
# Ph0P0001 AI-FSM Procedure

## Version 2.0

## Documentation Information

| | |
|---|---|
| **Contrat Number** | 101069595 |
| **Project Website** | www.safexplain.eu |
| **Contratual Deadline** | 31.03.2024 |
| **Dissemination Level** | SEN |
| **Nature** | R |
| **Author** | Javier Fernández |
| **Modified by** | Lorea Belategi |
| **Reviewed by** | Irune Agirre |
| **Approved by** | Irune Agirre |
| **Keywords** | AI, Functional Safety, FSM, AI-FSM Annex, Explainability |

This project has received funding from the European Union's Horizon Europe programme under grant agreement number 101069595.

# Table of Contents

# 1 Review / Modification History

| Version | Date | Description Change |
|---|---|---|
| **V2.0** | 15/02/2024 | Changes Applied as a result of TÜV Review 2024-01-19 |
| V1.0 | 04/12/2023 | First version after complete internal review |
| V0.2 | 01/12/2023 | Modifications and improvements based on internal review |
| V0.1 | 30/08/2023 | First draft |

*Note: Since Artificial Intelligence - Functional Safety Management (AI-FSM), utilizes templates from both the* traditional Functional Safety Management (FSM) *and its own templates, this annex distinguishes the AI-FSM templates by color-coding them in orange and the traditional FSM templates in green. Additionally, the folders' names will be enclosed in quotation marks and the files' names created from the templates are written in italics and underlined. These files' names are preceded by "REF_", which should be changed to reflect the specific safety project reference.*

# 2 Objective

This document describes the main procedure of the AI-FSM, which is complementary to the IEC 61508 SIL 3 IKERLAN's traditional FSM [1]. In particular, this document is an annex of the FSM procedure guideline. The FSM procedure guideline encompasses the main steps, actions, and safety considerations (procedures, documents, templates …) necessary to manage the information and the activities in a safety-related project. This annex offers supplementary guidance for cases where safety-critical systems integrate components based on Artificial Intelligence (AI), with a focus on Deep Learning (DL).

# 3 Scope

This document applies to all the information (procedures, documents, templates, tables) defined in the template of the AI-FSM.

# 4 Introduction

This document defines the additional steps, actions and considerations that shall be addressed in the FSM when incorporating DL components into a safety-critical system, as shown in Figure 2. The guidelines in this document are complementary to those defined in the traditional FSM. Therefore, both shall be followed when developing a DL-based safety system according to this FSM. The guidelines in this AI-FSM procedure only apply to systems corresponding to the definition of Subsection 4.1, and the documents composing the AI-FSM are the following:

- Main procedure (this document) provides a set of steps required to generate the basic structure for safety-related projects. It serves as an internal guideline for fulfilling the procedure template.
- Procedure template: This document compiles how functional safety has been assessed within the organization.
- Guidelines: These documents offer additional guidance for specific processes.
- Templates: Standard documents used to document the information consistently. They typically include examples and tables to be completed, serving as a starting point for collecting specific information. However, the proper fulfillment of these documents is subject to technical expert judgment for the specific application.
- Internal Reviews (IRs): reviews based on the activities of the left side of the safety lifecycle. The main objective is to check that the activities defined in each phase have been properly carried out, serving as a quality assurance.

The rest of this document is structured as follows:

- Section 4.1 defines the context of this procedure.
- Section 4.2 offers a concise definition of FSM, outlining the processes involved in the V-based lifecycle proposed by the Ikerlan FSM for SIL 3 applications.
- Section 5 defines a set of concepts to understand the AI-FSM, maps the AI functional safety lifecycle with the traditional V-based lifecycle, and depicts the structure of folders and documents associated with the AI-FSM.
- Section 0 collects and specifies the documents to be generated throughout the AI-FSM procedure.

## 4.1 Context

When referring to DL-based safety systems, this procedure considers the definitions of the European Aviation Safety Agency (EASA) concept paper for Machine Learning (ML) application [2], which makes the decomposition shown in Figure 1.

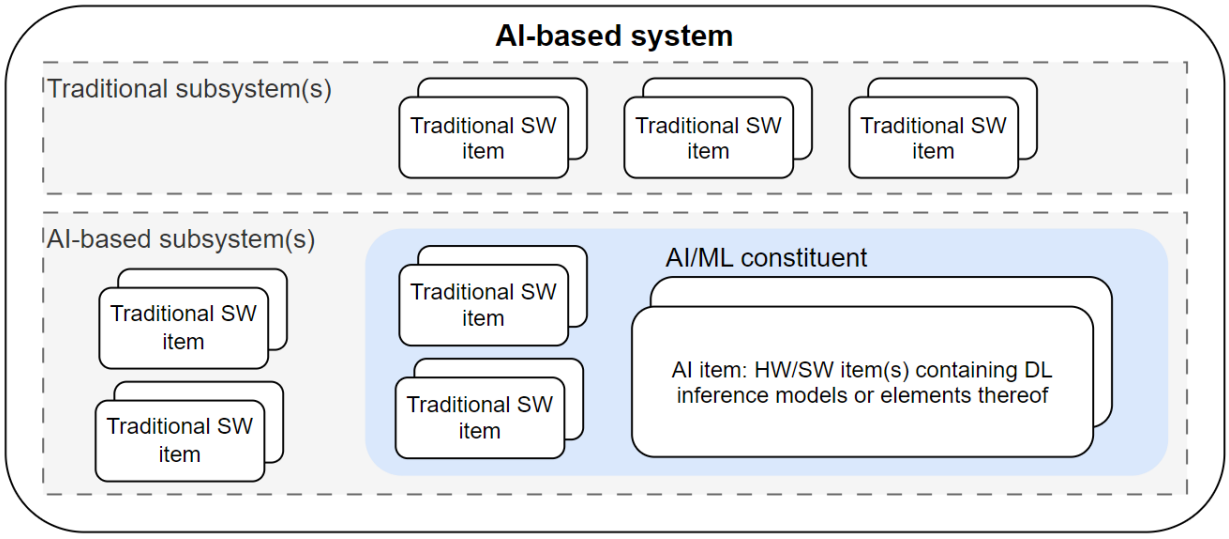


*Figure 1: AI-based system decomposition based on EASA concept paper [2]*

Based on this decomposition, the EASA concept paper makes the following definitions [2]:

- AI-based system: systems encompassing traditional subsystem(s) and incorporating at least one AI-based subsystem.
- AI-based subsystem: subsystem that involves one or more AI/ML constituents.
- AI/ML constituent: It is a combination of software and hardware items that include at least one specialized hardware or software item containing at least one ML model.
- AI/ML item: specialized hardware or software item that builds the ML model(s).
- Traditional subsystem: subsystem that does not include any ML model.
- Traditional SW/HW item: hardware or software items that do not include ML model(s).

This annex focuses on the DL constituents, a subfield of ML. As a result, we use the terms "DL constituent" and "DL item" instead of "AI/ML constituent" and "AI/ML items", respectively. The current version of this AI-FSM is restricted to DL constituents with the following features:

- DL algorithms based on supervised learning for visual perception classification tasks.
- Applications based on offline learning processes in which the model remains fixed at approval time, while excluding online learning processes.

