

FCA – Initial Data Management Plan

Deliverable ID:	D3.2
Project acronym:	FCA
Grant :	101114764
Call :	HORIZON-SESAR-2022-DES-IR-01
Topic :	Flight Centric ATC
Consortium coordinator :	DLR
Edition date:	31 August 2023
Edition:	01.00
Status :	Draft
Classification:	PU

Abstract

This Initial Data Management Plan (DMP) describes the procedures for data management that the FCA project will follow during and beyond its three-year duration. The plan provides an overview on the datasets expected to be generated by the project and details how compliance with the FAIR principles for data management will be ensured. This document does not describe how deliverables or other communication, dissemination or exploitation material will be managed and published. The document will be updated over the course of the project whenever significant changes arise (e.g. new data or new policies).

Authoring & approval

Author(s) of the document

Organisation name	Date
DLR	02 August 2023

Reviewed by

Organisation name	Date
AFT	08 August 2023
DLR	18 August 2023
ENAIRe/CRIDA	08 August 2023
EUROCONTROL	04 August 2023
INDRA	18 August 2023
Integra	22 August 2023

Approved for submission to the SESAR 3 JU by¹

Organisation name	Date
AFT	24 August 2023
DLR	25 August 2023
ENAIRe	Silent approval
EUROCONTROL	Silent approval
Indra	Silent approval
Integra	30 August 2023
UKSATSE	30 August 2023

Rejected by²

Organisation name	Date

¹ Representatives of all the beneficiaries involved in the project

² Representatives of the beneficiaries involved in the project

Document history

Edition	Date	Status	Company author	Justification
00.01	27 July 2023	Draft	DLR	Initial version
00.02	02 August 2023	Draft	DLR	Initial version for internal review
00.03	23 August 2023	Initial version	DLR	Initial version for approval
01.00	31 August 2023	Final	DLR	Initial DMP approved for submission

Copyright statement © 2023 – (AgentFly, DLR, ENAIRE, EUROCONTROL, INDRA, INTEGRA, UKSATSE). All rights reserved. Licensed to SESAR 3 Joint Undertaking under conditions.

FCA

FLIGHT CENTRIC ATC

FCA

This document is part of a project that has received funding from the SESAR 3 Joint Undertaking under grant agreement No 101114764 under European Union's Horizon Europe research and innovation programme.



Table of contents

1	<i>Data summary</i>	7
2	<i>FAIR data</i>	8
3	<i>Other research outputs</i>	10
4	<i>Allocation of resources</i>	11
5	<i>Data security</i>	12
6	<i>Ethics</i>	13
7	<i>Other issues</i>	14

List of figures

No table of figures entries found.

List of tables

Table 1: List of acronyms	6
Table 2: Data generated by the project	3
Table 3: Input data used within the project	3

List of acronyms

Acronym	Description
ATC	Air Traffic Control
BADA	Base of Aircraft Data
CDEP	Communication, Dissemination and Exploitation Plan
DLR	Deutsches Zentrum für Luft- und Raumfahrt e.V.
DMP	Data Management Plan
DOI	Digital Object Identifier

EC	European Commission
FAIR	Findable, Accessible, Interoperable, Re-usable
FCA	Flight Centric ATC
GDPR	General Data Protection Regulation
INTEROP	Interoperability Requirements
OSD	Operational Service and Environment Definition
PMP	Project Management Plan
SPR	Safety and Performance Requirements
VALP	Validation Plan
VALR	Validation Report
WP	Work Package

Table 1: List of acronyms

1 Data summary

Over the course of the FCA project, different types of datasets will be generated and managed by the beneficiaries. The purpose of these datasets is the demonstration of feasibility of the Flight Centric ATC (FCA) concept on the targeted maturity level V3 and the validation of the objectives that will be defined in the deliverable D2.6 Final Validation Plan (VALP).

An overview of the datasets that, at the time of submission of the Initial Data Management Plan (DMP), are expected to be generated throughout the FCA project is presented in Table 2. It should be noted that at this point, the Solution's Initial VALP and SPR-INTEROP/OSED have not yet been completed. Since these documents define the scope of the Solution and the validation activities of the exercise, it is only possible in this Initial Data Management Plan to vaguely indicate what type of data will be generated.

The consortium believes that the utility of all listed datasets to researchers outside of the project is limited. This is due to the fact that the datasets will be generated precisely to fit the scope of the FCA project. Even though it is a successor of previous SESAR projects addressing the Flight Centric ATC concept, the scope of this FCA project differs significantly and data collected in previous validation exercises cannot be re-used for the validation of the specific use cases and validation objectives.

