

IFAV3 - Initial Data Management Plan

Deliverable ID:	D1.2
Project acronym:	IFAV3
Grant :	101114683
Call :	HORIZON-SESAR-2022-DES-IR-01
Topic :	HORIZON-SESAR-2022-DES-IR-01-WA6-3
Consortium coordinator :	DLR
Edition date:	06 October 2023
Edition:	01.00
Status :	Final
Classification:	PU

Abstract

This document is the initial Data Management Plan of the IFAV3 project (Deliverable 1.2).

This DMP contains the data handling and data management procedures to be applied throughout the whole project, including general data collection, storage, use and access rules. Most of the data that is handled in IFAV3 will stem from validation exercises and preparatory activities, and will be used to demonstrate the benefits of IFAV3 concepts / enablers. This document does not include the procedures to manage and publish deliverables or any other communication, dissemination or exploitation material.

As this document is the initial DMP it will be updated / replaced by D1.6 in month 18 of the IFAV3 project.

Authoring & approval

Author(s) of the document

Organisation name	Date
DLR	05/10/23
NLR	05/10/23
ENAIRE	05/10/23
CRIDA	05/10/23
INECO	05/10/23
DBLUE	05/10/23
ENAV	05/10/23
INDRA	05/10/23
LFV	05/10/23
UPM	05/10/23
LIU	05/10/23
NATS	05/10/23

Reviewed by

Organisation name	Date
DLR	05/10/23
NLR	05/10/23
ENAIRE	05/10/23
CRIDA	05/10/23
INECO	05/10/23
DBLUE	05/10/23
ENAV	05/10/23
INDRA	05/10/23
LFV	05/10/23
UPM	05/10/23
LIU	05/10/23
NATS	05/10/23

Approved for submission to the SESAR 3 JU by¹

Organisation name	Date
DLR	05/10/23
NLR	05/10/23
ENAI	05/10/23
CRIDA	05/10/23
INECO	05/10/23
DBLUE	05/10/23
ENAV	05/10/23
INDRA	05/10/23
LFV	05/10/23
UPM	05/10/23
LIU	05/10/23
NATS	05/10/23

Rejected by²

Organisation name	Date

Document history

Edition	Date	Status	Company author	Justification
00.01	17/07/23	Draft	DLR	First Version
00.02	25/09/23	Draft	DLR	Version to be completed by partners
00.03	29/09/23	Final Draft	DLR / DBLUE	Final version for Review & Approval
01.00	06/10/23	Final Version	DLR	Version for submission

¹ Representatives of all the beneficiaries involved in the project

² Representatives of the beneficiaries involved in the project

Copyright statement © 2023 – DLR, NLR, ENAIRE, INECO, CRIDA, DBLUE, ENAV, INDRA, LFV, UPM, LIU, NATS. All rights reserved. Licensed to SESAR 3 Joint Undertaking under conditions.

IFAV3

INCREASED FLEXIBILITY OF ATCO VALIDATIONS – V3

IFAV3

This document is part of a project that has received funding from the SESAR 3 Joint Undertaking under grant agreement No 101114683 under European Union's Horizon Europe research and innovation programme.



Table of contents

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

List of figures

Es konnten keine Einträge für ein Abbildungsverzeichnis gefunden werden.

List of tables

Table 1: list of acronyms	6
Table 2: Solution 1 Data Summary	1
Table 3: Solution 2 Data Summary	4
Table 4: Used repositories for data publication	2
Table 5: Policies of used repositories	2

List of acronyms

Acronym	Description
ATC	Air Traffic Control
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable, Re-Usable
IFAV	Increased Flexibility of ATCO Validations

MB	Megabyte
MUAC	Maastricht Upper Area Control Centre
SESAR JU	Single European Sky ATM Research Joint Undertaking
WP	Work Package

Table 1: List of acronyms

1 Data summary

The activities performed in the IFAV3 project will include the generation / collection and use of data. The main purpose of data processing will be to provide evidence of the benefit of IFAV3 concepts and tools, in order to reach the targeted maturity level (Solution 1: V3, Solution 2: V1).