The current version of this AI-FSM is grounded in the emerging initiatives and early stages standards existent at the time of writing, including EASA Concept Paper [2], AMLAS [3] and ISO/IEC DTR 5469 [4][1]. In the future, the AI-FSM may be updated to extend the types of AI constituents addressed and to correspondingly conform to forthcoming iterations of emerging standards, such as ISO/CD PAS 8800 [5], IEC TS 6254 [6] or the Automotive SPICE 4.0 [7], which are under development during the creation of this AI-FSM.

[1] It is worth mentioning that not all the recommendations from emerging initiatives have been adopted. For example, in the case of EASA, it proposes a "W-model" in contrast with the AI-FSM model, which does not specifically refer to linear, V-shaped, or W-shaped development lifecycles, as will be explained later.

## 4.2 IEC 61508 SIL 3 FSM

An FSM defines a development strategy that consists of a set of procedures, guidelines, and templates that define how a project with functional safety considerations should be executed (planning, involved team, activities, documents, configuration management, modification procedures, etc.). The main goal of the FSM is to ease the definition, organization, and control of the information generated during safety-critical project development while fulfilling the requirements of functional safety standards. For instance, IKERLAN's FSM [1] has proven compliance with IEC 61508 SIL 3, and hence, any new functional safety project that aims to meet with IEC 61508 up to SIL 3 can directly follow the procedures described on it and reuse the prepared templates. This FSM, referred to as "traditional FSM" in this document to differentiate from the "AI-FSM", is based on the V-model development process and structured in the following lifecycle phases depicted in Figure 2:

- Ph0 Overall Life Cycle
- Ph1 System Concept Specification
- Ph2 System Architecture Specification
- Ph3 Module Detailed Design
- Ph4 Implementation
- Ph5 Module Testing
- Ph6 Integration Testing
- Ph7 Validation Testing

It can be observed that the system development process is broken down into two different development processes that also adhere to the V-model: i) the hardware development process, and ii) the software development process.

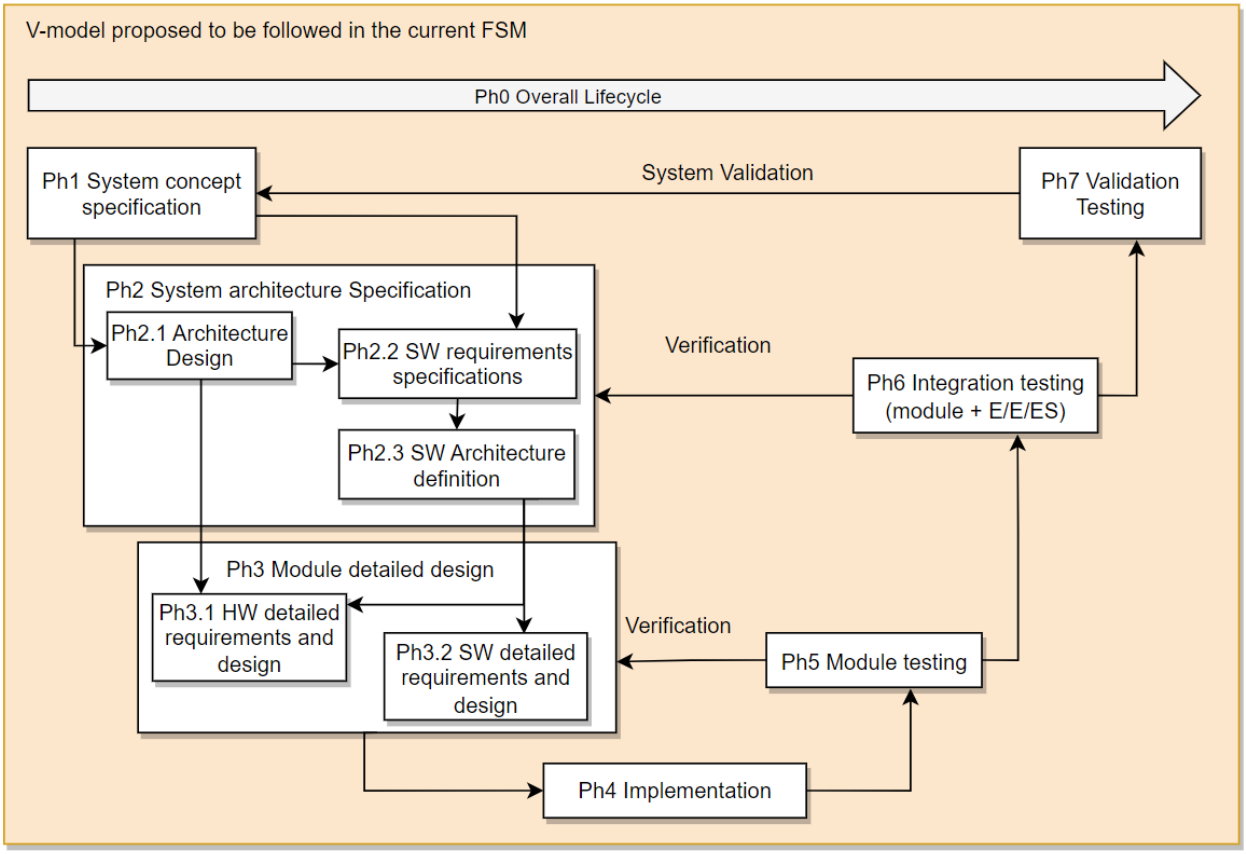


*Figure 2. V-Model followed by traditional FSM of [1].*

However, DL-based systems have some particularities concerning traditional functional safety systems that require new steps and considerations with respect to traditional safety systems. The main new challenges arise from the fact that DL systems result from data-driven learning processes, and some parts are not explicitly programmed as in traditional safety systems. This brings some new needs to the FSM, such as defining procedures for data management, dealing with sources of uncertainties, model bias, etc. [2]. These needs are covered by the AI-FSM introduced in next sections.

# 5 AI-FSM Overview

In order to structure the guidance for DL-based safety systems that match within the scope of previous Subsection 4.1, we further decompose the DL constituent as shown in Figure 2, and we differentiate between the learning stage and the inference stage.

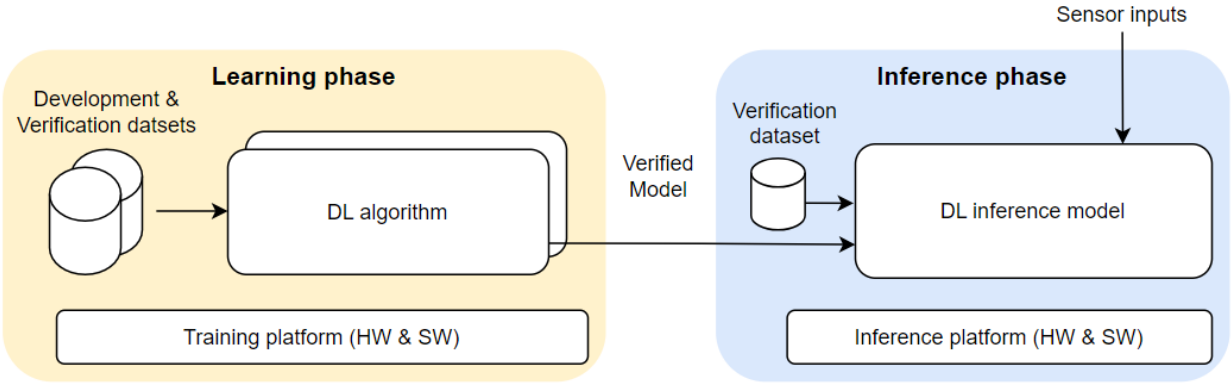


*Figure 3: DL constituent decomposition*

The main concepts of Figure 3 can be defined as follows:

