

FAIRiCUBE – F.A.I.R. INFORMATION CUBES

Work Package 2: Use
Deliverable 2.5: Use Case Validation

Deliverable Lead: space4environment
Deliverable due date: 31/07/2025

Version: 1.6
2025-08-06

Document Control Page

Document Control Page	
Title	Deliverable D2.5 Use Case Validation
Creator	space4environment
Description	The deliverable describes the different validation steps that UC must undertake to get reliable information on the output quality.
Publisher	"FAIRiCUBE – F.A.I.R. information cubes" Consortium
Contributors	S4E, 4SF, NIL, NHM, WER
Date of delivery	31/07/2025
Type	R — Document, report
Language	EN-GB
Rights	Copyright "FAIRiCUBE – F.A.I.R. information cubes"
Audience	<input checked="" type="checkbox"/> Public <input type="checkbox"/> Confidential <input type="checkbox"/> Classified
Status	<input type="checkbox"/> In Progress <input type="checkbox"/> For Review <input checked="" type="checkbox"/> For Approval <input type="checkbox"/> Approved

Revision History			
Version	Date	Modified by	Comments
0.1	22/03/2024	Mirko Gregor	Creation of template and structure
0.2	21/05/2024	Mirko Gregor	First draft
0.3	07/06/2024	Maria Ricci	Updated 3.2; added 3.2.1
1.0	21/06/2024	Jaume Targa	Internal review and format checking
1.1	24/06/2024	Maria Ricci	Added chapter 3.2.6
1.2	25/06/2024	Jaume Targa	Further review and format check
1.3	26/06/2024	Maria Ricci	Final version
1.4	27/06/2025	Maria Ricci	Reopened document for update
1.5	04/08/2025	Maria Ricci	Added Chapter 5
1.6	06/08/2025	María Colina	Further review and format check



Disclaimer

This document is issued within the frame and for the purpose of the FAIRiCUBE project. This project has received funding from the European Union's Horizon research and innovation programme under grant agreement No. 101059238. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the European Commission.

This document and its content are the property of the FAIRiCUBE Consortium. All rights relevant to this document are determined by the applicable laws. Access to this document does not grant any right or license on the document or its contents. This document or its contents are not to be used or treated in any manner inconsistent with the rights or interests of the FAIRiCUBE Consortium or the Partners detriment and are not to be disclosed externally without prior written consent from the FAIRiCUBE Partners. Each FAIRiCUBE Partner may use this document in conformity with the FAIRiCUBE Consortium Grant Agreement provisions.

Table of Contents

Document Control Page	2
Disclaimer	3
Table of Contents	4
List of Figures & Tables.....	5
1 Context	6
1.1 Overall objective of WP2	6
1.2 Description of WP2 work	6
1.3 Description of Task 2.6	6
2 Introduction	7
2.1 FAIRiCUBE and its Use Cases.....	7
2.2 Quality control and validation under UCs' context	9
3 Use Case Validation.....	10
3.1 Context.....	10
3.2 Elements to be validated	10
3.2.1 Use case specification	10
3.2.2 Assessment of usability/fitness-for-purpose	11
3.2.3 Data pre-processing and ingestion	12
3.2.4 Data processing, machine learning applications	14
3.2.5 Machine learning output/data and information sharing	14
3.2.6 AI ethics assessment.....	14
4 Validation protocol checklist.....	16
5 Implementation of Use Case validation	18
5.1 Example: UC5 Validation tracker	18
Annex I: Validation certificates	20



List of Figures & Tables

Figure 1 : Data quality control workflow template	13
Figure 2 : UC implementation steps that must be validated	16
Table 1 : UC validation checklist.	17
Table 2 : UC5 Validation tracker.	19

1 Context

1.1 Overall objective of WP2

The overall objective of Work Package 2: Use (WP2) is to ensure efficient execution of the Use Cases (UCs), assuring those potential synergies pertaining to both data and processing are identified and leveraged.

1.2 Description of WP2 work

WP2 focuses on the execution and cross-coordination of the different UCs on the FAIRiCUBE. Acting with an “outsider” role, the supervision crosscuts through all Use Case (UC) activities ensuring harmonisation with both upstream (data sources, ingestion, and processes) and downstream (results, promotion, and distribution of outputs) activities.

1.3 Description of Task 2.6

Controlling the quality of each UC is essential to ensure the success of the WP. This task will assure that all UCs address quantifiable goals, and that the interpretation of the results are in line with scientific methodology. This task will ensure that the outcomes of each UC are validated along the different steps from potential data pre-processing and ingestion over data processing to the final results and their representation. The task will validate that the UC objectives are met, the approach was followed, and the researched questions have been attempted to be answered. This deliverable will summarise the various single validation steps that are planned in the context of the FAIRiCUBE UCs and provide a general protocol for their overall quality control.

2 Introduction

2.1 FAIRiCUBE and its Use Cases

Implementing FAIRiCUBE leans on two main pillars. The first pillar is the development and provisioning of the FAIRiCUBE Hub which is a crosscutting platform and framework for data ingestion, provision, analysis, processing, and dissemination. The second pillar consists of the five Use Cases (UCs) that have been designed to illustrate how data driven projects can benefit from cube formats, infrastructure, and computational benefits.

The main role of the UCs in the context of the FAIRiCUBE project and the development of the FAIRiCUBE Hub is to test the system (from data ingestion and registration over data storage, manipulation and the development of the ML application to the visualisation of the outcomes), demonstrate what works and what does not and, thus, help to finetune the different elements of the FAIRiCUBE Hub during their development. While details of the UCs are already described in other deliverables (see list at the end of the Chapter), it is nevertheless beneficial for the further understanding of this report to provide a short overview of the objectives, data analysis plan and outcome expectations.