The data collected in this project can be categorized as following:

Data Cat 1) Data needed to drive the validation exercises (e.g. traffic data, airspace characteristics data, data about local procedures and constraints). In general this data is already available, and in some cases already published on various sources (open repositories, providers of specific types of data, deliverables, reports, scientific papers etc.). Cat 1 data can be further distinguished into dependent (data that follows from recent research of the same research partners) and independent data (data that is obtained from entities other than the project partners).

Data Cat 2) Data needed to drive the used technical enablers (e.g. further geographical data, data used to model controller competencies on local procedures). Depending on the concrete case, this data needs to be prepared / produced in the frame of this project.

Data Cat 3) Personal data needed to perform administrative activities, such as coordination with exercise participants, filling participant agreement forms or the concrete scheduling of exercise participants.

Data Cat 4) Data needed to measure the performance indicators in the experiments (e.g. questionnaire responses, raw data collected during a simulation, such as trajectories, workload measurement, behavioural data etc.). All this data will be produced in, short before or short after the experiments.

Data Cat 5) Processed data calculated to demonstrate concrete benefits or to proof how far validation objectives have been achieved.

In general, all data will be stored in *electronic format*. The concrete file format depends on the application that is used to generate / record / analyse the data. Access to the data will in general be restricted to *people involved in the project*. Personal data will be stored as long as needed, will be *anonymised* whenever possible (research data concerning a certain exercise participant) and will be *destroyed when no longer required*. The data will be stored on systems owned by the partners, and no cloud services from other organisations will be used for storage of experiment data.

This document does not include the procedures to manage and publish deliverables or any other communication, dissemination or exploitation material.

As a summary, the following table provides more details per solution about the foreseen use of data, the type, the data source, storage and data access procedure.

Table 2: Solution 1 Data Summary

Beneficiary	Task	Data description	Source	Format	Storage (Place / Access) or repository	Will be made available on public repository
All Exercise owners (ENAI, INDRA, NATS, DLR, NLR, LFV, LIU)	T2.3	Data Cat 3): Administrative data from exercise participants (name, company, licenses / operational experience, endorsements, email, phone number, participants id / pseudonym)	Participant	Electronic format (xlsx), or as part of participant agreement forms	Protected servers with limited access, in case of print-outs a locked office	No
DLR	T2.3	Data Cat 1): Recorded traffic samples of real air traffic as basis for scenario modelling; Data about airspace characteristics (sector data, fixes, airspace boundaries)	MUAC, or open sources	Electronic format (xml or similar)	Protected DLR servers with limited access	Not applicable, as these data sets are already public
DLR	T2.3	Data Cat 2): Data needed to model controller competence levels (recent working periods on the sector, experienced traffic volumes and complexity, experienced special occurrences); Data needed to predict minimum controller competence levels (traffic volume and complexity within a scenario, special occurrences, etc.)	MUAC	Electronic format (xml or similar)	Protected MUAC servers (based on cooperation contract DLR-MUAC)	No
DLR	T2.3	Data Cat 4): Data recorded during the validation exercise to measure the performance of the test person during the simulation runs, e.g. trajectories, workload / situation awareness measurements as well	MUAC	Electronic format (xml or similar)	Protected MUAC servers (based on cooperation contract DLR-MUAC)	No



		as data produced by an observer / assessor person				
DLR	T2.3	Data Cat 5): Data produced on the basis of processed data Cat 4), such as traffic complexity, traffic density, gain / loss in competence levels, safety impairments, etc.	DLR	Electronic Format	Protected servers with limited access	No
UPM	T2.3	Cat 5 data): Data developed from the original Cat 4 data), to assess conditions applicable in the different Common Unit Competence Scheme	UPM	Electronic Format	Protected servers with limited access	No
NATS	T2.4	Data Cat 4): Data recorded during the validation workshops to assess the performance of automation tools and any feedback and/or design considerations captured during the testing of the automation tools, e.g. workload / situation awareness measurements, feedback on defined workflow and HMI, as well as data captured in a survey during and/or post the testing session.	NATS	Electronic Format	Protected servers with limited access	No
ENAIRES	T2.3	Data Cat 1): Recorded traffic samples of real air traffic as basis for scenario modelling; Data about airspace characteristics (sector data, fixes, airspace boundaries)	ENAIRES sources	Electronic format (xml or similar)	Protected ENAIRES servers with limited access	Not applicable, as these data sets stem from other sources



