Innovative Multi-disciPlinary European Research training netwOrk on VolcanoEs

GA 858092 - H2020-MSCA-ITN-2019

Deliverable D7.3

Data management Plan V.2

Due by month 18 (February 2023)

Del.N.	Del. Title	Lead Beneficiary	Type	Dissemination level	Due date (month)	Due date
D7.3	Data management Plan V.2	INGV	Report	Confidential	18	February 2023

IMPROVE – Data Management Plan V. 2

This document constitutes the Deliverable 7.3 "Data Management Plan V2", representing the second version of the IMPROVE Data Management Plan (DMP).

1. Introduction/Objectives

According to the IMPROVE Grant Agreement (GA), the there-in defined IMPROVE data will be managed through a *Data Management Plan* and organized in a database fully available to the consortium and, after project closure, to the broad scientific community through the EPOS platform. EPOS will therefore provide long-term sustainability to storing, organization and access to the IMPROVE data. The data will be managed according to FAIR principles and rules, ensuring appropriate IPR for all partners and in particular for the industrial partners.

This report extends and integrates Deliverable 7.1 "Data Management Plan V1" (here-in-after DMP1) released at month six. DMP1 defined the constitution of a Data Management Team (DMT) within IMPROVE, which is managing any aspect of data collection, storage, and access, as well as the definition of each version of the DMP. Since then, a large IMPROVE field experiment has been conducted at the Krafla volcano, Iceland; plus surveys conducted both at Krafla and Etna volcanoes. The related data are being organized so as to make them accessible internally, as from the GA. This DMP is therefore aimed at i) describing the data already uploaded or being uploaded into the IMPROVE database; and ii) defining the regulations for accessing and utilizing those data.

2. Data survey

Data collection has been preceded by an internal survey in order for the DMT to know precisely the type and size of the multidisciplinary data being collected, and to ensure that the formats comply with the EPOS (and IMPROVE) standards.

The following table summarizes the data already shared and available (or being made available at the date of this report) through the IMPROVE database. The total size of data is >1 TB.

Table 1. List of the current datasets in the database

Partner owning the data	Data description
INGV	CO2 flux and soil temperature
UNIVBRIS	Gravity and GNSS
LMU	Petrological/Chemical
GFZ	Gravity
UI	GNSS
DIAS	Seismic arrays (Krafla and Etna)
USMB	ERT, MT, Seismic

3. Data access

This paragraph defines the rules of access and usage of data:

- All data are internally available to all IMPROVE partners through the Repository framework (described in the next section).
- At this stage of the project, the individual participants from the Partners are Authorized users, meaning that specific permission to access the data and data products in the Repository is granted to them, and the access is subject to a login and password.
- The individual participants who are granted access are those listed in Annex 2 of the IMPROVE Consortium Agreement (CA). Possible other participants from the IMPROVE Partners, proposed by the Supervisory Board (SB) member from that Partner, can be authorized after approval by the SB.
- The Repository first, and a structured database later on, are the instruments granting data sharing within the project consortium. Inside the Repository it is indicated the way to cite the shared datasets.
- The normal condition for using data from other owning partners is scientific cooperation. Other usage rights can be granted, depending on specific agreements between the user and the owner (see below).
- Visualization of the datasets (either the data files or their plots through dedicated software) by authorized users does not require any permission additional to the permission to access the database as outlined above. Instead, use of data of which the user is not the owner (as defined at Table 1 above) requires contacting the owner in order to obtain the related permission. Such a permission should be given in written form (e.g., an email or - depending on specific agreements - more advanced forms of written agreement).
- In case of multiple ownership, the request of using the data must be addressed to all owners as they are cited in the data repository.

4. Intellectual Property Rights

- Each partner holds the IPR of the data of which it is owner;
- Each partner can exploit IPR of other partners only when a specific agreement is in force with the IPR owner.
- IPR exploitation after project end will be specified in DMP V.3 which is to be provided at month 48.

5. Data Repository tool

The data repository (Repository) is a tool designed to meet the requirements of the Open Data model for the publication of scientific data in the form of catalogs, called

datasets. The implementation is based on the INVENIO RDM (Repository/Document Management) platform, a product of the INVENIO open-source project initially developed by CERN. The proposed platform allows researchers to archive, according to the Open Access paradigm and FAIR principles, scientific reports, presentations, images, datasets, software, and any other type of digital objects related to research activities. Each submission is associated with a landing page, resolvable through a public Digital Object Identifier (DOI) making the published product easily "citable". The developed software, forking from the INGV-OE DataRep solution [Torrisi et al., 2022], will be customized to the project scientific needs, such as restricted access to the consortium members. The system will also provide the integration with raw data repositories and the TSDSystem platform [Cassisi et al., 2015], both developed by the UFSO-IT group of INGV-OE.

February 28, 2023

For the Consortium

The Coordinator

Paolo Papale