D8.2: Data Management Plan

Revision: v.1.1

Work package	WP8
Task	T8.3
Due date	31/03/2024
Submission date	29/03/2024
Deliverable lead	LOR
Version	1.1
Authors	Iñaki Sainz (LOR)
Reviewers	Christos Papaioannou (LMS)
Abstract	This document collects the Data Management Plan of the COROB project. It gives details about the consortium policy and the data management responsible. It also specifies how data needs to be findable, accessible, interoperable and re-usable. It finally details the dataset to be generated during the project.
Keywords	Data, Data Management Plan, License

CHANGE CONTROL

Date	Version	Editor	Changes to document
2024/01/26	v1.0	Iñaki Sainz	First release
2024/03/13	v1.1	Iñaki Sainz	Data quality section added. Data summarie Funding corrected to co-funding.

Important remarks:

- The contributors listed in this table and on the front page are the report's primary editing authors. It is important to note that all COROB partners are contributing critical technical contributions to this ongoing work.

DISCLAIMER

Co-funded by the European Union

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

COROB (*Cooperative robotics powered by AI and data for flexible production cells*) project is co-funded by the EU's Horizon Europe programme under Grant Agreement number 101120640. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them. This work has received funding from the Swiss State Secretariat for Education, Research, and Innovation (SERI).

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

COPYRIGHT NOTICE

Project co-funded by the European Commission in the Horizon Europe Programme		
Nature of the deliverable:	**DMP**	
Dissemination Level		
PU	*Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project's page)*	✓
SEN	*Sensitive, limited under the conditions of the Grant Agreement*	
Classified R-UE/ EU-R	*EU RESTRICTED under the Commission Decision No2015/ 444*	
Classified C-UE/ EU-C	*EU CONFIDENTIAL under the Commission Decision No2015/ 444*	
Classified S-UE/ EU-S	*EU SECRET under the Commission Decision No2015/ 444*	

* *R: Document, report (excluding the periodic and final reports)*

DEM: Demonstrator, pilot, prototype, plan designs

DEC: Websites, patents filing, press & media actions, videos, etc.

DATA: Data sets, microdata, etc.

DMP: Data management plan

ETHICS: Deliverables related to ethics issues.

SECURITY: Deliverables related to security issues

OTHER: Software, technical diagram, algorithms, models, etc.

EXECUTIVE SUMMARY

In COROB several kinds of data will be generated, so Deliverable D8.2 aims at providing information and guidance for the correct management of that data. The following list provides an overview of the different type of data generated:

Design of arc-based flexible, cooperative and intelligent multi-robotic cells.

- Pre-process data including point clouds, images from cameras and others.
- Process-time data from the process actuators and sensors.
- Process-time inspection data.
- Post-process data from inspection activities.
- Guidelines and recommendations.

This document is based on a standard Data Management Plan created by LORTEK for its coordinated European projects.

TABLE OF CONTENTS

LIST OF TABLES

ABBREVIATIONS & GLOSSARY

DMP	Data Management Plan
DMPR	Data Management Plan Responsible
DOI	Digital Object Identifiers
FAIR	Findable, Accessible, Interoperable and Reusable
IPR	Intellectual Property Rights
OpenAIRE	Open Access Infrastructure for Research in Europe
ORP	Open Research Programme
SI	International System of Units
TBD	To Be Defined
WP	Work Package

1 DATA MANAGEMENT AND RESPONSIBILITY

1.1 CONSORTIUM POLICY

According to Open Research Programme (ORP) requirements, the COROB Data Management Plan (DMP) will be ruled by FAIR (Findable, Accessible, Interoperable and Reusable) Data Management Protocols. The ORD pilot applies primarily to the data needed to validate the results presented in scientific publications. Open access to other data is encouraged on a voluntary basis if it is not sensitive or subject to protection.

Publishable data will be made accessible within 6 months of publishing the data in peer reviewed scientific articles or similar, unless beneficiaries have outlined justifiable reasons for maintaining data confidentiality.

Each beneficiary is responsible for their records and documentation in relation to data generated, which must be in line with the accepted standards in the respective field (if do exist). To avoid losses, beneficiaries must take measures to ensure that data is backed-up.

The General Assembly will rule the IPR and will meet at each face-to face meeting as well as every time (via teleconference) any WP leader proposes open access of generated data.