ENAIRE	T2.3	Data Cat 3): Data needed to implement local sector specificities (operational procedures, size, shape, coordinations, traffic patterns, critical points, others) at IFAV tools.	ENAIRE AIP, ACC Operational Manuals	Electronic format (pdf and csv)	Public Information	It is already public
ENAIRE	T2.3	Data Cat 3): Personal data needed to filling participant agreement forms or the concrete scheduling of exercise participants.	Participant	Electronic Format	Protected servers with limited access	No
ENAIRE	T2.3	Data Cat 4): Data needed to measure the performance indicators in the experiments (e.g. questionnaire responses, raw data collected during a simulation, such as trajectories, workload measurement, etc.). All those data will be produced in, short before or short after the experiments.	ENAIRE	Electronic Format	Protected servers with limited access	No
ENAIRE	T2.3	Data Cat 5): Processed data calculated to demonstrate concrete benefits or to proof how far validation objectives have been achieved.	ENAIRE	Electronic Format	Protected servers with limited access	No



Table 3: Solution 2 Data Summary

Beneficiary	Task	Data Category according to the list above	Source	Format	Storage (Place / Access / Timeframe)	Will be made available on public repository
All Exercise contributors (DLR, NLR, DBLUE, LFV, LIU)	T3.3	Data Cat 3): Administrative data from exercise participants (name, company, licenses / operational experience, endorsements, email, phone number, participants id / pseudonym)	Participant	Electronic format (xlsx), or as part of participant agreement forms	Protected servers with limited access, in case of print-outs a locked office	No
DLR	T3.3	Data Cat 4): Data recorded during the validation exercise to estimate benefit of proposed solutions (questionnaire results, qualitative feedback)	Participant	Electronic format (xml or similar)	Protected servers with limited access, in case of print-outs a locked office	TBD
NLR	T3.3	Data Cat 4): Data recorded during the validation exercise to estimate benefit of proposed solutions (questionnaire results, qualitative feedback, bio behavioural data, performance indicators)	Participant and simulator	Electronic format (xml or similar)	Protected servers with limited access, in case of print-outs a locked office	No

2 FAIR data

2.1 Making data findable, including provisions for metadata

The data produced in IFAV3 will be made identifiable by using appropriate identifiers, using indications for their type, use and context. Whenever possible and expressive, indications for the following items should be included in the identifier and/or the metadata:

- The project acronym,
- The type of data contained in the data set,
- The location, area or ATC unit for which the data set contains data,
- The date and time frame for which the data set contains data,
- The format of the data set, when standard formats are used,
- The validation exercise to which the dataset belongs,
- The validation run to which the dataset belongs,
- The exercise participant id,
- An indication if the dataset is classified,
- The beneficiary acting as originator of the data set.

This general principle should be applied regardless of the concrete format, storage device, database or repository that is used.

However, a concrete naming convention as a standard will not be prescribed in IFAV3 as many different types of data for different purposes are collected by different partners in the project, and a standardization is neither necessary nor beneficial.

2.2 Making data accessible

2.2.1 Repository

The repositories used in the IFAV3 project are:

Teamsites / Sharepoints: will be used as a repository for the duration of the project. All project partners will have access to them when involved in the respective work package. The project partners will store all needed material regarding the implementation of the project, reports and deliverables – including small datasets (<100MB) when appropriate.

STELLAR platform: STELLAR is a repository provided by SESAR JU. Apart from uploading deliverables and managing the project, this platform might also be used to upload small datasets if useful.

The use of further public repositories, such as Zenodo, is not excluded for IFAV3 but currently not yet initiated.

2.2.2 Data

The following data sets will be made available:

Table 4: Used repositories for data publication

Beneficiary	Task	Data description	Embargos (if yes, until when)	Used repository

The table below summarizes the policy applied by the used repositories:

Table 5: Policies of used repositories

Repository	Free and standardised access protocol (Yes / No)	Any restrictions in the access (Yes / No), if yes then further specification	Will access be tracked, and are people accessing the data sets identifiable	Metadata available (yes / no)	Preservation time

2.2.3 Metadata

Information about metadata is included in Table 5.