All data collected throughout human-in-the-loop simulation activities within this project will be pseudonymized. Participants will be assigned unique identifiers. These identifiers will be utilized for all information gathered during the execution of the validation exercises, including performance data, questionnaires during, before or after simulation runs, feedback from de-briefing sessions, and individual participant interviews. This approach is essential to establish connections between data recorded using various methods (e.g., linking questionnaires with performance data) at a later stage. It is important to note that no documents will be created that directly associate the unique identifier with individual participants, ensuring their privacy and confidentiality.

As a general principle, the project will only gather and generate information that is relevant and necessary for the project (principle of data minimisation).



Beneficiary	Work Package (WP)	Type of data	Origin of the data	Purpose of the data	Data format	Size of data	Publication of the data
DLR	WP1	Project team contact list	Project members	Project management	.xlsx	< 1 MB	No – the contact list contains personal information and is protected by the GDPR.
DLR	WP2 (EXE002)	Simulation/workshop participants contact list	Simulation or workshop participants of EXE002	Coordination of validation activities and workshops	.xlsx	< 1 MB	No – the contact list contains personal information and is protected by the GDPR.
DLR	WP2 (EXE002)	Questionnaire responses	Simulation or workshop participants of EXE002	Validation of requirements and validation objectives	.xlsx	< 100 MB	Yes – anonymized questionnaire responses may be published.
DLR	WP2 (EXE002)	Verbatim transcript/summary of verbal feedback	Outcomes from meetings, workshops, interviews or simulation debriefings (EXE002)	Collection of stakeholder recommendations, ensuring a target-oriented development and validation of the concept	.docx .xlsx	< 500 MB	No – this data cannot be published, as it will most likely contain sensitive information regarding the simulation platform developed by DLR and is protected by intellectual property rights. However, a summary of the feedback received will be provided and



							selected anonymized quotes may be published.
DLR	WP2 (EXE002)	Flight performance data	TrafficSim platform	Validation of requirements and validation objectives related to performance	.csv	< 5 MB	Yes.
DLR	WP2 (EXE002)	Screen and voice recordings	TrafficSim platform	Evaluation of traffic situations during the simulations	.mp4 .wav	< 50 GB	No – this data cannot be published, as screen recordings contain sensitive information regarding the simulation platform protected by intellectual property rights as well as the traffic scenario provided by a third party, while voice recordings are protected by the GDPR.
ENAIRe	WP2 (EXE001)	Simulation/workshop participants contact list	Simulation or workshop participants of EXE001	Coordination of validation activities and workshops	.xlsx	< 1 MB	No – the contact list contains personal information and is protected by the GDPR.
ENAIRe	WP2 (EXE001)	Questionnaire responses	Simulation or workshop	Validation of requirements and validation objectives	.xlsx	< 100 MB	Yes – anonymized questionnaire responses may be published.

			participants of EXE001				
ENAIRE	WP2 (EXE001)	Verbatim transcript/summary of verbal feedback	Outcomes from meetings, workshops, interviews or simulation debriefings (EXE001)	Collection of stakeholder recommendations, ensuring a target-oriented development and validation of the concept	.docx .xlsx	< 500 MB	No – this data cannot be published, as it will most likely contain sensitive information regarding the simulation platform developed by INDRA and is protected by intellectual property rights.
ENAIRE	WP2 (EXE001)	Flight performance data	iTEC platform	Validation of requirements and validation objectives related to performance	.csv	< 10 MB	No – this data cannot be published, as it will most likely contain sensitive information regarding the simulation platform developed by INDRA and is protected by intellectual property rights. However, the data will be exploited and the results will be published within the solution's Validation Report (VALR).
ENAIRE	WP2 (EXE001)	Screen and voice recordings	iTEC platform	Evaluation of traffic situations during the simulations	.mp4 .wav	< 50 GB	No – this data cannot be published, since it contains sensitive information



							regarding controllers' way of working.
--	--	--	--	--	--	--	--

Table 2: Data generated by the project

Besides the data generated within the project, input data will be handled and is listed in Table 3 below. Since these datasets originate from external sources, they will not be published by this project.

Beneficiary	Work Package	Type of data	Origin of the data	Purpose of the data
DLR	WP2 (EXE002)	Traffic scenario	EUROCONTROL	Realistic simulation of air traffic for validation activities
DLR	WP2 (EXE002)	Aircraft performance data	EUROCONTROL	Realistic simulation of air traffic for validation activities
ENAIRe	WP2 (EXE001)	Traffic scenario	EUROCONTROL	Realistic simulation of air traffic for validation activities
ENAIRe	WP2 (EXE001)	Aircraft performance data	EUROCONTROL	Realistic simulation of air traffic for validation activities

Table 3: Input data used within the project



2 FAIR data

The FCA project is committed to the Open Science approach, aiming to maximize public access to research data as well as enabling re-use of the data by researchers outside of the project.