1.2 DATA MANAGEMENT RESPONSIBLE

The Project Data Contact will be the Project Coordinator, who is the direct contact with the European Commission. He will ensure that the DMP is respected with the support of the WP leaders. He will oversee:

- Ensuring the data is correctly uploaded into repositories through periodical checks.
- Completing the DMP with the links related to the data and its regular update.
- Ensuring the data availability.

Ensuring that information related to accessible data is in accordance with the produced data.

Table 1: Project Data Contact

Project Data Contact (PDC)	Iñaki Sainz
PDC affiliation	LORTEK
PDC email	isainz@lortek.es

1.3 DATA SUMMARY

Table 2 provides a general overview of the type of data that will be generated, and which will be open access. Each open access dataset will be described in detail according to the template showed in Appendix A. It includes the available information for each identified open access dataset. Updates of DMP will include any update in these tables.

Table 2: Data Summary

Data nature	Type*	Use	Origin (Partner)	Potential users	Data Format	Open access
Design of arc-based flexible, cooperative and intelligent multi-robotic cells.	DES	Development of the arc-based multi-robotic cells	LOR, DGH, LMS, GROB	Universities, research centres, Industrial companies	PDF, PPT	Partial
Pre-process data including point clouds, images from cameras and others.	NUM	Perform the necessary operations to be ready for the process.	LOR, DGH, LMS, GROB, INTRA	LOR, DGH, LMS, GROB, INTRA	TBD	Partial
Process-time data from the process actuators and sensors.	NUM	Training of the models.	LOR, DGH, LMS, GROB, INTRA	LOR, LMS	TBD	Partial
Process-time inspection data.	NUM	Training of the models.	LOR, DGH, LMS, GROB, INTRA	LOR, LMS	TBD	Partial
Post-process data from inspection activities.	NUM	Training of the models.	LOR, DGH, LMS, GROB, INTRA	Universities, research centres, Industrial companies	TBD	Partial
Guidelines and recommendations.	PRO	Provide recommendations and guidelines to universities, research centres and industrial companies	COROB consortium	Universities, research centres, Industrial companies	PDF	✓
RL models for adaptive impedance control	COD	Reduce the positioning error of manipulator robots during welding	LOR	LOR	TBD	No
Active learning algorithms	COD		LMS	LMS	TBD	No

Data nature	Type*	Use	Origin (Partner)	Potential users	Data Format	Open access
Human-centred UX Design and Evaluation	DES	Provide outcomes of the human centred design of the COROB UX	TU/e	Universities, research centres, Industrial companies	PDF	✓
Design of the COROB platform	DES	Development of the	INTRA	Universities, research centres, Industrial companies	PDF	✓
Final validation	PRO	Disseminate the results of the project.	COROB consortium	Universities, research centres, Industrial companies	PDF	✓
Scientific publications	DES	Disseminate the scientific results of the project.	LOR, LMS, DGH, GROB, INTRA	Universities, research centres, Industrial companies	PDF, TBD	✓
Contact Details		Communication and Dissemination activities	D4P	D4P for communication and dissemination purposes	Txt,xls,word. pdf	No

** NUM: numerical, EXP: experimental, COD: code, PRO: protocols, DES: design*

1.4 DATA QUALITY

The author(s) of the dataset are responsible for the quality assurance of the dataset, both in terms of format and actual data quality. The experimental quality of the datasets should be carefully checked before publication to ensure that only datasets with added value are published.

2 FAIR DATA

COROB will ensure the generation of FAIR data, so it will take measures to ensure that data is Findable, Accessible, Interoperable, and Reusable.

2.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

COROB will deposit generated and collected data in an open online research repository under Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC 0). The primary repository selected in COROB is ZENODO, which was developed by CERN as part of the OpenAIRE (Open Access Infrastructure for Research in Europe) project. ZENODO allows researchers to deposit both publications and data, providing tools to linking them to these through persistent identifiers and data citations. It facilitates the finding, assessing, re-using and interoperating of datasets which are the basic principles for FAIR data that projects must comply with. A dedicated ZENODO community has been created for COROB, where the created open data will be collected (https://zenodo.org/communities/corob).

The guidelines provided by ZENODO will be used by COROB to comply with FAIR principles.

In order to store and make findable any COROB openly accessible data, the chosen online repository needs to facilitate identification of data and refer to standard identification mechanisms (ideally persistent and unique identifiers such as Digital Object Identifiers (DOI)), which should be outlined.