2.3 Making data interoperable

Whenever possible and feasible, standard data formats will be used (e.g. Asterix Cat. 62 for Traffic data, or Eurocontrol data formats, or other most common and easy-to-access data formats). This will be indicated as such when publishing the data.

In all other cases, a description of the used format / ontology / vocabulary will be included. It shall be ensured that all published data are machine readable.

IFAV3 does not intend to introduce new project specific ontologies or vocabularies.

2.4 Increase data re-use

Owners of open results arising from the IFAV3 project are encouraged to release their work under a Creative Commons license, preferably Creative Commons Attribution 4.0 (CC-BY-4.0, <http://opendefinition.org/licenses/cc-by/>).

Authors of scientific publications arising from the IFAV3 project are encouraged to seek an agreement with the scientific publisher of the publication that allows the authors to (i) retain the ownership of the copyright for their work and to (ii) deposit the publication in an Open Access repository.

Whenever datasets are made freely available, a description shall be included, providing information about:

- A description of the data set,
- The format of the data set, and how to read / use it,
- The conditions, including license conditions, under which the data may be used by whom,
- When applicable, the purposes for which the data may be used,
- The context under which the data set has been created.

The publication of data sets shall be handled in a similar way as for the publication of scientific papers / articles, with the same quality assurance process if possible.

3 Other research outputs

Research outputs other than those covered in this document are:

- Official deliverables
- Publication of scientific articles
- Publication of news articles, press releases or posts on social media
- Participation in / presentations on conferences, workshops or similar events

No software or other material will be published.

The policy to be followed for official deliverables are laid down in the appropriate Grant Agreement, as well as in the available Horizon Europe Guidance Material.

The management of scientific articles, news, press releases, posts on social or the presence at a conference / workshop or similar events will be defined in D4.1 (Communication, Dissemination and Exploitation Plan).

4 Allocation of resources

Every beneficiary has reserved an appropriate amount of budget (up to 2% of total budget, at least one PM) for communication, dissemination and exploitation purposes. This includes, apart from the costs / effort needed to perform the task itself, to take part in public events or to produce any kind of publications, the costs that might be necessary to follow the FAIR principles in this research project.

As a general rule, every beneficiary is primarily responsible for the data produced and managed by them. Further roles and responsibilities with a relation to data management are listed below.

The **Project Manager** is responsible for:

- developing the data management plan and policy in cooperation with the project partners.
- monitoring data management activities.

The **Work Package / task leaders** are responsible for:

- the implementation of the data management policy in their respective WP / Tasks.
- asking partners for missing information or clarifications.
- providing input to the data management plan by analysing and summarising the WP-specific data handling procedures.

The **Quality Assurance Manager** is responsible for:

- performing a quality assurance of open data before their publication.
- keeping contact with work package leaders and exercise owners, and decide together on critical issues.

5 Data security

In principle, company-own rules and standards for a secure data storage will apply in case data is stored on company own servers.

However, as a minimum standard:

- Servers used for storing data shall be secured with state-of-the-art firewalls and passwords against unauthorized via online connections.
- Servers used for storing data shall be secured against unauthorized physical access.
- Online access rights shall be granted to the minimum required number of persons to the minimum required extent.
- Servers used for storing data shall be subject to regular backups to avoid an unintended loss of data.
- When no longer needed (e.g. after a defined period of preservation, or when prescribed by agreements between the data provider and the data user), access rights shall become invalid and/or the data shall be deleted.
- Whenever data needs to be exchanged between two partners, only trusted communication channels shall be used (encrypted email, trusted file transfer applications or sending storage devices via traditional mail).

6 Ethics

The consortium confirms that there are no ethics or legal issues in connection with data management. Informed consent for data sharing and long-term preservation will be included in questionnaires dealing with personal data. If personal data processing takes place on the basis of consent to personal data processing, a person is entitled to revoke consent to the data processing at any time, in the same way in which it has been given, and in such case any further data processing based on consent previously given for the relevant purpose will no longer be performed.

7 Other issues

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