



# Data management for Microbial Resources: Open Science perspectives in the infrastructure MIRRI-IT

Marco Beccuti<sup>1</sup>, Antonio d'Acierno<sup>2</sup>, Simone Donetti<sup>1</sup>, Sandro Gepiro Contaldo<sup>1</sup>,  
Paolo Romano<sup>3</sup>, Rossana Damiano<sup>1</sup>, Marco Aldinucci<sup>1</sup>, Giovanna Cristina Varese<sup>1</sup>

<sup>1</sup> University of Turin, IT

<sup>2</sup> ISA - CNR, Italy

<sup>3</sup> IRCCS Ospedale Policlinico San Martino, Genoa, Italy





# The SUS-MIRRI.IT project



- “Strengthening the MIRRI Italian Research Infrastructure for Sustainable Bioscience and Bioeconomy” ([www.sus-mirri.it](http://www.sus-mirri.it))
  - funded in 2022
  - 17M€
  - PNRR – Research Infrastructure
- It aims at providing a **single entry point** platform to the resources of the Italian microbial Biological Resource Centers databases, in order to strengthen the Italian Research Infrastructure for Microbial Resources (MIRRI-IT) and ensure its long-term sustainability,



# Online collaborative work environment



- Developing a **semantic data model** from existing initiatives and datasets at the Italian and European level to facilitate the collection of data from Italian collections
- Implementing a Harvard Dataverse instance to effectively store and deliver the semantically annotated microbial information according to the **FAIR paradigm**, supporting data sharing and discovery.



# Workflow and integration



- The project leverages the configurable system of **roles** and nested repositories made available by Dataverse to enforce clear and reliable **workflows** for data curation, preservation, and publishing
- It builds upon its integration with established platforms for the **identification of users**, such as ORCID, and of **digital objects**, such as DOI and FDO, to create an open, integrated environment



# 0001

**Draft** **Unpublished**



Contaldo, Sandro Gepiro, 2023, "0001", <https://doi.org/10.5072/FK2/DVEGCC>, Root, DRAFT VERSION 

Cite Dataset ▾

[Learn about Data Citation Standards.](#)

Publish Dataset

Edit Dataset ▾

Contact Owner

Share

Dataset Metrics 

0 Downloads 

## Description

Strain

## License/Data Use Agreement



CC0 1.0

Files

Metadata

Terms

Versions

 Add + Edit Metadata

This dataset has been configured to use **Mirri** as the language for all metadata entries.

