

D7.2 Data Management Plan

WP7 Management

Due M : 6

Submission date : 07/07/2023

Lead beneficiary : CNRS

Author : Agnès Deboulet

Executive summary

This DMP is a global scheme that applies to all Fairville labs involved in the Fairville project. It expands on the information contained in the Fairville Scientific Proposal, the paragraph on open access to research data in the Grant Agreement and the Consortium Agreement for Fairville. The Fairville Data Management Plan is based on the general Fairville guidelines set out in this document. In addition, it contains details of the type of data collected or generated and how it is made FAIR (findable, accessible, interoperable, reusable).

Research data that is generated as part of the Fairville project and is not restricted by law (GDPR, IPR) will be made available to the wider scientific community via publication in a data repository such as Huma-Num at the end of the project. Personal information concerning human individuals will be processed in accordance with European and national legislation, as specified in Fairville Scientific's proposal (see grant agreement). To ensure long-term access, research data will be stored as far as possible in sustainable formats and will be documented in accordance with discipline-specific standards and licenses.

1.Prepare

1.1 Data collection

Within the framework of the Fairville project, participants will collect and generate a variety of research data to address the research questions raised in their respective Fairville labs. They will gather existing data such as audiovisual material, interview transcripts, social media data, websites, archives, (digital copies of) printed materials, drawings and photographs, journals and articles. In addition, they will create interview data, survey data and video and audio recordings. At the end of the project, reusable research data will be archived in sustainable, interoperable data formats.

1.2 Data documentation

Where datasets are collected and stored as part of the project, they will be fully documented in accordance with EU standards, where these apply. Conventions on file names, folder hierarchies and "readme" file information can help to ensure that data sets can be understood in the long term. For data archived at the end of the project, metadata will be added in accordance with the repository standard (e.g. Dublin Core metadata). Any other documentation required to understand and reuse the datasets will be stored with the data.

2. Handling research Data-management

2.1 Data storage and back-up

To store and share data safely during the project, participants will use secured facilities

provided by two public institutions (e.g. My Core Space, provided by CNRS but also tentatively for daily storage, LaboMega performed by AMU partner. My Core Space complies with the CNRS information security policy, and backups are carried out regularly. Researchers should ensure that back-up procedures are in place. If not, they will make periodic backups themselves. They will follow local information security policies and, where necessary, seek support from local help desk staff, privacy officers and the institution's data protection officer.

2.2 Data access and security

Many of the data gathered in the Fairville project contain personal or sensitive information. Personal data will be replaced by code names as quickly as possible. Data re-identification keys will be kept in a separate, secure repository. Data will be anonymized as far as possible according to the needs of the research. Participants will be asked to agree to share datasets within the network, and access to these will be organised accordingly. For this reason, the CNRS provides My-Core, which enables encryption. Consent forms will mention the option of secondary use. If consent is not received for secondary use, data that can be traced back to individual persons will be deleted after completion of the proposal. Privacy officers in each country will provide advice on local procedures for handling and processing personal data, in conformity with national and European legislation.

3. Preserve and share

3.1 Data preservation and archiving

Copyrighted material and raw data containing personal or sensitive data cannot be

shared with the wider community. They are kept only within the framework and for the duration of the research project. Processed and anonymized datasets and all other datasets that cannot be traced back to individual persons will be archived in compliance with FAIR guidelines, which means that community standards are being adopted as well as licenses that allow reuse. The Fairville Data Management Plan will provide more details on how FAIR principles will be applied to each pilot project.

3.2 Data sharing and reuse

Research data resulting from the Fairville project that may be of interest to the scientific community or the general public will be made available for re-use via a data repository such as Huma-Num, with a licence authorising re-use and a reference to the grant number. The data repository will provide a permanent identifier enabling datasets to be linked and cited in publications. The data will be accompanied by all documentation and, where appropriate, software to enable reuse.