The dataset naming should be according to this scheme:

[Name of the project]_[Type of Data]_[Date]_[Name of dataset]

- Name of the project: “COROB”
- Type of data: “NUM”, “EXP”, “ENV”, “SOC”
- Date: YYYYMMDD
- Name of the dataset

Metadata of deposited data must be open under a Creative Common Public Domain Dedication (**CC 0**) or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the FAIR principles and provide information at least about the following:

- Datasets (description, date of deposit, author(s), and embargo).
- Horizon Europe funding.

Project name, acronym and number.

- Licensing terms.
- Persistent identifiers for the dataset.

- The organisations involved in the action.
- Where applicable, the metadata must include persistent identifiers for related publications and other research outputs.

2.2 MAKING DATA OPENLY ACCESSIBLE

In order to maximize the impact of COROB data, the project will facilitate sharing of results and data within and outside the consortium. Selected data and results will be shared with the scientific community and other stakeholders through publications in scientific journals and presentations at conferences, as well as through open access data repositories. There will be an open access policy applied to these following the rules outlined in the Grant Agreement.

The General Assembly will review and approve all data that is identified as appropriate for open access. This process will be carried out on an ongoing basis to facilitate the publication of appropriate data as soon as possible. The General Assembly is responsible for the IPR issues within COROB and their approval will avoid any possible conflicts between open access and IPR issues.

All data will be made available for verification and re-use, unless the data owner can justify why data cannot be made openly accessible. The General Assembly will assess such justifications and make the final decision, based on examination of the following elements regarding confidentiality of datasets:

- Commercial sensitivity of datasets
- Data confidentiality for security reasons
- Conflicts between open-access rules and national and European legislation (e.g. data protection regulations)

Sharing data could jeopardize the objectives of the project

Other legitimate reasons, to be validated by the General Assembly

Upon deciding that a database should be kept confidential, the reasons for doing so will be included in an updated version of the DMP. The data will be accessible through:

- Publications in scientific journals
- The Project website

ZENODO repository (or any other repository complying with statements in section 2.1)

To encourage re-use and further application of project results, all COROB data that underlies scientific publications will be made available via open-access online platforms, unless subject to protection, OR unless release of all or part of the data to open-access platforms could jeopardize the project's main objectives.

2.3 MAKING DATA INTEROPERABLE

Partners will observe OpenAIRE guidelines for online interoperability, including as set of guidelines that includes OpenAIRE Guidelines for Literature Repositories, OpenAIRE Guidelines for Data Archives, etc. These guidelines can be found at: https://guidelines.openaire.eu/en/latest/.

The project will use vocabularies, standards, formats and methodologies as defined in Appendix A to facilitate interoperability.

2.4 INCREASE DATA RE-USE

COROB is expected to produce a novel data and knowledge through experimental approaches that will be presented to the scientific community and industry, through a carefully designed portfolio of dissemination actions. Datasets uploaded in the ZENODO repository will be freely accessible after an embargo period determined per dataset, if required.

As the project progresses and data is identified and collected, further information on increasing data re-use will be outlined in subsequent versions of the DMP. In specific, information on how data will be licenced to permit the widest reuse possible, when the data will be made available for re-use, whether the data produced and/or used in the project is useable by third parties and specifications of length of time for which the data will remain reusable will be provided.

3 DATA SECURITY

COROB will ensure safety store of data by the following ways:

- Use of ZENODO.
- All along the project, data are shared and stored in a secured SharePoint hosted by the Project Coordinator.
- Each beneficiary will keep a back-up of the own generated data.

APPENDIX A. OPEN ACCESS DATASET INFORMATION

Dataset [name dataset]

Data Storage	ZENODO
Link to repository (only if public)	Link to the dataset in: https://zenodo.org/communities/corob
Dataset identifier	*DOI number*
Data license	CC by 4.0 (if public)
Date	*YYYY/MM/DD*
Data version	*Version number*
DMP version	*Version of the DMP*
Partner organizations involved in this dataset	*List of companies involved*
DMPR name	Iñaki Sainz
DPMR email	isainz@lortek.es
Description of the data	*Give a short description of the data, including amount (if known) and content.*
Nature of the data	*Select the nature the data:* • *Experimental data* • *Numerical data* • *Statistical data* • *Environmental data* • *Social data* • *Other (specify)*
Type of the data	*Select the type of data:* • *text,* • *numbers,* • *images,* • *3D models,* • *software,* • *audio files,* • *video files,* • *reports,* • *other (specify)*
Size of the data	*Size of the data in bytes (kB, MB or GB)*
Data file	*Data files names*
Unit measurement system	*SI*
Potential users	*Universities, Research Centres, Industrial companies…*
Ethical Issues	*N/A*