- UC 1: Urban adaptation to climate change
 - Cities face numerous challenges in combating climate change, including mitigating the Urban Heat Island effect, adapting to shifting precipitation patterns, and addressing urban biodiversity loss exacerbated by human activities.
 - Efforts are underway at both the European and local levels to address these challenges through comprehensive data collection and tailored strategies.
 - On the European scale, data-driven analyses such as cluster analysis help identify cities with similar characteristics and inform decision-making on adaptation strategies.
 - At the local level, cities prioritize the implementation of concrete actions guided by reliable data to mitigate climate impacts effectively.
 - Initiatives like data cubes offer promise in consolidating diverse datasets and providing stakeholders with customized information, with platforms like the FAIRiCUBE Hub poised to support experts in generating tailored solutions for immediate implementation.
- UC 2: Agriculture and Biodiversity Nexus
 - Investigation of farming activities' impact on biodiversity within agricultural landscapes by using the concept of the Dutch Biodiversity Monitor (DBM) for standardized biodiversity assessment. Enhance understanding of the relationship between agricultural practices and biodiversity at a large scale
 - Focus on identifying correlations and causal relationships between farm activities and biodiversity changes.
 - Application of interpretable AI and Causal Machine Learning to attribute biodiversity changes to specific agricultural practices.
 - Implementation on the FAIRiCUBE Hub for data collection, analysis, and accessibility, raising awareness among stakeholders in smart agriculture and biodiversity domains about data cubes and AI.

- Exploration of data cube-based infrastructure for improved access to biodiversity-related information. Refine biodiversity estimates within a spatial context and inform decisions on nature-inclusive practices.
- Provision of analysis tools within FAIRiCUBE for extracting causal relationships across different locations and questions.
- UC 3: Environmental Adaptation Genomics in *Drosophila*
 - Aims to provide insights into evolutionary dynamics and inform conservation strategies by understanding how organisms might respond to ongoing environmental changes. Intersects quantitative environmental and genomic datasets to understand how climatic and human-induced variations impact genetic diversity in *Drosophila melanogaster*.
 - Utilizes *D. melanogaster*, a widely studied genetic model organism with a global distribution, facilitating quantitative analysis of environmental influences on genetic variation. Integrates genomic data from diverse populations across different environments with high-resolution geospatial data to study ecological factors affecting local adaptation and identify genomic targets of selection.
 - Applies established population genetics theories and approaches to infer evolutionary history and predict future responses of *Drosophila* populations to changing environments.
 - Investigates the interplay between environmental factors and genetic diversity within *D. melanogaster* populations to uncover mechanisms driving local adaptation.
- UC 4: Spatial and temporal assessment of neighbourhood building stock
 - Buildings account for approximately 40% of the EU's energy demand and 36% of its greenhouse gas emissions, highlighting the importance of reducing their environmental impact.
 - Policy initiatives such as the "Renovation wave strategy" and "fit for 55" aim to enhance the energy efficiency and sustainability of European building stocks.
 - A transition to circular use of building materials is crucial to mitigate environmental impacts and ensure resilience to supply chain disruptions.
 - Varied levels of data clarity exist for buildings, with newly constructed ones often having detailed information compared to older structures.
 - UC 4 aims to develop models using FAIR-compliant data to estimate material use intensity, energy performance, and greenhouse gas emissions of building stocks.
 - These models will facilitate informed decision-making at the national level, prioritizing investments and promoting sustainable building practices.
- UC 5: Validation of Phytosociological Methods through Occurrence Cubes
 - Phytosociology classifies vegetation communities based on species cover but lacks full explanation of community formation due to limited consideration of environmental conditions.
 - The main objective is to validate traditional phytosociological methods by linking distribution data of plant species from sources like GBIF and botanical collection platforms.
 - Additionally, the aim is to develop a new phytosociological approach using satellite and occurrence data to predict the presence of plant communities in unknown areas.
 - Integration of distribution data with environmental factors enhances understanding of vegetation community formation and classification.
 - This UC contributes to improved conservation and management strategies by advancing the ability to characterize and predict plant communities.
 - FAIRiCUBE offers a framework for ground-truthing by comparing known environmental factors with areas where plant communities occur.

More details on the UCs can be explored in the deliverables of WP2 mainly dealing with their data needs, data analysis plans and synergies between UCs regarding ingestion and processing, and WP3 looking at the Machine Learning approaches and their implementations. Key deliveries to provide context to this report are:

- D2.1 Report on UC data sources
- D2.2 Report on data analysis plan
- D3.1 UC Exploratory data analysis
- D3.2 Machine learning strategy specific for each use case.

2.2 Quality control and validation under UCs' context

Quality control (QC) refers to the application of methods or processes that determine whether data, processing outputs or infrastructure tools meet overall quality goals and defined quality criteria. To determine whether something is 'good' or 'bad' - or to what degree they are so - one must have a set of quality goals and specific criteria against which data and outputs are evaluated.¹ In general, three different types of quality can be distinguished:

- Thematic quality guarantees that the thematic correctness of the results of the data processing chains will meet the quality requirements. The principal indicators used to assess thematic quality of the results are standardized quality measures (such as overall, user's and producer's accuracy etc.) based on a comparison of the results and independent reference data summarized in a form of validation protocol. This is the standard implemented in EO processing projects.
- Technical quality guarantees that the technical characteristics of the processing outputs agree with the technical specification of the given product (e.g. pixel or grid size, coordinate system, acceptable ranges of the thematic variables, etc.). Technical characteristics of any output data layer produced during operational use of the given data processing chain are finally compared with the corresponding product specification list to check that the output layer meets the technical specifications.
- Scientific quality assesses the developed processing chains from the scientific perspective. It guarantees that the used workflows and processing methods agree with the current state-of-the-art.