- **DL algorithm**: A DL algorithm refers to the computational process that employs Neural Networks (NNs) to learn patterns or features from data. It encompasses the mathematical and computational operations involved in training a NN, adjusting its parameters (weights and biases), and optimizing its performance. DL algorithms include mechanisms like backpropagation, gradient descent, and various optimization techniques to minimize prediction errors during training. The algorithm defines the structure of the NN, the activation functions used, and how the network's parameters are updated based on the data.
- **DL inference model or DL model:** A DL model is the trained NN that has learned patterns and relationships from the training data. The model consists of the architecture, weights, and biases of the NN, which are determined during training.
- **Dataset**: In DL, a dataset refers to a collection of input data examples that are used to train, evaluate, and verify the DL model(s). These examples consist of input data and corresponding target or output values, allowing the model to learn patterns and relationships in the data to make accurate predictions or classifications in case of being employed in the training or allowing to verify the expected output once the model(s) have been trained. Datasets are a foundational component in the training and verification of DL models.
- **Training and inference platform:** The former relates to the underlying platform on which the DL model is developed, refined, and optimized using the datasets. The latter refers to the platform on which the DL model is finally deployed to make predictions.

In addition, the reader can observe two main stages at Figure 3:

1. **Learning stage**: This stage refers to the process of training a model and includes two main phases:

- **Data Management.** In DL constituents, Data Management is one of the most labor-intensive and crucial processes in DL development. This phase splits into four: i) data requirements specification, ii) data collection, iii) data preparation, and iv) data verification. Emphasize the significance of Data Management within every individual subphase. For instance, according to the data collection:
    - On one hand, the training data set establishes the behavior of the DL component, and its adequacy determines the desired behavior.
    - On the other hand, verifying dataset entail check whether the requirements defined are met. The proper identification of the cases more prone to jeopardize safety is essential.
- **Learning Management.** Learning management is performed simultaneously with Data Management. It can be decomposed into four main steps: i) model requirements specifications, ii) model design, ii) model training, iii) model evaluation and iv) model verification. This phase is performed in the training platform.

2. **Inference stage:** This stage refers to the adequation of the trained model to be implemented in the final platform where it will perform the inference:
    - **Inference Management**: Once the model has been trained, evaluated and verified, it must be deployed over the final platforms where it will perform the inference. This platform may not be the same as the one used for the training. In case of not being the same, this phase requires additional model verification.

The V-based lifecycle, as traditionally followed by FSM, has been expanded considering these concepts, as depicted in Figure 4. For improved visual distinction, the conventional lifecycle is denoted by white boxes, whereas DL components are illustrated using colored boxes. It is worth noting in Figure 4 that a sequence of numbered blue rhombuses symbolizes datasets originating from the Data Management phase. Additionally, there is a red rhombus that serves as a condition to check the results of the model evaluation. In case the model evaluation does not meet the criteria, a new iteration of the model design, model training and model evaluation steps must be performed until the model is successfully validated. These elements will be elaborated further in the forthcoming documents that comprise this AI-FSM.

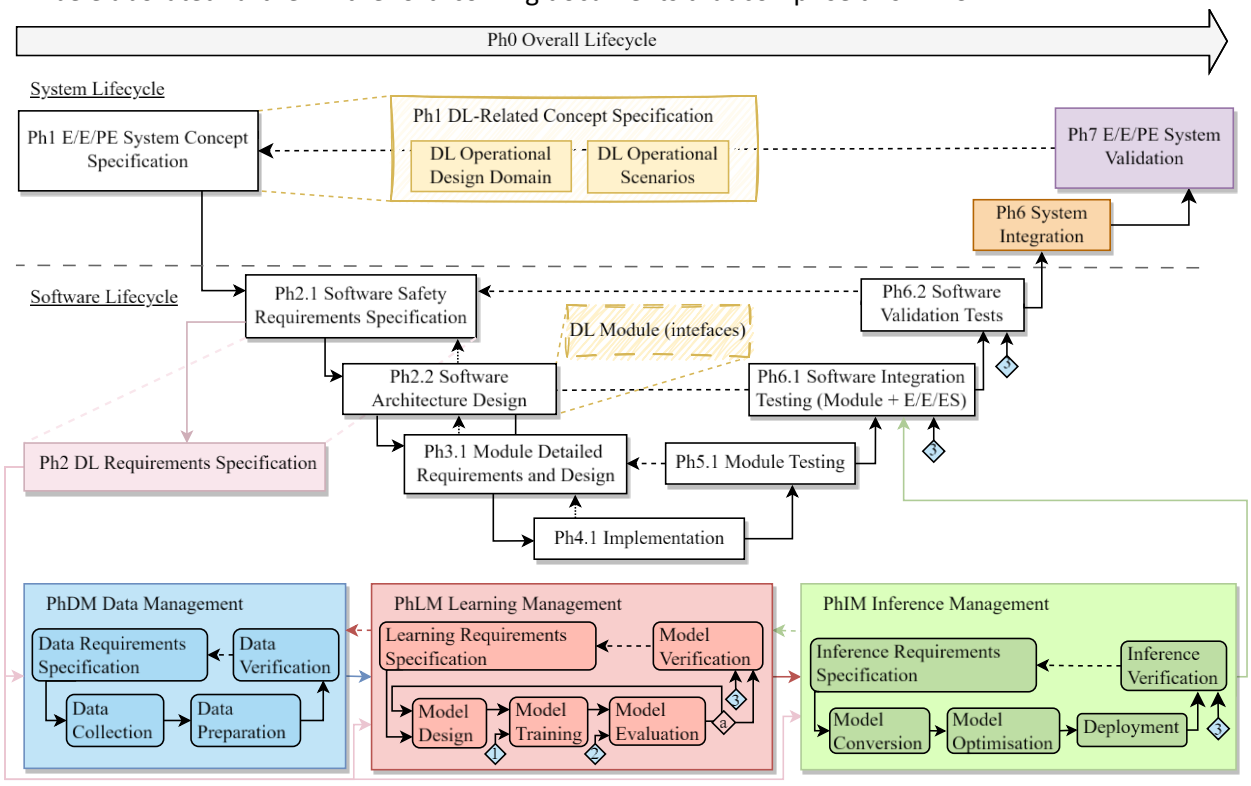


*Figure 4. Mapping AI lifecycle with traditional functional safety lifecycle*

Following the previously defined V-lifecycle this AI-FSM provides a new set of guidelines, templates, and internal review documents to complement the traditional FSM as it can be seen in Figure 5.