Data collected for research purposes (including questionnaires, interviews, observations, audio/screen recordings (where applicable)), as well as promotional materials gathered during data collection activities and events that may be video recorded or photographed, will be subjected to current European regulations on matters of data handling and privacy (GDPR, Regulation (EU) 2016/679). The research outcomes will be reported without contravening the right to privacy and data protection.

2.1 Making data findable, including provisions for metadata

In the project, data intended for publication will be uploaded to a trusted repository. At this stage of the project, a specific repository for publication of data has not been selected. It will be identified in a subsequent iteration of the data management plan. However, candidates for a repository have been identified, such as Zenodo, which applies DataCite's Metadata Schema. For any set of data uploaded to the repository, metadata must be submitted, containing at least the following properties:

- Identifier (Digital Object Identifier (DOI), assigned by the repository)
- Creator (name of the institution or person that created the data)
- Title (descriptive name of the dataset)
- Publisher (name of the beneficiary that holds the data)
- Year of publication (year in which the data is made publicly available. In case the data is protected by an embargo period, this would be the year in which the embargo expires.)
- Type of data (e.g. "dataset", "report",...)
- License (outlining the conditions for usage of the data)

Besides the above-listed entries, further properties can be described. If Zenodo is used as a repository, and thereby DataCite's Metadata Scheme is applied, keywords will be added whenever possible, using the free-text property "Subject" to add keywords, in order to enhance possibilities for discovery and re-use. In case the project ultimately chooses another repository, it will be one that ensures the provision of indexed and harvestable metadata and keywords.

As the project will generate a wide range of types of data, there is no strict rule for the naming of files. However, file names shall contain the following elements:

- Project ID and acronym
- Descriptive title
- Version number

Names of datasets created within an exercise should furthermore contain the following information:

- Exercise ID
- Validation run
- Participant ID

Validation run and participant ID are however only to be added to the file name when applicable, since it is likely that some datasets will not be specific to one run and/or participant.

2.2 Making data accessible

2.2.1 Repository

The FCA project uses the following repositories:

Protected institutional servers: Each project partner generating data within the project is responsible for storing the data internally on protected servers and controlling access to it.

Teamsites: For the duration of the project, a teamsite is used to work in a collaborative manner on deliverables as well as exchange information. It may furthermore be used to store and make small sets of data (< 100 MB) available to the partners. Access to the teamsite is limited to persons who are actively involved in the project.

STELLAR platform: As part of the SESAR programme, the FCA project will use STELLAR as a repository provided by SESAR 3 JU. Besides its purpose as a tool for project management and information exchange, it might be used as a repository for uploading small datasets.

Open-access repository: Data that the project intends to publish will be uploaded to a trusted repository. A specific repository has not yet been chosen but will be named in a later iteration of the data management plan. However, data will only be published via a repository that follows the FAIR principles for data management, assigns a persistent identifier to each uploaded dataset and requires the provision of metadata. One repository that fulfils these criteria and is under consideration is Zenodo, a freely accessible repository. It was developed as part of the European OpenAIRE program, commissioned by the European Commission, and allows researchers of all scientific fields to collaborate, manage, and publish their data and software openly.

2.2.2 Data

Whenever necessary and unless the data is sensitive, the project partners will share research data within the consortium. Furthermore, the FCA project is committed to making data generated in the project as openly available as possible, meaning that it will be uploaded to an open-access trusted repository, as long as the publication does not violate the General Data Protection Regulation (GDPR) or commercial interests of the individual project partners such as intellectual property. For the publication, a repository will be chosen that allows access through a free and standardized access protocol, such as the OAI-PMH protocol used by Zenodo.

Table 3 in chapter 1 of this document shows which datasets are intended to be published or not, along with a justification for not publishing certain data.

Data that is published via an open-access repository will not be made available to the public before the end of the project and might be restricted by an embargo for a certain period beyond the end of the project in order to grant the project partners priority for exploitation and dissemination of the generated data.

In the case of access requests to sensitive or personal data, it is the responsibility of the beneficiary holding the data to evaluate according to the GDPR and, if appropriate, approve the request. This should be supported by the appointed data manager of the respective beneficiary, and, if necessary, the project's data management lead. No further data access committee will be established. If sensitive data is exchanged, the beneficiaries will use trusted communication channels (such as encrypted e-mail, trusted file transfer applications, or secure delivery via traditional mail) and access rights shall be invalidated and/or data shall be deleted when they are no longer needed (e.g., after a defined preservation period or as per agreements between the provider and the user of the data).

2.2.3 Metadata

Metadata will be submitted along with any data uploaded to an open-access repository and contain information as described in section 2.1 of this document. Besides the properties that are mandatory to be submitted as metadata, the FCA project aims to increase findability of the data by providing additional metadata whenever possible.