Quality control can be qualitative or quantitative, whereby the qualitative systematic accuracy assessment consists of a systematic qualitative survey conducted as a preceding step to the statistically rigorous quantitative accuracy assessment. The qualitative systematic accuracy review can already provide feedback to the production or development team and help to improve the processing chain at an early stage. In addition, quality control should also be able to answer the question whether the products are fit-for-purpose or whether there are any limitations of the products with respect to their intended uses. This last step can be complemented by qualitative checks by user organisations.

¹ <https://www.usgs.gov/data-management/quality-control-qc-detecting-and-repairing-data-issues-recommended-practices#:~:text=By%20Data%20Management,quality%20criteria%20for%20individual%20values>

3 Use Case Validation

3.1 Context

This deliverable will summarise the various single validation steps that are planned in the context of the FAIRiCUBE UCs and provide a general protocol for their overall quality control. The FAIRiCUBE project foresees several validation steps along the way, i.e., from smaller validation tasks (e.g., data ingestion validation or processing validation) to bigger ones (e.g., the validation of the FAIRiCUBE Hub or the validation of the entire FAIRiCUBE project as such). The UC validation sits somewhere in the middle between the smaller and larger validation activities as it subsumes the smaller tasks that cover the QC of single steps that the UCs must undertake. Strictly speaking, a full UC validation will only be possible once all single steps have been implemented and carried out, i.e., it also includes the validation of the FAIRiCUBE Hub which contains, e.g., the visualisation or final data provision to the users. Some of the validation steps are documented in dedicated deliverables. These are, next to this overall description of the UC validation, the following:

- D5.3 Validation of data ingestion routines
- D3.6 Validation of processing and ML applications
- D4.6 Validation of sharing
- D6.11 AI ethics assessment

Since UCs go through most or all steps of the FAIRiCUBE chain (i.e., ingestion, data processing/machine learning, output/sharing), this UC validation protocol can be understood as the overarching document for the entire QC chain when executing a UC via the FAIRiCUBE Hub. It might, however, be more correct to call it a UC auditing, as the protocol itself does not deal with the validation of the UCs, but rather controlling that the different steps have been implemented and the UC is conformed to its main questions and the user requirements.

3.2 Elements to be validated

Each FAIRiCUBE UC follows its own logical flow of processes to convert data into information. However, several of the processing steps fall under the same headings which are, by consequence, applicable to all UC implementations. The following subsections provide an overview of the validation of these steps, with reference to the deliverables in which they are explained in detail. Chapter 4 presents the final schematic protocol of the UC validation.

3.2.1 Use case specification

The first step for a UC is to define the specifications about objectives, target users, resources (data and processing), methods, and final product. Validation of the UC specifications ensures that UCs have a clearly defined plan before starting to collect, process and analyse the data. A priori statement of these specifications enhances trustability in the UC results. The following requirements should be met for a successful UC:

- **Clear goal defined:** the problem to be addressed and its relevance is stated in a precise way. The objective is specific, measurable, achievable, realistic and time-bound (SMART principles)
- **Target users identified:** the users or user groups who are going to benefit from the outcomes of the UC have been identified and contacted; the goals have been adjusted to the user needs and feedback.
- **Required datasets identified:** the most suitable datasets to reach the goals have been identified; an estimate of the required resources for acquiring and storing them has been made.
- **Required processing and ML/AI approaches identified:** the best available processing and ML/AI approaches to reach the defined goals have been identified; an estimate of the required processing resources has been made.
- **Workflow designed:** how the processing and ML/AI resources are applied to the datasets have been outlined; these steps are optimally documented by means of a workflow diagram; regular updates of the workflow are carried out.
- **Visualization of outputs designed**

Finally, the UC specifications should be transparently documented and communicated through the designated outlets (e.g. Project website, GitHub repository).

3.2.2 Assessment of usability/fitness-for-purpose

Product quality² is not an absolute measure, but a relative one because it depends on the intended use of the product. The following collection of criteria helps to increase the general understanding on the users' perspective on product quality, and product design should aim at covering as much of it as is feasible and realistic:

- Products should support the users' work, i.e., they should clearly address the policy or thematic area within which the respective user operates. This is a prerequisite for products to be included in the user's working practices, e.g., support decision making or be applicable within (compulsory) monitoring.
- Production should be service-oriented, i.e., put in place a transparent service chain to be able to completely, and in detail, retrace all the steps of alteration that were applied to the original data. It should be possible for users to be involved in their development if they are technically capable to do so.
- Products must be reliable, i.e., make the production method publicly available and attach complete meta-data; moreover, the QC of the final products should be independently executed, and results published.
- Products must be applicable, i.e., they should be fit-for-purpose and appropriate within the users' work environment; this relates to both spatial and temporal coverage, the completeness, scaling, timeliness, resolution, as well as quality and the balance between quality and costs.
- Data systems with which the products are created and shared should be stable and reliable, including being interoperable.