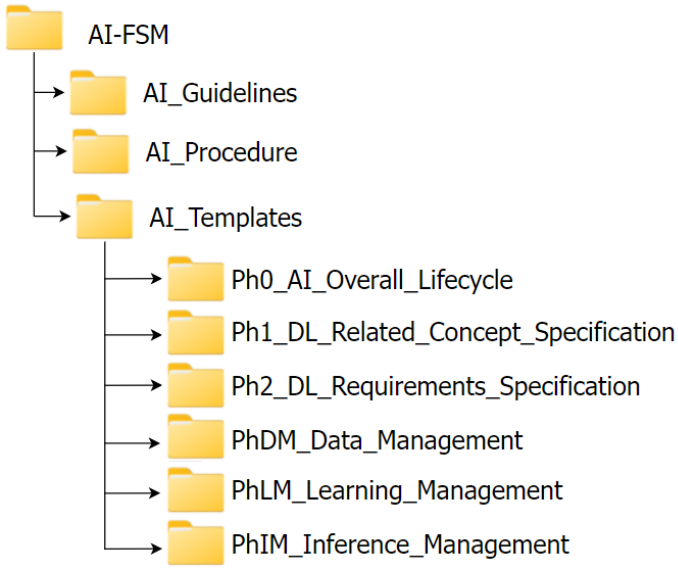


*Figure 5. Folder structure*

The structure of the documents that will be created throughout the AI-FSM and the nomenclature to denote them is defined in the *Ph0G0001_Doc_Structure.docx*. However, as explained before, in AI-FSM appear new phases that are not covered with the traditional *Ph0G0001_Doc_structure.docx* document. For these phases, we define the following identifiers:

- PhDM relates to the Data Management phase.
- PhLM relates to the Learning Management phase.
- PhIM relates to the Inference Management phase.

From this point on, this document only refers to the information or documents that differ from the traditional FSM. The rest of documents should be generated and fulfilled following the traditional FSM.

The following tables describe the inputs and outputs of each of the steps of the AI lifecycle as follows:

1. Table 1 collects the steps, inputs, outputs and templates associated with the Overall Lifecycle phase (Ph0).
2. Table 2 collects the steps, inputs, outputs and templates associated with the DL-Related Concept Specification phase (Ph1). Traditional FSM requires the definition of the software operating conditions to ensure that the safety-related system is used within the intended scope including factors such as temperature ranges, input conditions or process variables. However, within the AI domain, the array of input variables and operational scenarios can be exceptionally vast. Hence, in this phase, we incorporate the definition of the Operational Design Domain (ODD) and the operational scenarios to highlight what might require further engineering efforts.
3. Table 3 gathers the steps, inputs, outputs and templates associated with the DL Requirements Specification phase (Ph2). This includes the definition of the DL requirements and the DL component description.
4. Table 4 collects the steps, inputs, outputs and templates associated with the Data Management phase (PhDM).
5. Table 5 collects the steps, inputs, outputs and templates associated with the Learning Management phase (PhLM).
6. Table 6 collects the steps, inputs, outputs and templates associated with the Inference Management phase (PhIM).

*Table 1. Inputs and outputs of the overall lifecycle phase (Ph0)*

| Phase | Step | Inputs | Outputs | Corresponding templates |
|---|---|---|---|---|
| Ph0 AI Overall Life Cycle | Generate the AI-FSM document | *REF_FSM procedure* | *REF_Ph0D0001_AI-FSM_Procedure* | *Ph0T0001_AI_FSM_template* |
| | V&V the AI-FSM document | *REF_Ph0D0001_AI-FSM_Procedure* | *REF_Ph0D0002_AI-FSM_Procedure_IR* | *Ph0T0001_AI_FSM_template_IR* |
| | Generate the AI_Document_List | *REF_Document_list* | *REF_Ph0D0003_AI_Document_List* | *Ph0T0002_AI_Document_List_template* |
| | V&V the AI_Document_List | *REF_Ph0D0003_AI_Document_List* | *REF_Ph0D0004_AI_Document_List_IR* | *Ph0T0002_AI_Document_List_template_IR* |
| | Generate AI version tracking | *REF_version_tracking* | *REF_Ph0D0005_AI_Version_Tracking* | *Ph0T0003_AI_Version_Tracking_template* |
| | V&V the AI version tracking | *REF_Ph0D0005_AI_Version_Tracking* | *REF_Ph0D0006_AI_Version_Tracking_IR* | *Ph0T0003_AI_Version_Tracking_template_IR* |
| | Generate AI organizational chart | *REF_organizational_chart* | *REF_Ph0D0007_AI_Organizational_Chart* | *Ph0T0004_AI_Organizational_Chart_template* |
| | V&V AI organizational chart | *REF_Ph0D0007_AI_Organizational_Chart* | *REF_Ph0D0008_AI_Organizational_Chart_IR* | *Ph0T0012_Organizational_chart_template_IR* |
| | Generate the AI log of tests | - | *REF_Ph0D0009_AI_Log_of_Tests* | *Ph0T0006_Log_of_Test_template* |
| | V&V the AI log of test | *REF_Ph0D0009_AI_Log_of_Test* | *REF_Ph0D0010_AI_Log_of_Tests_IR* | *Ph0T0006_Log_of_Test_template_IR* |
| | Generate the AI selection of tools | - | *REF_Ph0D0011_AI_Tools_Selection* | *Ph0T0010_Tools_selection_template* |
| | V&V the AI selection of tools | *REF_Ph0D0011_AI_Tools_Selection* | *REF_Ph0D0012_AI_Tools_Selection_IR* | *Ph0T0010_Tools_selection_template_IR* |
| | Generate the AI traceability matrix | - | *REF_Ph0D0013_AI_Traceability_Matrix* | *Ph0T0011_Traceability_matrix_template* |
| | V&V the AI traceability matrix | *REF_Ph0D0013_AI_Traceability_Matrix* | *REF_Ph0D0014_AI_Traceability_Matrix_IR* | *Ph0T0011_Traceability_matrix_template_IR* |

*Table 2. Inputs and outputs of the DL-Related Concept Specification phase (Ph1)*

| Phase | Step | Inputs | Outputs | Corresponding templates |
|---|---|---|---|---|
| Ph1 DL-Related Concept Specification | ODD definition | REF_System_Requirements_Specifications | REF_Ph1D0001_DL_Operational_Design_Domain | Ph1T0001_DL_Operational_Design_Domain_template |
| | V&V the ODD | REF_Ph1D0001_DL_Operational_Design_Domain | REF_Ph1D0002_DL_Operational_Design_Domain_IR | Ph1T0001_DL_Operational_Design_Domain_template_IR |
| | Operational scenarios definition | REF_System Requirements Specifications<br>REF_Ph1D0001_DL_Operational_Design_Domain | REF_Ph1D0003_DL_Operational_Scenarios | Ph1T0002_DL_Operational_Scenarios_template |
| | V&V the operational scenarios | REF_Ph1D0003_DL_Operational_Scenarios | REF_Ph1D0004_DL_Operational_Scenarios_IR | Ph1T0002_DL_Operational_Scenarios_template_IR |

