

# OPEN REPOSITORIES 2020

DOI [10.15161/oar.it/23688](https://doi.org/10.15161/oar.it/23688)

## The Reproducibility and Reusability Platform

R. Barbera<sup>[1,2]</sup>, R. Bruno<sup>[2]</sup>, M. Fargetta<sup>[2]</sup>, R. Rotondo<sup>[2]</sup>, A. Anagnostou<sup>[3]</sup>, S. J. E. Taylor<sup>[3]</sup>

[1] Department of Physics and Astronomy "E. Majorana" of the University of Catania - Italy, [2] Italian National Institute of Nuclear Physics, Division of Catania - Italy, [3] Brunel University London – United Kingdom

# Motivations and overview

For Open Science to become a common practice, its enabling technologies must demonstrate to be useful and easy to use. Building and executing software on distributed computing infrastructures (DCIs), with input data related to Open Access publications and coming from FAIR repositories, should hence be as easy as surfing the web. The Reproducibility and Reusability Platform (**RRP**) consists of several interconnected components addressing this.

- **DAMS**

[DOI 10.19161/oar.it/23465](#)

The [INFN Open Access Repository](#) (**OAR**) is based on [Invenio](#) and on an enhanced version of the code of [Zenodo](#). It exploits the concept of “communities”, which is central in the Zenodo architecture.

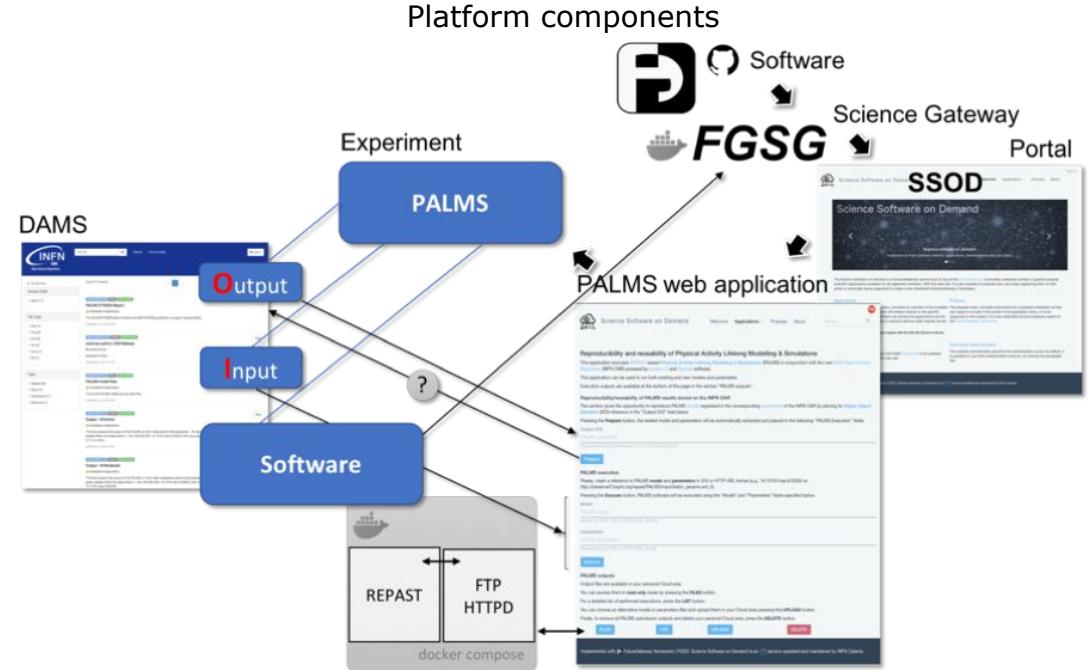
For the RRP, output and input datasets together with the necessary software are published with a registered DOI. Relationships among DOIs allow to reuse or reproduce the experiment results.

- **Experiment**

Physical Activity Lifelong Modelling & Simulations (**PALMS**) is an [agent-based](#) simulation application that predicts the lifelong physical activity behaviour of a population considering individual characteristics and their effect on physical activity over time. The model produces individual and aggregated quantitative outputs for quality of life and health conditions related costs.

- **FGSG + SSOD**

The [Future Gateway Science Gateway](#) (**FGSG**) is a virtualised environment capable to host one or more Science Gateways exploiting the capabilities of the [FutureGateway Framework](#). The [Science Software on Demand](#) (**SSOD**) is a science gateway running on top of the FGSG infrastructure.