² Description including the criteria have been taken from the Deliverable Number 2.4a - Quality criteria for GMES products prepared in the context of the FP6 project GMES Network of Users (GNU) that ran from 2007-2010 (see <https://www.copernicus.eu/en/gmes-network-users>).

Based on those criteria, a final step of the validation process should be a qualitative assessment of the usability and fit-for-purpose of the output, both when serving internal (i.e., other UCs) or external users:

- In case of a UC synergy, the UC that integrates data from another UC recognises immediately whether the data is of use (i.e., fit for purpose).
- If the processing chain is executed to serve an external user, this user should be able to provide feedback on the product usability as well.

The collection of such feedback should be standardised and formalised as much as possible, e.g., by using a questionnaire that is provided to each user. In general, the quality control and validation of the UC work will be largely qualitative as oftentimes reference data are lacking to implement a qualitative validation of the output products, only some of the processing steps can be quantified and compared to benchmarks (e.g., use of computational resources).

3.2.3 Data pre-processing and ingestion

Related deliverable: D5.3 Validation of data ingestion routines.

Short description: Data pre-processing consists of actions that are undertaken by users before or during the data ingestion process, oftentimes on their own machines, e.g., resampling or calculation of indices from raw imagery. From a quality check standpoint, preprocessing does not differ from any further processing undertaken after the data is ingested in the system. Hence the validation steps pertaining data processing (algorithm implementation validation, benchmarking and comprehensive documentation, see below) outlined in deliverable D3.6 should be applied during pre-processing.

The ingestion of new datasets into FAIRiCUBE Hub depends on the dataset nature and on the target platform (rasdaman or EOX). For a detailed explanation of the ingestion pipeline refer to deliverable D5.2 Ingestion Pipelines. The validation of ingestion is however independent of the target platform. The proposed method includes the following key aspects:

- A list of characteristics to be checked after ingesting the dataset,
- Calculation of descriptive statistics to support checking the defined characteristics,
- Spatial validation,
- An automatic anomaly detection method to identify deviations from previously ingested data, by means of attribute validation,
- Comparison of source and ingested metadata,
- Error labelling and data incorporation,
- Reporting and logging.

This approach does not require domain experts to define data quality constraints or provide valid examples. The implementation of the data ingestion validation can be partially automated by hooking "validation routines" into the ingestion pipeline. A first draft of such automated validation routine within

FAIRiCUBE is available [online](#)³. There is currently a Python file that performs quality checks on ingested raster data and a [readme file](#)⁴. And a data quality control workflow as shown in Figure 1.