*Table 3. Inputs and outputs of the definition of the DL requirements (Ph2)*

| Phase | Step | Inputs | Outputs | Corresponding templates |
|---|---|---|---|---|
| Ph2 DL Requirements Specification | DL Requirements Specification | REF_Software Requirements Specifications | REF_Ph2D0001_DL_Requirements_Specifications<br>REF_Ph2D0003_DL_Requirements_Verification_Tests | Ph2T0001_DL_Requirements_Specifications_template<br>Ph0T0009_Test_definition_and_results_template |
| | | REF_Ph2D0001_DL_Requirements_Specifications<br>REF_Ph2D0003_DL_Requirements_Verification_Tests | REF_Ph2D0002_DL_Requirements_Specifications_IR<br>REF_Ph2D0004_DL_Requirements_Verification_Tests_IR | Ph2T0001_DL_Requirements_Specifications_template_IR<br>Ph0T0009_Test_definition_and_results_template_IR |

*Table 4. Inputs and outputs of each step of the Data Management phase (related to Ph3, Ph4 and Ph5 of the traditional lifecycle)*

| Phase | Step | Inputs | Outputs | Corresponding templates |
|---|---|---|---|---|
| PhDM Data Management | Data Requirements Specifications | *REF_Ph2D0001_DL_Requirements_Specifications*<br>*REF Ph1D0001_DL Operational Design Domain*<br>*REF_Ph1D0003_DL_Operational_Scenarios* | *REF_PhDMD0001_Data_Requirements_Specifications*<br>*REF_PhDMD0007_Data_Requirements_Verification_Tests* | *PhDMT0001_Data_Requirements_Specifications_template*<br>*Ph0T0009_Test_definition_and_results_template* |
| | | *REF_PhDMD0001_Data_Requirements_Specifications*<br>*REF_PhDMD0007_Data_Requirements_Verification_Tests* | *REF_PhDMD0002_Data_Requirements_Specifications_IR*<br>*REF_PhDMD0008_Data_Requirements_Verification_Tests_IR* | *PhDMT0001_Data_Requirements_Specifications_template_IR*<br>*Ph0T0009_Test_definition_and_results_template_IR* |
| | Data Collection | *REF_PhDMD0001_Data_Requirements_Specifications* | *REF_PhDMD0003_Data_Collection_Log*<br>Collected data structured in datasets[2] | *PhDMT0002_Data_Collection_Log_template* |
| | | *REF_PhDMD0003_Data_Collection_Log* | *REF_PhDMD0004_Data_Collection_Log_IR* | *PhDMT0002_Data_Collection_Log_template_IR* |
| | Data Preparation | *REF_PhDMD0001_Data_Requirements_Specifications*<br>*REF_PhDMD0003_Data_Collection_Log*<br>Raw data files structured in datasets[2] | *REF_PhDMD0005_Data_Preparation_Log*<br>Prepared data structured in datasets[2] | *PhDMT0003_Data_Preparation_Log_template* |
| | | *REF_PhDMD0005_Data_Preparation_Log* | *REF_PhDMD0006_Data_Preparation_Log_IR* | *PhDMT0003_Data_Preparation_Log_template_IR* |
| | Data Verification | *REF_PhDMD0001_Data_Requirements_Specifications*<br>*REF_PhDMD0007_Data_Requirements_Verification_Tests*<br>Datasets[2] | *REF_PhDMD0007_Data_Requirements_Verification_Tests*<br>Verified datasets[2] | *Document previously generated* |

*(*) Datasets include: i) Development (training and validation), ii) verification datasets.*

[2] Datasets include: i) Development (training and validation) datasets and ii) verification dataset.

*Table 5. Inputs and outputs of each step of the Learning Management phase (related to Ph3, Ph4 and Ph5 of the traditional lifecycle)*

| Phase | Step | Inputs | Outputs | Corresponding templates |
|---|---|---|---|---|
| PhLM Learning Management | Learning Requirements Specifications | *REF_Ph2D0001_DL_Requirements_Specifications* | *REF_PhLMD0001_Learning_Requirements_Specifications*<br>*REF_PhLMD0005_Learning_Requirements_Evaluation_Tests*<br>*REF_PhLMD0007_Learning_Requirements_Verification_Tests* | *PhLMT0001_Learning_Requirements_Specifications_template*<br>*Ph0T0009_Test_definition_and_results_template*<br>*Ph0T0009_Test_definition_and_results_template* |
| | | *REF_PhLMD0001_Learning_Requirements_Specifications*<br>*REF_PhLMD0005_Learning_Requirements_Evaluation_Tests*<br>*REF_PhLMD0007_Learning_Requirements_Verification_Tests* | *REF_PhLMD0002_Learning_Requirements_Specifications_IR*<br>*REF_PhLMD0006_Learning_Requirements_Evaluation_Tests_IR*<br>*REF_PhLMD0008_Learning_Requirements_Verification_Tests_IR* | *PhLMT0001_Learning_Requirements_Specifications_template_IR*<br>*Ph0T0009_Test_definition_and_results_template_IR*<br>*Ph0T0009_Test_definition_and_results_template* |
| | Model Design | *REF_PhLMD0001_Learning_Requirements_Specifications* | *REF_PhLMD0003_Model_Election_Log* | *PhLMT0002_Model_Election_Log_template* |
| | | *REF_PhLMD0003_Model_Election_Log* | *REF_PhLMD0004_Model_Election_Log_IR* | *PhLMT0002_Model_Election_Log_template_IR* |
| | Model Training | *REF_PhLMD0003_Model_Election_Log*<br>Training dataset | Trained Model(s) | *There is not a template, it should be considered as an implementation.* |
| | Model Evaluation | *REF_PhLMD0005_Learning_Requirements_Evaluation_Tests*<br>Trained Model(s)<br>Validation dataset [3] | *REF_PhLMD0005_Learning_Requirements_Evaluation_Tests*<br>Evaluated Model(s) | *Document previously generated* |
| | Learning Model Verification | *REF_PhLMD0007_Learning_Requirements_Verification_Tests*<br>Evaluated Model(s)<br>Verification dataset | *REF_PhLMD0007_Learning_Requirements_Verification_Test*<br>Verified Learning Model(s) | *Document previously generated* |

[3] Although this document maintains the name "validation" according to AI nomenclature, it would not correspond to validation in the context of safety

*Table 6. Inputs and outputs of each step of the inference stage (related to Ph3, Ph4 and Ph5 of the traditional lifecycle)*

| Phase | Step | Inputs | Outputs | Corresponding templates |
|---|---|---|---|---|
| ***PhIM Inference Management*** | *Inference Requirements Specifications* | *REF_Ph2D0001_DL_Requirements_Specifications*<br>*REF_PhLMD0001_Learning_Requirements_Specifications* | *REF_PhIMD0001_Inference_Requirements_Specifications*<br>*REF_PhIMD0007_Inference_Requirements_Verification_Tests* | *PhIMT0001_Inference_Requirements_Specifications*<br>*Ph0T0009_Test_definition_and_results_template* |
| | | *REF_PhIMD0001_Inference_Requirements_Specifications*<br>*REF_PhIMD0007_Inference_Requirements_Verification_Tests* | *REF_PhIMD0002_Inference_Requirements_Specifications_IR*<br>*REF_PhIMD0008_Inference_Requirements_Verification_Tests_IR* | *REF_PhIMD0002_Inference_Requirements_Specifications_IR*<br>*Ph0T0009_Test_definition_and_results_template_IR* |
| | Model Conversion | *REF_PhIMD0001_Inference_Requirements_Specifications*<br>Verified Learning Model | *REF_PhIMD0003_Model_Conversion_Log*<br>Converted Model | *PhIMT0002_Model_Conversion_Log* |
| | | *REF_PhIMD0003_Model_Conversion_Log* | *REF_PhIMD0004_Model_Conversion_Log_IR* | *PhIMT0002_Model_Conversion_Log_IR* |
| | Model Optimization | *REF_PhIMD0001_Inference_Requirements_Specifications*<br>Converted Model | *REF_PhIMD0005_Model_Optimization_Log*<br>Optimized Model | *PhIMT0003_Model_Optimization_Log* |
| | | *REF_PhIMD0005_Model_Optimization_Log* | *REF_PhIMD0006_Model_Optimization_Log_IR* | *PhIMT0003_Model_Optimization_Log_IR* |
| | Inference Model Verification | *REF_PhIMD0007_Inference_Requirements_Verification_Tests*<br>Optimized Model or Converted Model<br>Verification dataset | *REF_PhIMD0007_Inference_Requirements_Verification_Tests*<br>Verified Inference Model | *Document previously generated* |

