of a pipeline, we modelled most of the proteins of the four serotypes of DENV using the SWISS-MODEL platform, to identify and map the viral target proteins suitable for drug development.

Figure 3: A ModelArchive entry from AVITHRAPID consortium member Ingrid Guarnetti Prandi

While dissemination primarily targets open-access and community repositories, the Fraunhofer ownCloud also plays a role in securely exchanging data with collaborators and external partners. It provides controlled access for sharing pre-publication datasets or sensitive materials under specific agreements, bridging the gap between internal work and open dissemination.



The screenshot shows the AVITHRAPID GitHub organization page. At the top, there is a navigation bar with 'Overview', 'Repositories 1', 'Projects', 'Packages', 'Teams', 'People 5', 'Insights', and 'Settings'. Below this is the organization's profile, including the AVITHRAPID logo, a description: 'AVITHRAPID aims to establish a pipeline for novel antiviral agents against emerging infectious diseases with pandemic potential', and social media links for 4 followers, a website, and a LinkedIn company page. The main content area displays a 'README.md' file with the following text:

Welcome to AVITHRAPID

- Our mission: Establishing a pipeline of early stage drug candidates based on a set of small molecules whose bioactivity against various viral targets has been already shown and which can be rapidly translated into novel antiviral agents against emerging infectious diseases with pandemic potential.
- Our lifetime: Jan 2024 - Jun 2028
- Our funders: European Union EU and Swiss State Secretariat for Education, Research and Innovation SERI
- More information: <https://cordis.europa.eu/project/id/101137192>

About the project

The AVITHRAPID consortium "Antiviral Therapeutics for Rapid Response Against Pandemic Infectious Diseases" is part of the Horizon Europe Framework Programme topic "Pandemic preparedness and response: Broad spectrum anti-viral therapeutics for infectious diseases with epidemic potential". It was launched in January 2024 with a budget of 50 million Euro to combat and prevent pandemics more efficiently. The Swiss Partners in the AVITHRAPID consortium are funded by the Swiss State Secretariat for Education, Research and

Figure 4: The AVITHRAPID GitHub organisation

Together, these environments form a coherent and transparent data flow that spans the full lifecycle of research within AVITHRAPID—from experimental generation and secure storage to analysis, publication, and long-term dissemination. Although further harmonisation and automation are planned, the existing resources already ensure that project data, models, and publications are captured, curated, and shared in a FAIR-compliant and reproducible manner.

6.3 Accordance with the FAIR principles

The FAIR principles—Findable, Accessible, Interoperable, and Reusable—form the conceptual and practical framework guiding all data management and dissemination activities in AVITHRAPID. They

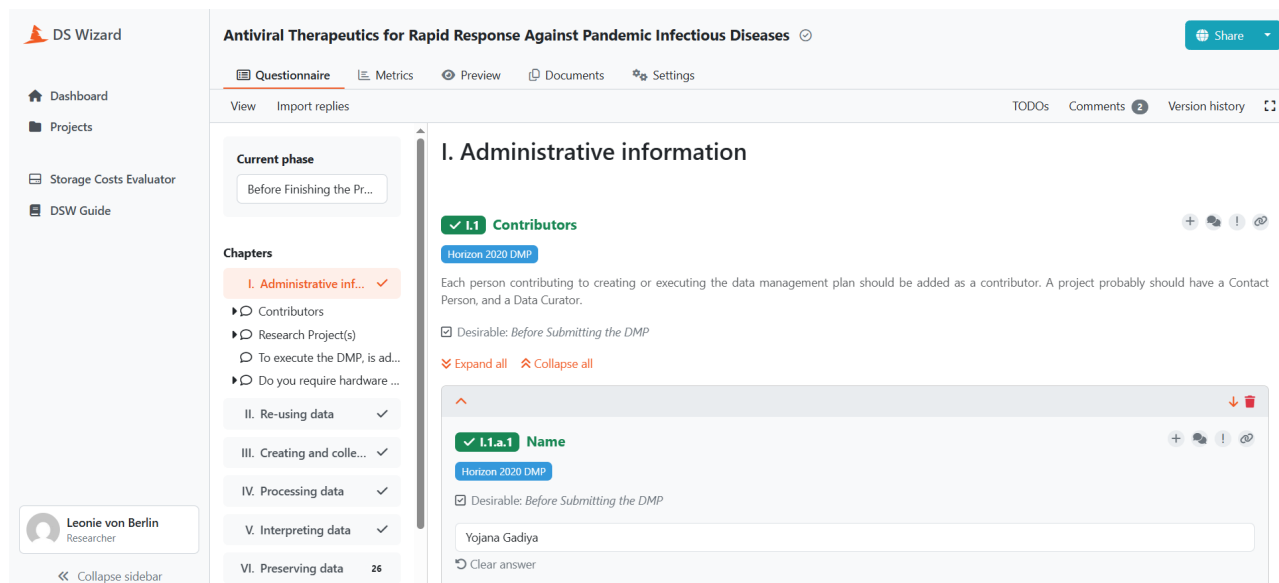


Figure 5: The Data Stewardship Wizard to create the AVITHRAPID DMP

ensure that research outputs are not only preserved but also discoverable, understandable, and reusable to accelerate research and improve pandemic preparedness. Within WP7, these principles are embedded throughout the data lifecycle, from the initial documentation of experiments in the ELN to the open publication of datasets, code, and models in public repositories. To operationalise these principles, AVITHRAPID used the [Data Stewardship Wizard](#) (DSW, Figure 5) to create its FAIR-compliant Data Management Plan. The DSW provides a structured and standardised framework for describing datasets, metadata, storage, and sharing strategies, ensuring alignment with European Open Science Cloud (EOSC) and Horizon Europe requirements. Furthermore, the project applies the [FAIR Dataset Maturity Assessment Tool](#) (FAIR-DSM) to evaluate and improve the FAIRness of individual datasets, enabling targeted optimisation of FAIR implementation.