Data quality control workflow

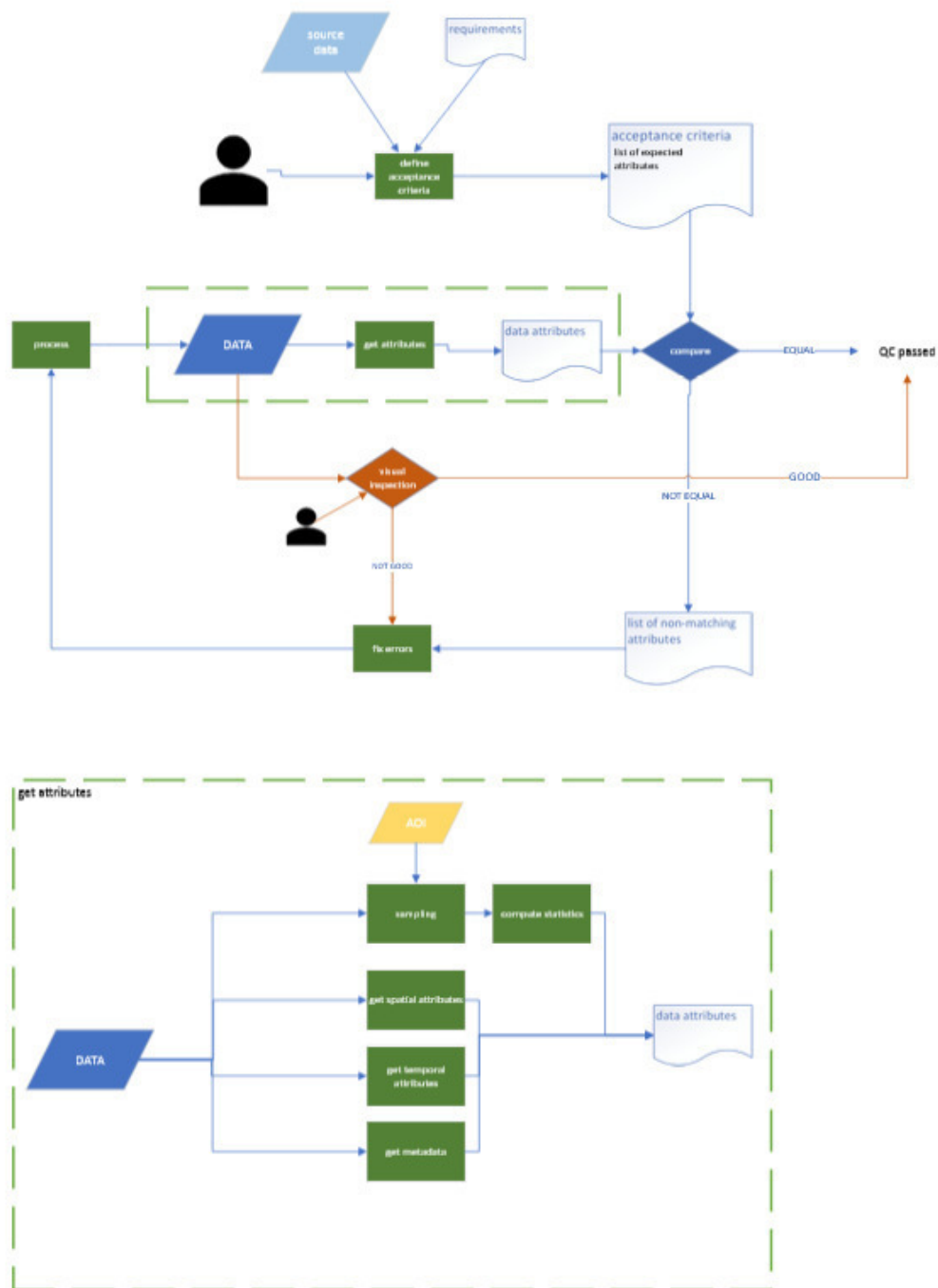


Figure 1 : Data quality control workflow template

³ https://GitHub.com/FAIRiCUBE/common-code/blob/main/quality_check/

⁴ https://GitHub.com/FAIRiCUBE/common-code/blob/main/quality_check/quality_check_readme.md

3.2.4 Data processing, machine learning applications

Related deliverable: *D3.6 Validation of processing and ML applications.*

Short description: The validation of processing and ML applications can be subdivided into data, model, performance, and ethical/bias validation, hence covers part of the value adding chain of FAIRiCUBE, i.e., algorithm implementation validation, machine learning validation, benchmarking, comprehensive documentation. The above-mentioned deliverable concludes with a validation checklist supporting users in their validation efforts.

3.2.5 Machine learning output/data and information sharing

Related deliverable: *D4.6 Validation of sharing.*

Short description: Validation of FAIRiCUBE Hub assures that all components of FAIRiCUBE Hub work correctly, both individually as well as in interaction with each other. To this purpose, validation steps have been described for all FAIRiCUBE Hub components:

- **Information:** within FAIRiCUBE Hub, information is provided both in the form of documentation (using read-the-docs) as well as via the Knowledge Base.
- **Data:** while validation of individual datasets is described in D5.3, here the focus is on the systems describing and serving data. This section includes metadata editors, catalogs and their search functionality as well as web services and APIs for data access.
- **Processing:** while validation of individual processing routines is covered in D3.6, here the focus is on the systems enabling the processing. These approaches must be verified, to assure that they allow for correct execution of scripts and tools developed within FAIRiCUBE.
- **Portrayal:** different tools are utilized by different UC for portrayal of the results of their work. These tools must be validated to assure they correctly display the data generated by the UCs.

In addition, validation of the interaction of these individual components is foreseen.

3.2.6 AI ethics assessment

Related deliverable: *D6.11 OEI - Requirement No. 2 Ethics Board review.*

Short description: In the rapidly evolving field of artificial intelligence (AI) and data science, ethical considerations have become increasingly critical. The EC mandates the use of the Assessment List for Trustworthy AI (ALTAI) to assure that any ML/AI tools created in Horizon projects align with the requirements of ethical AI. Within FAIRiCUBE, a socio-technical scenario template was created using a predefined set of questions:

- Aim of the system
- Actors
- Actors' Expectations and Motivation
- Actors' Concerns and Worries
- Context where the AI system is used



- Interaction with the AI system
- AI Technology used
- Clinical studies /Field tests/ Other Evidence
- Intellectual Property
- Legal framework
- Ethics oversight and/or approval.

These questions have been answered for all FAIRiCUBE UCs, assuring that none of the principles for trustworthy AI have been broken.

4 Validation protocol checklist

The protocol mainly serves the purpose to remind UCs (and later everyone who runs a process via the FAIRiCUBE Hub) of the different quality control and validation steps they need to take to assure that the output products are the best quality possible. The main aim is to increase trust in the output products, make sure that the products deliver what was requested and that they are scientifically sound. This checklist needs to be read together with the more detailed checklists of the single validation steps for each of the sub-processes as described in the previous chapter. To avoid repetition, the protocol/checklist below will be high-level and without too much detail.



Figure 2 : UC implementation steps that must be validated

As for now, this checklist has only been conceptionally verified with the UCs under FAIRiCUBE and is mostly meant to provide guidance on which steps to follow. Further, the guidance shall be applicable to all current and future UCs that will be executed under FAIRiCUBE and in principle every data science framework. No UC specifics are therefore included and the checklist as the simplified and potentially main output of this deliverable will not contain results of the UC validation. It is foreseen to make a webservice form from *Table 1* with an optional field for comments for each checkbox item and the guidance results with comments can be harvested and provisioned as part of the Knowledge Base services.

UC Specifications (Section 3.2.1) and User assessment / fitness-for-purpose (Section 3.2.2) are validation checks described and developed within this delivery. These required under UC validation. However, the others are validation processes developed and described in other deliverables. These, however, are key and will form part of the overall UC validation which is overarching this deliverable. These are:

- D5.3 – Validation of ingestion
- D3.6 - Validation of processing and ML applications
- D4.6 - Validation of sharing
- D6.11 - AI ethics assessment

Moreover, this report (D2.5 UC validation), together with the four reports above, are part of the overall FAIRiCUBE validation document D1.2 .

Table 1 : UC validation checklist.