Fairville's Data Management Plan

Name of the scientific coordinator	Agnès Deboulet
Email Address	deboulet@gmail.com
Name of the Project	FAIRVILLE
Funding body(ies)	European Commission — Research Executive Agency
Grant number of Fairville	RIA-FAIRVILLE-CL2 DEMOCRACY GA N°101094991
1a. How will new data be collected or produced and/or how will existing data be re-used?	Within the framework of Fairville, data are produced in several ways • participatory action research (co-production of data determined by each collective in each field; field reports drawn up mainly by researchers) • semi-directive interview surveys (transcriptions and reports). • The project will continue to collect data, to build up (and subsequently publicise) the archives of the several actors involved in the co-production processes studied. • to complete the interview surveys (grey literature, press articles, photographic images, etc.). • All textual documents are titled, geo-located, dated and list the authors.
1b. What data (for example the kind, formats, and volumes), will be collected or produced?	Different types of data will be produced: • video recordings: format to be defined • audio recordings: .wav and .mp3 formats to be preferred • images: .jpg formats to be preferred • text documents: preferred .doc and .pdf-a formats • Fairville favours widely used formats that are compatible with open software. • At this stage, the expected volumes are: ◦ Around one hundred semi-directive interviews (audio files and transcripts) ◦ Around 40 observation reports (.doc format). • The volumes of video and audio rushes have yet to be determined, pending the definition of participatory action research.
2a. What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany the data?	• The research methodology, including that of data collection, will be spelled out in the documents relating to the various project tasks. • In the collective work tool (Fairville website, see point 3), files are organised by task and by field. A naming convention has also been created: FV (for FairVille) + XX (task or terrain) - XX (authors' initials) - Date - Title. • In the next months, we plan to begin discussions with Huma-NUM regarding the secondary storage of data from the project. At that point, metadata standards will be more precisely defined to ensure compatibility with their platform. In September, we will start having the data stored in Lanomega (managed by Aix Marseille University) or by Nextcloud. This double level of storage will ensure the good preservation of data

2b. What data quality control measures will be used?	<i>It seems to us that quality control issues mainly concern observation reports. A grid has been drawn up collectively. In addition, all reports are drawn up collectively - at the very least, writing/proofreading is shared between the pair of researchers present at the meeting.</i>
3a. How will data and metadata be stored and backed up during the research?	<i>Data will be stored progressively on the Fairville Website. All team members have access to it.</i>
3b. How will data security and protection of sensitive data be taken care during the research	<i>The Fairville website allows the definition of differentiated access and editing rights, by file, so as to respect confidentiality issues. The data producer retains control over how the data is shared within the collective.</i> <i>The data is a priori non-public - no data will be made public without the explicit agreement of the people concerned. In view of the subject of the research (co-production with local residents, sometimes in the context of mobilizations against development projects), some data are politically sensitive, as they reveal internal debates that should not step out of the community of co-researchers</i>
4a. If personal data are processed, how will compliance with legislation on personal data and on security be ensured?	<i>The project envisages the following provisions with regard to the RGPD:</i> <i>seeking the consent of data subjects to the preservation/sharing of personal data and/or its anonymization (including indirect personal data).</i>
4b. How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?	<i>This question was discussed at a session organised by the Ethics Committee. Data are accessible to members of the Fairville project.</i>
4c. What ethical issues and codes of conduct are there, and how will they be taken into account?	<i>An ethics committee made up of researchers and field workers has been set up and can be consulted by all team members at any time.</i>
5a. How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?	<i>These questions have not yet been decided, but the intentions are clear:</i> <i>To make as much data as possible publicly accessible, while preserving the integrity of the individuals and collectives concerned.</i> <i>The team's aim is for data with a public vocation to be hosted by Huma-Num.</i> <i>The archives of the other partners involved in the project will be stored at the LAVUE office (Nanterre University).</i>
5b. How will data for preservation be selected, and where data will be preserved long-term (for example a data repository or archive)?	<i>The project team operates on a collegial basis, with decisions made at monthly team seminars. Decisions will be taken at these times, respecting the opinions of the different work teams (task and/or field).</i> <i>The data may be used by any researcher in the team, in particular the work teams that were responsible for their production, for the duration of the project and for two years thereafter.</i>
5c. What methods or software tools are needed to access and use data?	<i>The aim is to use formats that are as widespread and accessible as possible.</i> <i>Long-term access conditions will be defined with the Huma-NUM team.</i>
5d. How will the application of a unique and persistent identifier (such as a Digital	<i>We will refer to Huma-NUM practices.</i>

Object Identifier (DOI)) to each data set be ensured?	
6a. Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?	• Roles to be defined with Huma-NUM. • During research, data management is done in a collegial manner. Audrey Debargue (Project Manager of Fairville) is responsible for the technical management of the data storage in the Fairville Website. • The data management plan is the responsibility of the project's scientific coordinator Agnès Deboulet.
6b. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?	• The Fairville website hosting costs are covered by the Fairville project budget. • Preparing data for long-term storage has not yet been discussed or decided.

About this Data Management Plan :

Date written	01/05/2023
Date last update	20/06/2023
Version	1