Before delving into the details of the AI-FSM procedure, it's important to clarify some of the terminology adopted throughout the AI-FSM. This is necessary because certain terms may have different meanings or interpretations across different domains, leading to potential conflicts or misunderstandings:

- The definitions of “validation” and “verification” can vary across different technology areas or domains. In the realm of AI, “validation” typically refers to a step in the process aimed at ensuring the convergence of the developing model to terminate the AI training process. This differs significantly from the “V&V” (Verification and Validation) concepts commonly used in the functional safety community. This AI-FSM employs the definition of validation in the context of functional safety, where validation refers to the process of evaluating whether a system or component meets its intended use and satisfies the specified requirements under real operating conditions. However, this AI-FSM maintains the term “validation” dataset, although it conflicts with the functional safety definition of validation, in order to prevent misunderstanding among AI specialists.
- As will be explained later in the Learning Management phase, the "validation" dataset is used to assess the performance of the model and determine when the AI training process is complete. Since this process does not constitute verification or validation, we refer to it as evaluation.
- We encompass under the term "test" all the methods employed as verification/validation criteria, which include inspections, analysis, and actual tests. These methods are defined as follows:
    - Inspection: It takes place when a direct examination of the design documentation is enough to show that the requirement has been implemented. The requirements must be traced to the proposed design for the system. This design should be defined in the documents.
    - Analysis: This is a non-functional validation, which may include simulation, quantitative analysis, statistical analysis and comparison between the results obtained analytically and numerically. The requirements will be tested by a dedicated analysis.
    - Test: Implies a functional validation, which can verify the suitability of the implemented requirement by direct measurement. The requirements will be verified using a test procedure. A test is defined by a set of input parameters, an environment, a procedure and a set of outcomes or outputs.

# 6 AI-FSM Detailed Procedure

This section guides the user in the generation of the folders and documents to be generated and fulfilled during the development process.

Every time a new file is generated, first, it is required to replace the name of the project words in the header and in the front cover of the file with the name of the specific project, and secondly, the content (in **blue**) of the table in the Front cover (responsible of preparing, reviewing and approving the template). The corresponding revision number must be set for the specific project and the Review/Modification History table shall also be modified. The contract number, project website, contractual deadline, dissemination level (PU=Public, SEN=Sensitive) and the nature (R=Report or OTHER).

New documents generated in the AI-FSM should be consolidated within a single folder. To achieve this, within the repository of the dedicated functional safety project, generate a new folder specific to the AI-FSM with the name “AI-FSM”. In the same way than in the traditional FSM, the AI-FSM folder should be divided into subfolders according to AI life cycle phases. Therefore, within AI-FSM folder, the subsequent subfolders should be created:

1. “Ph0 AI Overall Lifecycle” folder. It will contain the documents resulting from the activities described in Section 6.1.
2. “Ph1 DL-Related Concept Specification” folder. It will contain the ODD and operational scenarios documents described in Section 6.2. These documents can be stored in the specific folder of the traditional FSM. However, to easily identify the documents relative to the AI-FSM we recommend including them in this folder.

3. "Ph2 DL Requirements Specification" folder. It will contain the documents resulting from the activities described in Section 6.3, such as the DL requirements specifications.
4. "PhDM Data Management" folder. It will contain all the information relative to the data. We refer the reader to the *PhDMG0001_Data_Management_guideline.docx* document that provides all the information related to the Data Management phase.
5. "PhLM Learning Management" folder. It will contain all the information relative to the learning process. We refer the reader to the *PhLMG0002_Learning_Management_guideline.docx* document that provides all the information related to the Learning Management phase.
6. "PhIM Inference Management" folder. It will contain all the information relative to the inference process. We refer the reader to the *PhIMG0003_Inference_Management_guideline.docx* document that provides all the information related to the Inference Management phase.

Subsection 6.1 explains the modifications to be performed in the overall lifecycle (Ph0). The new documents to be generated regarding phase 1 (Ph1) in Subsection 6.2. The documents relative to the DL Architecture Specifications phase in Subsection 6.3 and the documents associated with Data, Learning and Inference Management phases in Subsections 6.4, 6.5, and 6.6 respectively. It should be noted that the steps performed in the last three phases of the AI-FSM (PhDM Data Management, PhLM Learning Management, and PhIM Inference Management) correspond to three phases in the traditional lifecycle (Ph3 Module detailed design, Ph4 Implementation, and Ph5 Module testing), as will be explained later.

## 6.1 AI Overall Lifecycle – Phase 0 (Ph0)

The section presents the documents relative to the overall lifecycle:

**Phase Definition**

1. Create the *REF_Ph0D0001_AI-FSM_Procedure.docx* from *Ph0T0001_AI_FSM_template.docx*. This document is generated in order to specify the procedure and project specific information. The current document (*Ph0P0001_AI_Procedure.docx)* eases the generation and organization of the required information.
2. The *Document List.docx* file lists all the files generated throughout the project. In the traditional FSM, the document is generated from the *Ph0T0002_Document_List_template.docx* template. To differentiate between projects including AI and those that do not, create a new document list to gather the documents relative to AI-FSM using the *Ph0T0002_AI_Document_List_template.docx* template. This *REF_Ph0D0003_AI_Document_List.docx* file should either be merged within the *Ph0T0002_Document_List_template.docx* template from the traditional FSM or explicitly explained in the *Document List.docx* that those documents related to AI are gathered in the *REF_Ph0D0003_AI_Document_List.docx* document.

3. The *Version_Tracking.docx* file collects the relationship between the different elements of a safety project. In the traditional FSM, this document is generated from *Ph0T0001_Version_Tracking_template.docx* template, and its fulfillment is guided by *Ph0G0003_FSM_Version_Tracking_guide.docx* from the traditional FSM. To differentiate between projects including AI and those that do not, create a new version tracking document to gather the relationship relative to AI-FSM using the *Ph00T0003_AI_Version_Tracking_template.docx* template. *REF_Ph0D0005_AI_Version_Tracking.docx* document should either be merged within the *Version_Tracking.docx* from the traditional FSM or explicitly explained in the *Version tracking.docx* that those relationship between the different elements of the AI project are gathered in the *REF_Ph0D0005_AI_Version_Tracking.docx* document.