UC implementation step	Check type
UC specifications	Clear Goal defined Required datasets identified Required ML/AI approaches identified Workflow designed Visualisation of outputs designed All details in Section 3.2.1
User assessment / fitness-for-purpose	Support the users' work Service orientation Reliability Applicability Data systems stability, reliability, and interoperability All details in Section 3.2.2
Data pre-processing and ingestion	List of characteristics Descriptive Statistics Calculation Spatial Validation Anomaly detection Error Labelling and Data Incorporation Reporting and Logging All details in D5.3 – Validation of ingestion
Processing and Machine Learning	Algorithm implementation validation Machine learning validation Benchmarking Comprehensive documentation All details in D3.6 – Validation of processing and ML applications
Data sharing	Information Data Processing Portrayal All details in D4.6 – Validation of sharing
AI ethics assessment	Ethics (Trustworthy AI) GDPR applicability All details in D6.11 – OEI - Requirement No. 2 Ethics Board review

5 Implementation of Use Case validation

The validation protocol was implemented in two phases to ensure a structured and transparent validation process across all use cases.

Self-Assessment Survey: First, an online self-assessment survey was distributed to the leads of each use case. This survey was designed to identify which validation steps were applicable and to capture the status of their implementation. The responses provided the basis for generating a summary in the form of validation certificates. These certificates, one per use case, are included in the annex of this deliverable.

Use Case Validation Tracker: Based on the survey results, the second phase involved completing the Use Case Validation Tracker, an extension of the validation protocol checklist that documents the implementation step, the type of check applied, and references to material used for the validation. Given the diverse nature of the validation criteria, the references vary and may include links to project documentation or sections of the software (data processing, models etc.).

The full set of Use Case Validation Trackers is published on the [FAIRiCUBE Hub website](https://hub.fairicube.eu/validation-use-cases.html)⁵. In the following section, we present the Use Case Validation Tracker for Use Case 5 as an illustrative example.

5.1 Example: UC5 Validation tracker

The table below illustrates how the Use Case Validation Tracker supports the validation of each implementation step in a Use Case. By extending the checklist with links to supporting documentation, it ensures a comprehensive and traceable validation process. Typically, the use case lead completes the tracker. To validate the *Use Case specifications*, the tracker references relevant sections of the FAIRiCUBE Digital Library. The UC5 Scrollytelling webpage confirms that the output visualization has been successfully designed. For *User assessment*, the UC lead included a user-friendly version of the model implementation. Similarly, links to other components of the processing chain validate the steps for *Data pre-processing and ingestion*, and *Machine learning*.

The tracker also uses various elements of the FAIRiCUBE Hub platform to confirm data sharing. These include the STAC metadata catalog for data and processing resources, the GitHub repository hosting the output dataset, and related publications on Zenodo. Finally, the AI ethics assessment relies on the results of the Trustworthy AI self-assessment, published on the FAIRiCUBE Hub platform. GDPR applicability is noted as not relevant for this use case.

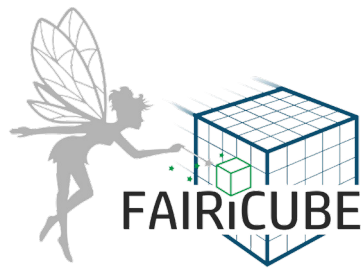
⁵ <https://hub.fairicube.eu/validation-use-cases.html>

Table 2 : UC5 Validation tracker.

UC implementation step	Check type	Reference
UC specifications	Clear Goal defined	FAIRiCUBE Digital library – UC5 – Research questions
	Required datasets identified	FAIRiCUBE Digital library - UC5 – Data retrieval
	Required ML/AI approaches identified	FAIRiCUBE Digital library - UC5 - Processing Steps and ML Applications
	Workflow designed	FAIRiCUBE Website – UC5 workflow
	Visualisation of outputs designed	UC5 Scrollytelling
User assessment / fitness-for-purpose	Support the users' work	GitHub – User friendly model
	Service orientation	
	Reliability	
	Applicability	
	Data systems stability, reliability, and interoperability	
Data pre-processing and ingestion	List of characteristics	GitHub – Data pre-processing and ingestion
	Descriptive Statistics Calculation	
	Spatial Validation	
	Anomaly detection	
	Error Labelling and Data Incorporation	
	Reporting and Logging	
Processing and Machine Learning	Algorithm implementation validation	GitHub – Model workflow GitHub – Model output statistics
	Machine learning validation	GitHub – Model validation
	Benchmarking	GitHub – Model benchmarking
	Comprehensive documentation	FAIRiCUBE Digital Library – UC5 – Processing steps and ML Applications
Data sharing	Information	FAIRiCUBE Data Catalog – GBIF Data
	Data	GitHub – Datasource table
	Processing	FAIRiCUBE Data Catalog – GBIF Data
	Portrayal	Zenodo – Poster for the GDDS event
AI ethics assessment	Ethics (Trustworthy AI)	FAIRiCUBE Hub – Validation of AI Ethics
	GDPR applicability	Not applicable

Annex I: Validation certificates

1. UC1_validation_certificate.pdf
2. UC2_validation_certificate.pdf
3. UC3_validation_certificate.pdf
4. UC4_validation_certificate.pdf
5. UC5_validation_certificate.pdf



FAIRiCUBE - Validation certificate

Validation of UC1 - In progress stage

Project Horizon Europe No 101059238

FAIRiCUBE - Validation certificate

Validation of UC1 - In progress stage

Jaume Targa, María Colina & Jaume Archilés

© 4sfera Innova
Carrer Cor de Maria 2, àtic 1a
17002 Girona
www.4sfera.com
contact@4sfera.com