Citation Metadata ▾

Strain 

<b>Strain ID</b> 	45660
<b>Collection</b> 	TUCC
<b>Accession Number</b> 	001





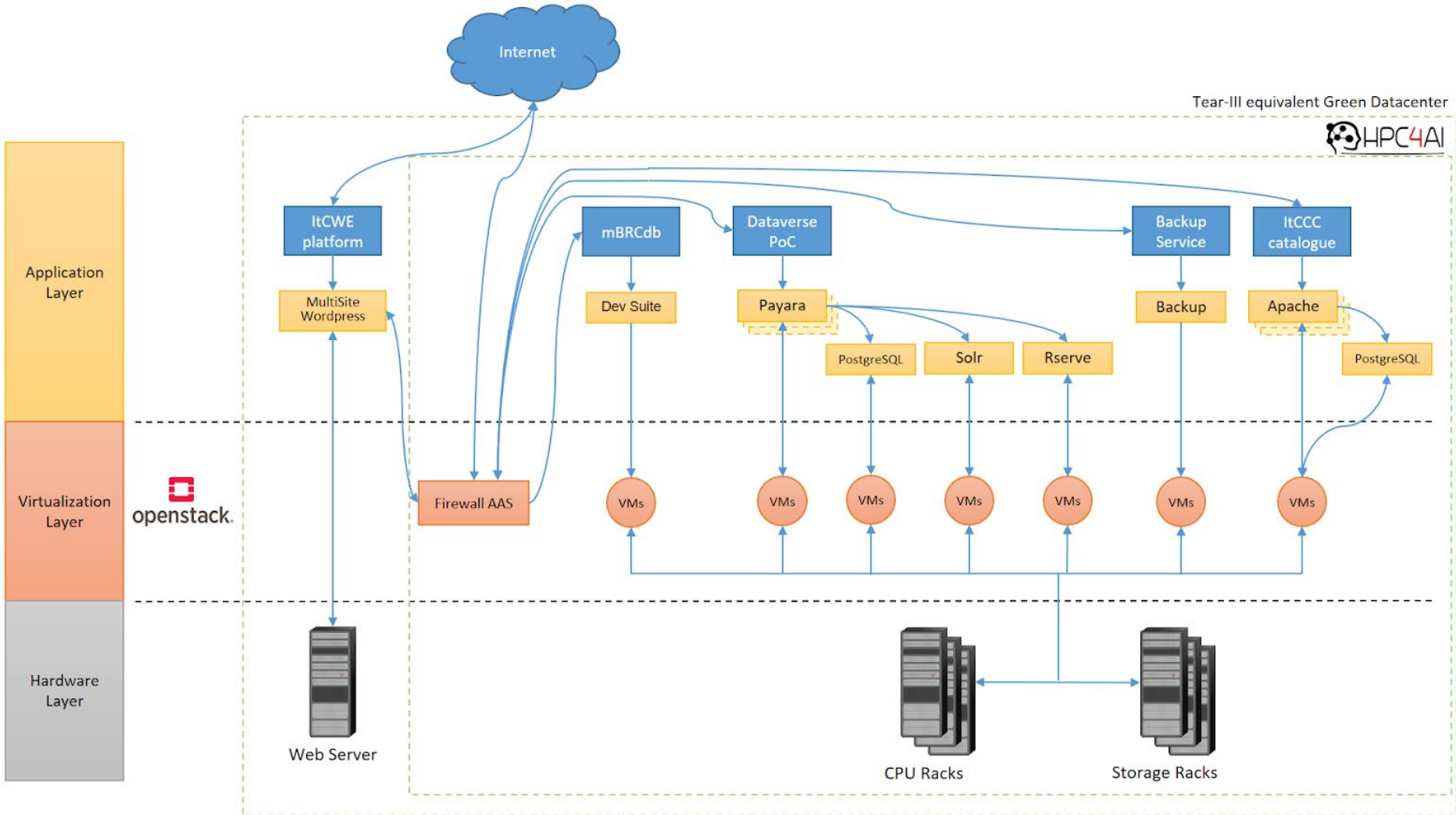
# Computational Infrastructure



- **Hardware level**, set up on the HPC4AI data centre (<https://hpc4ai.unito.it>)
- **Virtualization level**, based on the open-source OpenStack cloud technology (co-developed with Canonical - <https://canonical.com/>, the company behind Ubuntu)
- **Application level**, corresponding to the *macro-services* (e.g., PostgreSQL, Apache) that support the project *micro-services* (e.g., Wordpress, Dataverse).



# SUS-MIRRI.IT Computational Infrastructure





# Open access training in SUS-MIRRI project

- Open-access **courses** during the project will
  - present the methodologies and technologies developed during the project;
  - introduce theoretical and practical steps for opening up research practices and meet the expectations relating to openness
- An example of these is the "OSA2Micro – An Open Science Approach for Microbiology data integration" organized in collaboration with EOSC-life ([https://qbio.di.unito.it/cool\\_timeline/osa2micro/](https://qbio.di.unito.it/cool_timeline/osa2micro/))
- Moreover, we are working on developing **data analysis workflows** that will be compliant with the EU guidelines on open-access software.





# Cooperation with professional training at the University of Turin

SUS-MIRRI will be a case study in the **Professional Training Course for Data Stewards** organized by the University of Turin (40 ECTS, 2023-24)

- Objectives
  - supporting researchers in data management according to FAIR principles, in accordance with European policies and the principles of Open Science
  - managing platforms for data sharing and reproducibility of experiments according to FAIR principles and methods
- Modules:
  - **IT area**: training on IT architectures, protocols and representation paradigms; archival principles and models
  - **Ethical-legal and social area**: training on ethical aspects in data collection and sharing; legal aspects at national and community level
  - **Infrastructures**: Infrastructure and real-world case studies in the Humanities, Social Sciences, Life Sciences, and Sciences

# Thank you for your attention

<https://www.sus-mirri.it/>

<http://www.mirri-it.it/>

SUS-MIRRI.IT: [cristina.varese@unito.it](mailto:cristina.varese@unito.it),  
[marco.beccuti@unito.it](mailto:marco.beccuti@unito.it)

Professional Training Course for Data  
Stewards@Unito: [rossana.damiano@unito.it](mailto:rossana.damiano@unito.it)