4. The *Organizational_Chart.docx* file outlines the relationship between the company organisation and the methodology, identifies the main roles involved in a safety or cybersecurity project, and the relationships between these roles. In the traditional FSM, this document is generated from *Ph0T0012_Organizational_Chart_template* template, and its fulfillment is guided by *Ph0G0004_Organizational_Chart_guide.docx* from the traditional FSM. To differentiate between

projects including AI and those that do not, create a new organizational chart document to gather the relationship relative to AI-FSM using the *Ph00T0004_AI_Organizational_Chart_template* template. *REF_AI_organizational_chart.docx* document should either be merged within the *Organizational_Chart.docx* from the traditional FSM or explicitly explained in the *Organizational_Chart.docx* that those relationship between the different participants of the AI project are gathered in the *REF_AI_Organizational_Chart.docx* document.

5. The *Log_of_Tests.docx* file collects all the tests performed during the project and is generated from the from *Ph0T0006_Log_of_Test_template* template. To differentiate between projects including AI and those that do not, create a new log of tests document to monitor all tests relative to AI-FSM using the same template than in the traditional FSM. The content of this *AI_Log_of_Tests.docx* should either be included in the *Log_of_Tests.docx* or explicitly explained in the *Log_of_Tests.docx* that those tests relatives to AI-FSM are stored in the *AI_Log_of_Tests.docx* document.

6. In the traditional FSM, the *Tools_Selection.docx* file is generated including all the tools or frameworks employed through the lifecycle of the project, using the *Ph0T0010_Tools_Selection_template.docx* template. To prevent inconsistencies or omission of information, create a *REF_Ph0D0011_AI_Tools_Selection.docx* file from the traditional template to include AI tools and frameworks. Again, this file should either be merged within the *Selection_of_Tools.docx* file from the traditional FSM or explicitly explained in the traditional *Selection_of_Tools.docx* that those related to AI are gathered in the *REF_Ph0D0011_AI_Tools_Selection.docx*.

7. In the traditional FSM, the interdependences of the requirements at different levels of the development process, as well as the relationship between requirements and verification or validation mechanisms, are documented in the *Traceability_Matrix.docx* document, using the *Ph0T0011_Traceability_Matrix_template.docx* template. The use of DL involves the apparition of the following interdependencies (as well as the testing mechanisms associated):

    a. Software requirements specifications and DL requirements specifications.
    b. DL requirements specifications and data requirements specifications.
    c. DL requirements specifications and learning requirements specifications.
    d. DL requirements specifications and inference requirements specifications.

    As before, create a *REF_Ph0D0013_AI_Traceability_Matrix.docx* file from the traditional template. This file should either be integrated into the *Traceability_Matrix.docx* file or clearly explained in the traditional *Traceability_Matrix.docx* that interdependencies related to AI are documented in the *REF_Ph0D0013_AI_Traceability_Matrix.docx.*

**Verification and Validation activities:**

- Generate the *REF_Ph0D0001_AI-FSM_Procedure_IR.xlsx*, *REF_Ph0D0004_AI_Document_List_IR.xlsx*, *REF_Ph0D0010_AI_Log_of_Tests_IR.xlsx*, *REF_Ph0D0012_AI_Tools_Selection_IR.xlsx* and *REF_Ph0D0014_AI_Traceability_Matrix_IR.xlsx* from *Ph0T0001_AI_FSM_IR.xlsx*, *Ph0T0002_Document_List_IR.xlsx*, *Ph0T0006_Log_of_Tests_template_IR.xlsx*, *Ph0T0010_Tools_Selection_IR.xlsx* and *Ph0T0011_Traceability_Matrix_IR.xlsx*, respectively.

## 6.2 DL-Related Concept Specification – Phase 1 (Ph1)

This section presents the information relatives to the System Concept Specification phase:

**Phase Definition**

The documents to be generated in the system folder are the following ones:

- Generate the *REF_Ph1D0001_DL_Operational_Design_Domain.docx* file from the *Ph1T0001_DL_Operational_Design_Domain_template.docx* template and save it with the name of the specific project.

- Generate the *REF_Ph1D0003_DL_Operational_Scenarios.docx* file from the *Ph1T0002_DL_Operational_Scenarios_template.docx* template and save it with the name of the specific project.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx.*

**Verification and Validation activities:**

- Generate the *REF_Ph1D0002_DL_Operational_Design_Domain_IR.xlsx* and the *REF_Ph1D0004_DL_Operational_Scenarios_IR.xlsx* from *Ph1T0002_DL_Operational_Design_Domain_IR.xlsx* and *Ph1T0004_DL_Operational_Scenarios_IR.xlsx*, respectively.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx.*

## 6.3 DL Requirements Specification – Phase 2 (Ph2)

This section presents the information relatives to the System Concept Specification phase:

**Phase Definition**

- Generate the *REF_Ph2D0001_DL_Requirements_Specifications.docx* file from the *Ph2T0001_DL_Requirements_Specifications_template.docx* template and save it in the repository of the specific project with the name of the file for the specific project.
- Generate the *REF_Ph2D0003_DL_Requirements_Verification_Tests.docx* file from the *Ph0T0009_Test_definition_and_results_template.docx* template and save it in the repository of the specific project with the name of the file for the specific project.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx.*
*- Update the REF_Ph0D0013_AI_Traceability_Matrix.docx.*

**Verification and Validation activities**

- Generate the *REF_Ph2D0002_DL_Requirements_Specifications_IR.xlsx* and the *REF_Ph2D0004_DL_Requirements_Verification_Tests.xlsx* internal review documents from *Ph2T0001_DL_Requirements_Specifications_IR.xlsx* and *Ph0T0009_Test_definition_and_results_IR.xlsx,* respectively.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx.*

## 6.4 Data Management – Phase DM (PhDM)

As previously mentioned, this document refers the reader to the *Ph3G0001_Data_Management_guideline.docx* for further guidance on this phase. The subsequent documents should be stored in the “PhDM Data Management” folder, located within the “AI-FSM” folder.

**Phase Definition**

- Generate the *REF_PhDMD0001_Data_Requirements_Specifications.docx* file from the *PhDMT0001_Data_Requirements_Specifications_template.docx* template and store it in the repository of the specific project with the name of the file for the specific project. This step would relate to Phase 3 in the traditional FSM.
- Generate the *REF_PhDMD0009_Data_Requirements_Verification_Tests.docx* file from the *Ph0T0009_Test_definition_and_results_template.docx* template and save it in the repository of the specific project with the name of the file for the specific project. Defining the test of this template

corresponds with Phase 3 of traditional FSM while the implementation and the collection of results correspond to Phase 5.