Published on: 16/07/2025

Contents

Introduction	3
Validation certificate	4
1. General Use Case Validation	5
1.1. LEVEL 1	5
1.1. LEVEL 2 & 3	6
1.1.1. UC Specifications - Clear Goal defined (level 2)	6
1.1.2. UC Specifications - Required datasets identified (level 2)	6
1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)	6
1.1.4. UC Specifications - Workflow designed (level 2)	7
1.1.5. UC Specifications - Visualisation of outputs designed (level 2)	7
1.1.6. User assessment/fit-for-purpose (level 2)	7
2. Data ingestion (WP5)	8
2.1. LEVEL 1	8
2.2. LEVEL 2 & 3	8
2.2.1. Data validation checks (level 2)	8
2.2.2. Statistical Comparison of Characteristics with Source Dataset (level 2)	9
2.2.3. Characteristics Comparison with Source Dataset (level 2)	9
3. Processing and ML applications	10
3.1. LEVEL 1	10
3.2. LEVEL 2 & 3	10
3.2.1. Algorithm implementation validation (level 2)	10
3.2.2. Benchmarking (level 2)	11
3.2.3. Comprehensive documentation (level 2)	11
3.2.4. Machine learning validation (level 2)	11

Introduction

The **FAIRiCUBE project** empowers stakeholders beyond traditional Earth Observation (EO) communities by enabling them to provide, access, process, and share gridded data and algorithms in a **FAIR** (Findable, Accessible, Interoperable, Reusable) and TRUSTable manner. By leveraging Machine Learning (ML) on multi-thematic data cubes, FAIRiCUBE opens up new opportunities for governance and research institutions that have previously struggled to utilize EO resources effectively.

Central to this mission is the development of the **FAIRiCUBE Hub**—a crosscutting platform for data ingestion, analysis, processing, and dissemination across European data spaces. Through five innovative Use Cases that address EU Green Deal priorities at urban and regional scales, FAIRiCUBE demonstrates practical applications of this infrastructure:

- **UC1** – Urban adaptation to climate change
- **UC2** – Biodiversity and agriculture nexus
- **UC3** – Environmental Adaptation Genomics in *Drosophila*
- **UC4** – Spatial and temporal assessment of neighbourhood building stock
- **UC5** – Biodiversity occurrence cubes

This report presents the **validation process** established for each Use Case, focusing on both thematic consistency and the evaluation of fitness for purpose. The process involves two key components: (1) the formal specification and implementation of each Use Case, and (2) user-centered assessments to determine relevance, usability, and practical value.

By aligning technical outputs with specific user needs and the overarching goals of the EU Green Deal, this process supports the issuance of structured **validation certificates**. These certificates document the stakeholder's engagement and evaluation outcomes, ensuring traceability, transparency, and credibility in the co-creation of FAIR and TRUSTable data-driven solutions across environmental, biodiversity, and climate domains.

More information on **FAIRiCUBE validation** can be found here: <https://hub.fairicube.eu/validation.html>

Validation certificate

- **Name:** Maria Ricci
- **Date of response:** 15/05/2025 11:04:12
- **Level of validation:** UC1
- **Specific data set of process to validate:**
- **Stage of validation:** In progress stage

This validation certificate certifies the following items:

- General Use Case
- Data Ingestion
- Processing and ML Applications

1. General Use Case Validation

The **validation process** to be performed for each Use Case at a thematic level. Covers both the Use Case specification as well as wider User assessment & fit-for-purpose checks.

1.1. LEVEL 1

Table 1 highlights the key general checks on use case specifications and user assessment/fit-for-purposes.

Table 1: Validation Summary – General Checks

Check	Value
General Checks - Use Case Specifications	Clear Goal defined
General Checks - Use Case Specifications	Required datasets identified
General Checks - Use Case Specifications	Required ML/AI approaches identified
General Checks - Use Case Specifications	Workflow designed
General Checks - Use Case Specifications	Visualisation of outputs designed
General Checks - User assessment/fit-for-purpose	Support the users' work
General Checks - User assessment/fit-for-purpose	Service orientation

1.1. LEVEL 2 & 3

1.1.1. UC Specifications - Clear Goal defined (level 2)

The detailed aspects validated by the UC on the clear definition of goals are addressed in table 2

Table 2: Clear Goal Definition – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Specific - The objective is clearly defined and addresses a precise problem.		X			
Realistic - The objective is practical and relevant.	X				
Time-bound - The objective has a clear deadline.			X		

Level 3 Detail - (Additional stakeholder comments):

1.1.2. UC Specifications - Required datasets identified (level 2)

Table 3 summarises the aspects related to the required datasets identified.

Table 3: Required Datasets Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Relevant datasets	X				
Resource estimation	X				

Level 3 Detail - (Additional stakeholder comments):

1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)

Table 4 summarises the general validation of ML/AI approaches identification.

Table 4: Required ML/AI Approaches Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Suitable approaches	X				
Processing resources estimation	X				

Level 3 Detail - (Additional stakeholder comments): clustering, gap-filling, maxent, shap

1.1.4. UC Specifications - Workflow designed (level 2)

In relation to workflow designed, table 5 summarises its validation level.

Table 5: Workflow Designed – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Workflow outline	X				
Documentation			X		

Level 3 Detail - (Additional stakeholder comments):

1.1.5. UC Specifications - Visualisation of outputs designed (level 2)

Table 6 lists the status of the visualisation of outputs designed.