INFN Open Access Repository

[INFN](#) adheres to the [PlanS](#) initiative since 2018 and the [INFN Open Access Repository](#) has born as a demonstrative platform to perceive the benefits of Open Access in Science.

DOI 10.15161/oar.it/23572

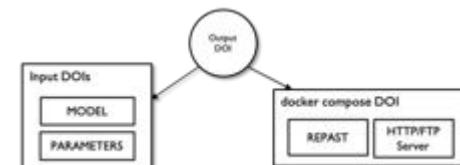
The repository is open for testing by all INFN staff, associated researchers and people from other organisations are free to use it.



**PALMS** related artifacts (*papers, reports, datasets, etc.*) are stored in the repository and each content is tagged with a [DataCite DOI](#) ensuring that all research outputs are citable, discoverable and, more importantly, linkable to each other to allow complete reproducibility and reusability.

The repository is also integrated with [GitHub](#) source code repository so that, input and output datasets together with the necessary software are accessible for reproducibility or using alternative input dataset obtain new results for reusability.

The key element in the repository for the RRP platform is the possibility to establish relationships among registered DOIs.

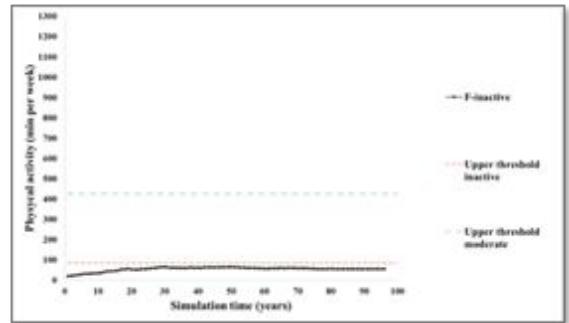
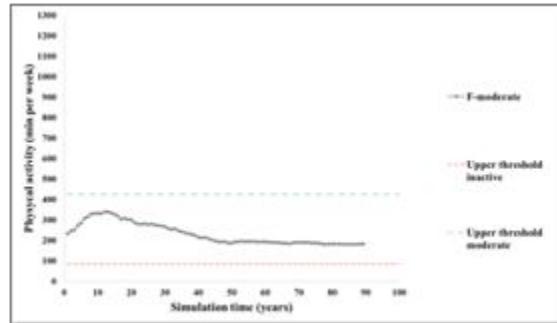
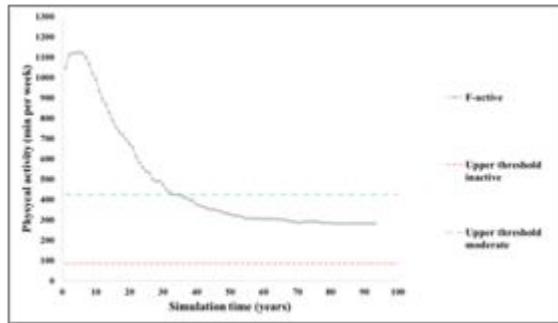


# Physical Activity Lifelong Modeling System

**PALMS** is a micro-simulation that predicts the lifelong physical activity behaviour of a population taking into account individual characteristics and their effect on physical activity over time. The model produces individual and aggregated quantitative outputs for quality of life and health conditions related costs.

The simulation software uses [REPAST](#) an open source agent-based modeling. For it a specific [docker-hub image](#) exists for PALMS executions. Each run requires two inputs, a model file for REPAST and a parameters' file. Both elements are registered into the INFN Open Access repository.

Graphs below represent the minutes spent for physical activity over time of three different kind of female profiles: Active, Moderate and Inactive.



# The FutureGateway Framework

The **FutureGateway** is an INFN [software project](#) aiming to build reliable and secure [Science Gateways](#).