- Generate the "Datasets" folder to store the data relative to each dataset generated in the Data Management process. Inside this folder:
  - Generate the "Development dataset" folder and within it:
    - Generate the "Training dataset" folder.
    - Generate the "Validation dataset" folder.
  - Generate the "Verification dataset" folder.
- Inside each of the datasets:
  - Generate a "Collected Data" folder to store the raw data and predefined datasets collected during the data collection step.
  - Generate a "Prepared Data" folder to store the data after being prepared in the data preparation step.
- Generate the *REF_PhDMD0003_Data_Collection_Log.docx* document from *PhDMT0002_Data_Collection_Log_template.docx* and store it in the "PhDM Data Management" folder. This document collects information related to the description of the data collected in the project as well as information of the data generated. Completing this step is analogous to Phase 4 in the traditional FSM.
- Generate the *REF_PhDMD0005_Data_Preparation_Log.docx* file from the *PhDMT0003_Data_Preparation_Log.docx* template and store it in the "PhDM Data Management" folder. This template has been generated in order to collect all actions and decisions taken when preparing data. This file includes a guide that eases the generation and organization of the required information. Fulfilling this step would relate to the Phase 4 in the traditional FSM.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx.*
*- Update the REF_Ph0D0013_AI_traceability_matrix.docx.*

**Verification and Validation activities**

- Generate the *REF_PhDMD0002_Data_Requirements_Specifications_IR.xlsx*, *REF_PhDMD0010_Data_Requirements_Verification_Tests_IR.xlsx*, *REF_PhDMD0004_Data_Collection_Log_IR.xlsx* and *REF_PhDMD0006_Data_Preparation_Log_IR.xlsx* from *PhDMT0001_Data_Requirements_Specifications_IR.xlsx*, *Ph0T0009_Test_definition_and_results_IR.xlsx*, *PhDMT0002_Data_Collection_Log_IR.xlsx* and *PhDMT0003_Data_Preparation_Log_IR.xlsx*, respectively.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx.*

## 6.5 Learning Management – Phase LM (PhLM)

As previously mentioned, from this process we refer the reader to the *Ph3G0002_Learning_Management_guideline.docx* for further guidance. The subsequent documents should be stored in the "Learning Management" subfolder that is part of the "AI-FSM" folder.

**Phase Definition**

- Generate the *REF_PhLMD0001_Learning_Requirements_Specifications.docx* file from the *PhLMT0001_Learning_*Requirements_Specifications_template*.docx* and store it in the repository of the specific project with the name of the file for the specific project.
- Generate the *REF_PhLMD0005_Learning_Requirements_Evaluation_Tests.docx* file from the *Ph0T0009_Test_definition_and_results_template.docx* and save it in the repository of the specific project with the name of the file for the specific project.
- Generate the *REF_PhLMD0007_Learning_Requirements_Verification_Tests.docx* file from the *Ph0T0009_Test_definition_and_results_template.docx* and save it in the repository of the specific project with the name of the file for the specific project.
- Generate the *REF_PhLMD003_Model_Election_Log.docx* file from *PhLMT0002_Model_Election_log_template.docx* and save it in the repository of the specific project with the name of the file for the specific project.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx.*
*- Update the REF_Ph0D0013_AI_Traceability_Matrix.docx.*

**Verification and Validation activities**

- Generate the *REF_PhLMD0002_Learning_Requirements_Specifications_IR.xlsx*, *REF_PhLMD0006_Learning_Requirements_Evaluation_Tests_IR.xlsx*, *REF_PhLMD0008_Learning_Requirements_Verification_Tests_IR.xlsx* and *REF_PhLMD0004_Model_election_log_IR.xlsx* from *PhLMT0001_Learning_Requirements_*Specifications*_IR.xlsx, Ph0T0009_Test_definition_and_results_IR.xlsx, Ph0T0009_Test_definition_and_results_IR.xlsx and PhLMT0002_Model_Election_log_IR.xlsx*, respectively.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx*

## 6.6 Inference Management – Phase IM (PhIM)

As it was previously mentioned, we refer the reader to the *PhIMG0003_Inference_Management_guideline.docx* for further guidance on this process. The subsequent documents should be stored in the "Inference Management" subfolder, located in the "AI-FSM" folder.

**Phase Definition**

- Generate the *REF_PhIMD0001_Inference_Requirements_Specifications.docx* file from the *PhLMT0001_Inference_*Requirements_Specifications_template*.docx* and save it in the repository of the specific project with the name of the file for the specific project.
- Generate the *REF_PhIMD0007_Inference_Requirements_Verification_Tests.docx* file from the *Ph0T0009_Test_definition_and_results_template.docx* and save it in the repository of the specific project with the name of the file for the specific project.
- Generate the *REF_PhIMD003_Model_Conversion_Log.docx* file from *PhIMT0002_Model_Conversion_Log_Template.docx* and save it in the repository of the specific project with the name of the file for the specific project.
- Generate the *REF_PhIMD005_Model_Optimization_Log.docx* file from *PhIMT0003_Model_Optimization_Log_template.docx* and save it in the repository of the specific project with the name of the file for the specific project.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx.*

*- Update the REF_Ph0D0013_AI_Traceability_Matrix.docx.*

**Verification and Validation activities**

- Generate the *REF_PhIMD0002_Inference_Requirements_Specifications_IR.xlsx*, *REF_PhLMD0004_Model_Conversion_Log_IR.xlsx*, *REF_PhLMD0006_Model_Optimization_Log_IR.xlsx* and *REF_PhLMD0008_Learning_Requirements_Verification_Tests_IR.xlsx* and from *PhIMT0001_Learning_Requirements_*Specifications*_IR.xlsx, PhIMT0002_Model_Conversion_Log_IR.xlsx, PhIMT0003_Model_Optimization_Log_IR.xlsx and Ph0T0009_Test_definition_and_results_IR.xlsx*, respectively.

*Reminder: -Update the state of REF_Ph0D0003_AI_Document_List.docx*

# 7 Acronyms and Abbreviations

Below is a list of acronyms and abbreviations employed in this document:

- AI – Artificial Intelligence
- AI-FSM – Artificial Intelligence - Functional Safety Management
- DL – Deep Learning
- EASA – European Aviation Safety Agency
- FSM – Functional Safety Management
- ML – Machine Learning
- NN – Neural Network
- ODD – Operational Design Domain
- V&V – Verification and Validation

# 8 Bibliography

[1] T. Gantevoort, «Functional Safety Management certificate related to IEC 61508 Parts 1-7:2010 - Phase 10 (E/E/PE safety related Systems Realisation), Certified Company: IKERLAN Technological Research Centre, Certificate No. 968/FSM 138.01/16,» TÜV Rheinland Industrie Service GmbH Automation and Functional Safety, 2016.

[2] European Union Aviation Safety Agency (EASA), «EASA Concept Paper: guidance for Level 1 & 2 machine learning applications,,» 2023.

[3] H. Richard, P. Colin, P. Chiara, J. Yan, C. Radu y H. Ibrahim, «Guidance on the Assurance of Machine Learning in Autonomous Systems (AMLAS),» Assuring Autonomy International Programme (AAIP), University of York, 2021.

[4] ISO/IEC JTC 1/SC 42 Artificial intelligence, «ISO/IEC DTR 5469 Artificial intelligence — Functional safety and AI systems,» 2023.

[5] ISO/TC 22/SC 32 Electrical and electronic components and general system aspects, «ISO/CD PAS 8800 Road Vehicles — Safety and artificial intelligence,» 2023.

[6] IEC, «IEC TS 6254 - Information technology — Artificial intelligence — Objectives and approaches for explainability of ML models and AI systems, Under Development».

[7] VDA QMC Working Group 13, «Automotive SPICE Process Assessment / Reference Model,» 06 06 2023. [En línea]. Available: https://vda-qmc.de/wp-content/uploads/2023/06/Automotive-SPICE-PAM-40-Gelbbandrelease.pd. [Último acceso: 09 2023].