The FCA project will use a repository that ensures access to metadata through an open, free and universally implementable protocol. Furthermore, a repository will be used that ensures free access to metadata and stores metadata as long as the repository exists.

Since all datasets are expected to be provided in commonly used formats, no special software will be required to access those, and consequently, no documentation regarding the software will be included in the metadata.

2.3 Making data interoperable

All datasets that are expected to be generated in the FCA project will be provided in standard data formats (see Table 3). In the case that an uncommon format is unavoidable, a description of the used format will be included.

For datasets building onto other datasets, the links will be described and for any data that is published via an open-access repository, the datasets will be cited using the respective metadata such as the persistent identifiers.

2.4 Increase data re-use

The data that will be generated within the FCA project has the purpose of validating specific objectives and requirements defined within the project. Consequently, the utility of the data generated by the project to external users is likely to be limited. However, the project will make efforts to support re-

usability. Whenever data is uploaded and made publicly available, a description of the dataset will be provided either through metadata or, if a more comprehensive description is necessary, by providing a readme file containing all information that is necessary for someone external to the project to understand the scope of the data and be able to re-use it. That includes information on the format of the data, the context of its creation, and, if applicable, specific software that would be necessary to read and process the data.

Every project partner generating data within the project will assign a person to be responsible for data management, who will assess the quality and content of any dataset before it is uploaded to a public repository.

A license outlining the conditions for re-use will be specified as part of the metadata provided with each uploaded dataset. A license that conforms to the Open Definition will be chosen, which is in line with the obligations set out in the grant agreement.

3 Other research outputs

This document covers mainly raw research data generated in this project. Other research outputs are the following:

- Software (e.g. prototypes and simulation software) – software created or further developed by project partners within this project will not be published.
- Official deliverables – procedures for the creation, quality assurance, publication and storage are described in the deliverable D1.1 Project Management Plan (PMP). They conform to the SESAR 3 Joint Undertaking Project Handbook and the conditions laid out in the grant agreement.
- Scientific articles
- Newsletters, press releases or social media posts
- Presentations at conferences, workshops or similar events

Procedures for the management of scientific articles, news, press releases, posts on social as well as the presence at conferences or similar events are defined in the deliverable D3.1 Communication, Dissemination and Exploitation Plan (CDEP).

4 Allocation of resources

A data management lead is appointed in D1.1 Project Management Plan and will steer and monitor the data management according to the Data Management Plan throughout the project. However, each beneficiary has ownership of the data generated by them and is thus primarily responsible for managing it. That includes an internal assessment which datasets may be made publicly available and, in cooperation with the project's quality lead (appointed in D1.1 Project Management Plan), ensuring a high level of quality for the published data.

The project intends to choose a trusted repository that is free of charge for the publication of datasets. Hence, the cost for data management according to the FAIR principles is limited to labour cost due to the activities described above and will be covered by the budget reserved for WP3 (Communication, Dissemination and Exploitation).

Long-term preservation of published data will be ensured by choosing a trusted repository that guarantees longevity of the stored data. This will be further detailed in subsequent versions of the DMP. Generated research data that is not published will be stored by the responsible beneficiary for a period of at least ten years after the end of the project, in accordance with the principles of good scientific practice.

5 Data security

In accordance with the grant agreement, beneficiaries storing data on their internal servers will implement suitable measures on a technical and organizational level to guarantee that the data remains securely stored with controlled access. The Data Protection Officers from each beneficiary organisation will provide support in this task, ensuring that the processing of personal data complies with the EU regulation 2016/679.

However, as a minimum standard among the project consortium, the following rules apply:

- Data storage servers shall be protected with state-of-the-art firewalls and strong passwords to prevent unauthorized access through online connections.
- Measures shall be implemented to secure data storage servers against unauthorized physical access.
- Online access privileges shall only be granted to the minimum number of individuals and to the minimum extent necessary.
- Regular backups shall be performed on data storage servers to prevent unintentional data loss.
- Access rights shall be invalidated and/or data shall be deleted when they are no longer needed (e.g., after a defined preservation period or as per agreements between the provider and the user of the data).
- Trusted communication channels (such as encrypted email, trusted file transfer applications, or secure delivery via traditional mail) shall be exclusively utilized for data exchange between two partners.

For published data, a trusted repository will be chosen that implements appropriate measures to ensure data security as well as long-term preservation and curation.

6 Ethics

Questionnaires that contain personal data will include informed consent for data sharing and long-term preservation. Consent to personal data processing can be revoked at any time, in the same way in which it has been given. In such cases, any further data processing based on consent previously given for the relevant purpose will no longer be performed. No other ethics or legal issues related to data management were identified.

7 Other issues

Currently, the FCA project does not make use of any procedures for data management other than those described in this data management plan.