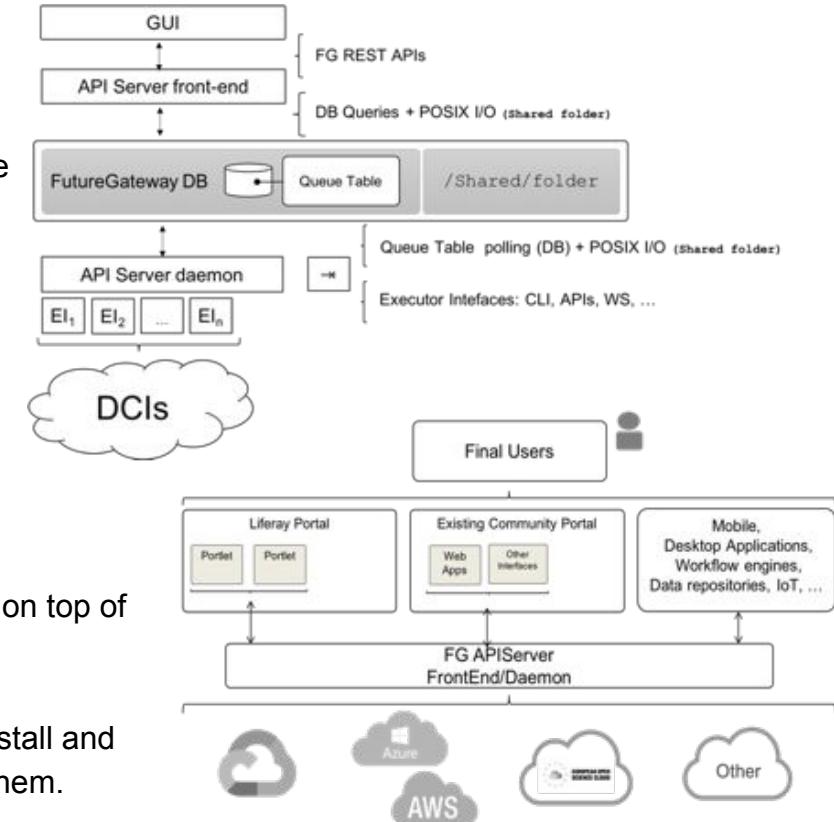
Science Gateways are today one of the most satisfying answers to execute software on top of distributed computing infrastructures by wide communities. This kind of platforms are normally requiring a non common set of experience and technical skills, often unavailable inside communities such as: Biomedicine, Chemistry, ...

FG is a mature product essentially providing a set of RESTful APIs to interact with one or more DCIs using a pluggable set of components called 'Executor Interfaces'.

The FG has three components:

- An API front-end to process incoming REST APIs from GUIs
- A database + a folder belonging to a sharable file system
- A daemon polling a queue table responsible to target execution on top of DCI using the correct EI

The FG is a complete framework, providing: APIs, scripts, tools to install and manage a Science Gateway and easily develop user interfaces for them.

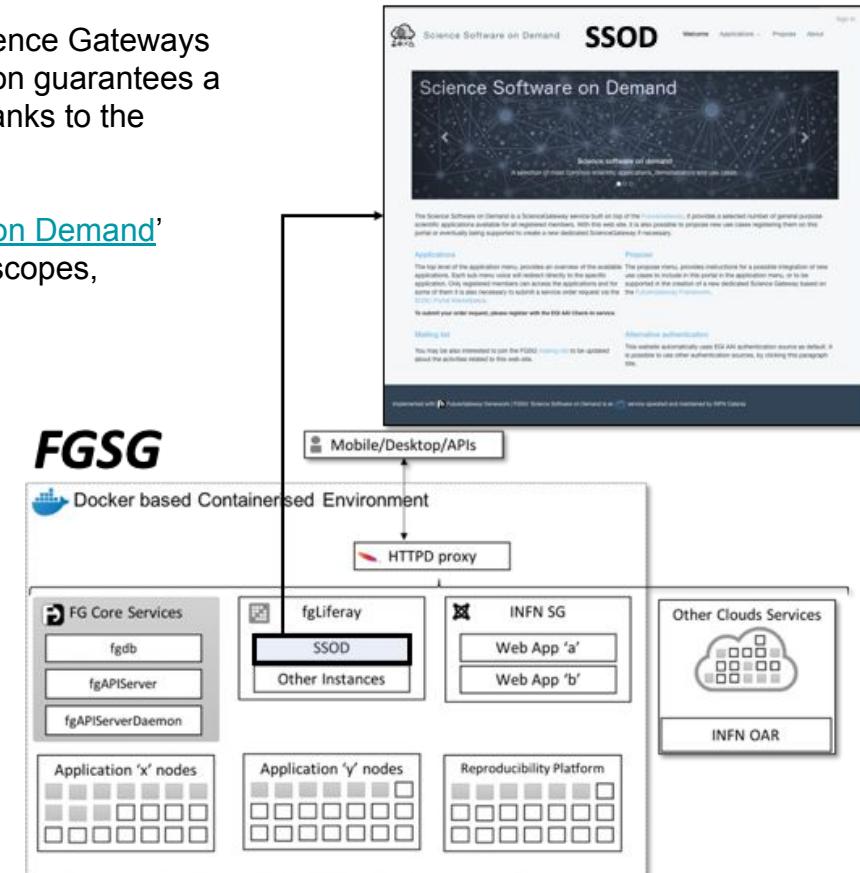
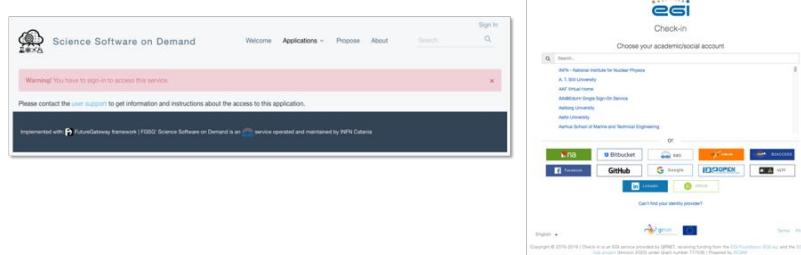