Findable

AVITHRAPID ensures that all public outputs can be easily located and cited through the assignment of DOIs via Zenodo, open access journals and other repositories. Metadata accompanying datasets and models describe the content, provenance, and context in a structured way, following repository-specific standards. The data catalogue serves as an internal indexing system that lists datasets, formats, and locations.

Accessible

All publications are released in open access journals, and datasets, SOPs, and models are made available through trusted repositories such as Zenodo, ChEMBL, PDB, and ModelArchive. Repository licenses and access conditions follow the principle of “as open as possible, as closed as necessary”, allowing temporary restrictions only when justified by intellectual property or confidentiality requirements. Metadata remain openly available even when data access gets restricted, and the secure ownCloud environment supports controlled access for internal and external collaboration.

Interoperable

To support exchange and integration, the project uses community-recognised data formats, for example comma-separated values (CSV), Tagged Image File Format (TIFF), Portable Document Format (PDF) and standards such as Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE), Minimum information about a proteomics experiment - Mass Spectrometry Informatics (MIAPE-MSI), and MicroArray Gene Expression Tabular Format (MAGE-TAB). Controlled identifiers like PDB Identifiers (IDs) link datasets to external databases and the usage of SMILES (Simplified Molecular Input Line Entry System) for describing the structure of molecules that can be read by most molecule editors allows sophisticated search procedures in databases.

Reusable

AVITHRAPID promotes reusability through comprehensive documentation, version control, and transparent provenance tracking. SOPs accompany datasets to describe experimental methods and analysis steps, while GitHub repositories preserve the full history of scripts and models. Public outputs are distributed under open licenses that allow reuse, and the chosen dissemination platforms guarantee long-term availability beyond the project duration.

Together, these measures ensure that all data handling and dissemination activities within AVITHRAPID are aligned with the FAIR principles and Horizon Europe's open science requirements. While some processes are still being refined, the established infrastructure, use of the Data Stewardship Wizard, and continuous FAIR-DSM assessments already provide a robust foundation for long-term accessibility, transparency, and reuse of the project's outputs. Continuous improvement of these workflows under WP7 will further enhance the project's contribution to a FAIR and sustainable European research ecosystem.

7 Conclusion and Outlook

The activities described in this deliverable demonstrate that AVITHRAPID has established a solid foundation for FAIR data management and open dissemination across the consortium. Within WP7, key infrastructures such as the ELN, the Fraunhofer ownCloud, the data catalogue, and domain-specific repositories have been implemented and connected to form a functional data workflow. Complementary measures such as standardised data sheets, onboarding materials, and the creation of the Data Champion Group have strengthened the operational framework and fostered a shared understanding of data-handling procedures among partners.

While some aspects still require fine-tuning, the project already ensures that data, models, and publications are traceable, securely stored, and publicly accessible. The use of the Data Stewardship Wizard for maintaining the Data Management Plan and the FAIR-DSM assessments for evaluating dataset quality support a process of continuous improvement and alignment with the FAIR and Open Science principles.

Building on this foundation, the next steps will focus on consolidating and refining the practical implementation of these principles, beginning at the roots of data recording. The ELNs play a central role here: assessing the FAIRness of ELN entries will help ensure that experimental data and metadata are structured, complete, and ready for integration. As part of this process, WP7 will explore the use of standardised terminology and public ontologies to harmonise the language used across ELN records and datasets, improving interoperability and consistency.

Another priority will be to explore developing Standard Operating Procedures (SOPs) derived from ELN records to formalise and share best practices across partners. These SOPs will not only document experimental processes but also support the reproducibility and comparability of results. Furthermore, the project will continue to refine code and data presentation workflows, ensuring that once scripts and models are deposited on GitHub, they are accompanied by clear documentation and can be easily linked to related datasets, publications, or dashboards. Such connections will enhance transparency, facilitate reuse, and demonstrate the coherence between computational and experimental components of the project.

Together, these ongoing developments will strengthen the FAIR data ecosystem within AVITHRAPID and ensure that the project's outputs remain sustainable, interoperable, and accessible long after its completion. WP7's work continues to serve as both an operational backbone for data management and a catalyst for open, collaborative research in the antiviral discovery community.

8 Glossary of Terms

Abbreviation	Full term
AVITHRAPID	Antiviral Therapeutics for Rapid Response Against Pandemic Infectious Diseases
CSV	Comma-separated values
D	Deliverable
DMP	Data Management Plan
DOI	Digital Object Identifier
DSM	DataSet Maturity assessment tool
DSW	Data Stewardship Wizard
ELN	Electronic Lab Notebook
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Reusable
ID	Identifier
MAGE-TAB	MicroArray Gene Expression TABular format
MIAPE-MSI	Minimum Information About a Proteomics Experiment - Mass Spectrometry Informatics
MIQE	Minimum Information for Publication of Quantitative real-time PCR Experiments
PDB	Protein Data Bank
PDF	Portable Document Format
SMILES	Simplified Molecular Input Line Entry System
SOP	Standard Operating Procedure
TIFF	Tagged Image File Format
WP	Work Package