SUPPORT4LHS

Process Mining and Knowledge Representation technologies to Support the Learning Health System

Deliverable 3.1 - Specification of the Data Management Plan (DMP)

Project deliverable

Authors: Mikel Egaña Aranguren (UPV/EHU), ???,

[1.- Executive summary 3](#_Toc132391246)

[2.- Introduction 5](#_Toc2132926119)

[3.- Datasets 5](#_Toc921869820)

[4.- Architecture and processes 5](#_Toc1043703698)

[5.- Conclusion 5](#_Toc103628451)

[6.- References 5](#_Toc1440617340)

[Annex A: Support4LHS FAIR data questionnaire 6](#_Toc353318005)

# 1.- Executive summary

# 2.- Introduction

The objective 3.1 ("Design and implementation of a FAIR Data Management Plan") of the project Support4LHS aims at publishing the datasets of the project following the FAIR principles1. In order to accomplish this, as specified in the grant agreement, two main tasks need to be fulfilled:

**(1) Design and creation of a Data Management Plan (DMP).** The DMP is presented in this deliverable: *Deliverable 3.1 - Specification of the Data Management Plan (DMP)*. This DMP is conceived as means to support the whole life cycle of the project data that will be collected, processed or generated. Therefore, the document is alive, and it will be updated through the life span of the project, accommodating changes that will be presented in the *Deliverable 3.2 - Final report on the Data Management Plan and degree of accomplishment of FAIR principles*.

**(2) Implementation and deployment of the DMP.** The implementation and deployment of the DMP will be carried out during the project, and the evaluation of its results presented in the *Deliverable 3.2 - Final report on the Data Management Plan and degree of accomplishment of FAIR principles*. However, an architectural overview of the processes designed to capture and publish the datasets is provided in this DMP.

The remainder of the document is organised as follows:

**Section 3 ("Datasets")** provides an overview of the datasets that are expected to be produced in the project, with information collected through the form described in Annex A.

**Section 4 ("Architecture")** describes the overall technical setting designed to capture and publish the data, including details on the resources allocated for the processes.

**Section 5 ("Conclusions")** wraps the document with final considerations for the future development of the DMP.

# 3.- Datasets

# 4.- Architecture

The basic architecture of the "FAIRification" framework that will be implemented also illustrates how the most salient points of the grant agreement will be realized (w.r.t objective 3.1). The main task of the FAIRification framework is to process the data produced in the project and publish it according to FAIR principles, as illustrated in Figure 1 (More details are provided in following figures).

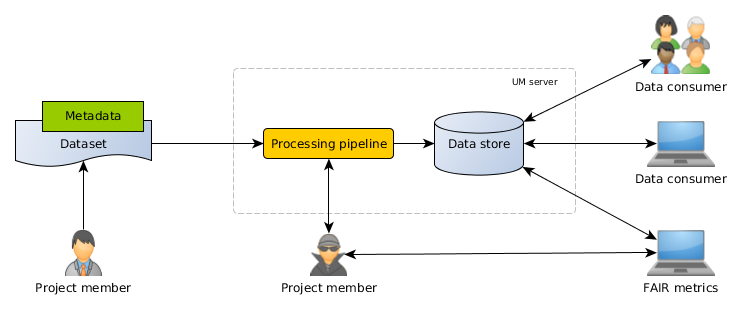


Figure 1: basic architecture of the FAIRification framework.

The process is divided into two main steps:

1) Processing pipeline: it acquires the data from the project members, it processes it, and it stores it in the data store.

2) The data is published according to FAIR principles and consumed by other, external scientists, or, more importantly, computational agents.

starts when a project member provides a new dataset, from the ones described in section 3, in a shared resource like a OneDrive folder.

# 5.- Conclusions

ok

6.- References   
1. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship : Comment. *Scientific data*. 2016;3:1-9.

# Annex A: Support4LHS FAIR data questionnaire