# The **FGSG** and the egi.eu **SSOD** service

The FGSG consists of a platform capable to host one or more Science Gateways using the [FGE](#) and totally based on Docker containers. This solution guarantees a rapid SG development and a secure and high scalable solution thanks to the adoption of Docker compose and Docker Swarm.

One of the available Science Gateways is the: '[Science Software on Demand](#)' (SSOD), a general purposes SG having principally disseminative scopes, developed in the context of the [EOSC-hub project](#).

The SSOD service is integrated with the [EGI](#) Authentication and Authorisation Infrastructure ([EGI Checkin](#)). Only authorised users can access the portal and use their applications.



# The RRP use case with PALMS

The SSOD has a section dedicated to the OpenScience' Reproducibility and Reusability platform that currently contains the PALMS experiment.

The application allows to retrieve the necessary inputs to reproduce a given output referenced in the form of a DOI.

1. To “**reproduce**” a given output, its DOI is ‘*solved*’ querying the INFN Open Access Repository, then the interface places the related input file references, still in the form of DOIs, in the execution input form below.
2. The execution input form accepts also HTTP based URLs to input files.
3. Once the input fields have been filled, it will be possible to execute the PALMS simulation reproducing the given output.
4. Generated output files will be available in a special area dedicated to the Science Gateway user allowing IO operations, thus opening the way to the “**reusability**”.

Science Software on Demand

Welcome Applications - Propose About

Reproducibility and reusability of Physical Activity Lifelong Modelling & Simulations

This application executes REPROD-based Physical Activity Lifelong Modelling & Simulations (PALMS) in conjunction with the new INFN Open Access Repository (INFN OAR) powered by Immortal v3 and Zenodo software.

This application can be used to run both existing and new models and parameters.

Execution outputs are available at the bottom of this page in the section "PALMS outputs".

Reproducibility/reusability of PALMS results stored on the INFN OAR

This section gives the opportunity to reproduce PALMS results registered in the corresponding community of the INFN OAR by placing its Digital Object Identifier (DOI) reference in the "Output DOI" field below.

Pressing the **Prepare** button, the related model and parameters will be automatically extracted and placed in the following "PALMS Execution" fields.

Output DOI

PALMS output DOI  
Paste Output DOI reference (10.15161/oar/it/23504)

**Prepare**

PALMS execution

Please, insert a reference to PALMS model and parameters in DOI or HTTP URL format (e.g., 10.15161/oar/it/23504 or http://pobeview2.hpcfpo.org/repast/PALMS/input/batch\_params.xml).

Pressing the **Execute** button, PALMS software will be executed using the "Model" and "Parameters" fields specified below.

Model  
PALMS model  
Model as DOI, HTTP, or URL format

Parameters  
PALMS parameters  
Parameters as DOI, HTTP, or URL format

**Execute**

PALMS outputs

Output files are available in your personal Cloud area.

You can access them in **read-only** mode by pressing the **FILES** button.

For a detailed list of performed executions, press the **LIST** button.

You can choose an alternative model or parameters files and upload them in your Cloud area pressing the **UPLOAD** button.

Finally, to remove all PALMS submission outputs and delete your personal Cloud area, press the **DELETE** button.

FILES LIST UPLOAD DELETE

Implemented with: FutureGateway framework | FOG3: Science Software on Demand is a service operated and maintained by INFN Catania