Table 6: Visualisation of Outputs – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Visualisation of outputs designed			X		

Level 3 Detail - (Additional stakeholder comments): scrollytelling, dashboards

1.1.6. User assessment/fit-for-purpose (level 2)

In relation to user assessment/fit-for purposes, table 7 provides level 2 validation summary for the UC.

Table 7: User Assessment – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Support the users' work - User assistance		X			
Service orientation - Transparent service chain		X			
Reliability - Method documentation				X	
Applicability - Fit-for-purpose		X			
Stable and interoperable systems		X			
Standardised feedback collection		X			

Level 3 Detail - (Additional stakeholder comments):

2. Data ingestion (WP5)

Validation process for **dataset ingested** to FAIRiCUBE. This covers data validation checks as well as statistical and spatial comparison of dataset characteristics with source dataset.

2.1. LEVEL 1

Table 8 covers the data validation checks as well as statistical and spatial comparison of dataset characteristics with source dataset.

Table 8: General Checks – Data Ingestion (WP5.2.6)

	Included
List of characteristics	X
Descriptive Statistics Calculation	X
Spatial Validation	X
Anomaly detection	X
Error Labelling and Data Incorporation	X
Reporting and Logging	X

2.2. LEVEL 2 & 3

2.2.1. Data validation checks (level 2)

Data validation checks performed, and their status are summarised in table 9.

Table 9: Data Validation Checks – WP5 (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Duplicates - Check for duplicate entries.	X				
Date overlap - Check for overlaps in the date column	X				
Date gaps - Check for missing dates between start and end date (only for timeseries).	X				
No data values - Verification of the correct use of no data.	X				
Value types - Check if data types are correct	X				
Value encoding - Check if the encoding of the data is correct	X				

Level 3 Detail - (Additional stakeholder comments):

2.2.2. Statistical Comparison of Characteristics with Source Dataset (level 2)

The proposed approach for data quality assessment and anomaly detection involves a systematic way to identify discrepancies in the ingested data. Table 10 summarises the characteristics validated.

Table 10: Statistical Comparison with Source Dataset – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Completeness	X				
Count of distinctive values	X				
Ratio of the most frequent value	X				
Maximum	X				
Minimum	X				
Standard deviation	X				
Number of records	X				
Date range	X				

Level 3 Detail - (Additional stakeholder comments):

2.2.3. Characteristics Comparison with Source Dataset (level 2)

Validating the spatial integrity of the ingested data involved checking whether the spatial information associated with the datasets has been preserved accurately during the ingestion process. Table 11 summarises the spatial validation checks performed.

Table 11: Characteristic Comparison with Source Dataset – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Grid boundaries	X				
Data completeness	X				
Projection/CRS	X				
Pixel size	X				
Number of bands	X				
Number of attributes	X				
Datatype	X				
Data format	X				
Centre coordinates and total area	X				

Level 3 Detail - (Additional stakeholder comments):

3. Processing and ML applications

Validation process on **processing and ML applications** to be performed for Use Case at an analysis and processing level. Covers data processing validation including algorithm implementation validation, comprehensiveness of documentation and benchmarking; machine learning validation; ethical and bias validation.

3.1. LEVEL 1

Table 12 highlights the key general checks on processing and ML applications.

Table 12: General Checks – Use Case Specifications

	Included
Algorithm implementation validation	X
Benchmarking	
Comprehensive documentation	X
Machine learning validation	

3.2. LEVEL 2 & 3

3.2.1. Algorithm implementation validation (level 2)

The detailed aspects validated on algorithm implementation are addressed in Table 13.

Table 13: Algorithm Implementation Validation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Technical Robustness and safety				X	
Assess the interactions				X	
End-to-end testing				X	
Cross-validation			X		

Level 3 Detail - (Additional stakeholder comments):

3.2.2. Benchmarking (level 2)

Table 14 lists the performance validation (benchmarking) aspects.

Table 14: Benchmarking – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Monitor compute resources	X				
Re-run and compare	X				

Level 3 Detail - (Additional stakeholder comments): NA

3.2.3. Comprehensive documentation (level 2)

Documentation and transparency aspects are included in table 15

Table 15: Comprehensive Documentation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Documentation and transparency		X			
Meta-data	X				

Level 3 Detail - (Additional stakeholder comments): NA

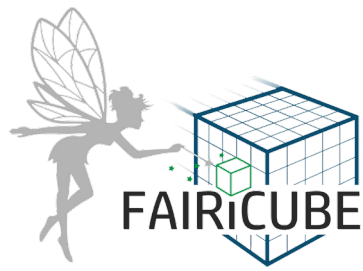
3.2.4. Machine learning validation (level 2)

Validation of Machine learning features are summarised in table 16.

Table 16: Machine Learning Validation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Dataset preparation for training	X				
Define appropriate validation metrics	X				
Prevent/Test overfitting and underfitting		X			
Statistical bias validation			X		
Human agency and oversight	X				

Level 3 Detail - (Additional stakeholder comments): NA



FAIRiCUBE - Validation certificate

Validation of UC2 - In progress stage

Project Horizon Europe No 101059238

FAIRiCUBE - Validation certificate

Validation of UC2 - In progress stage

Jaume Targa, María Colina & Jaume Archilés

© 4sfera Innova
Carrer Cor de Maria 2, àtic 1a
17002 Girona
www.4sfera.com
contact@4sfera.com

Published on: 16/07/2025

Contents

Introduction	4
Validation certificate	5
1. General Use Case Validation	6
1.1. LEVEL 1	6
1.1. LEVEL 2 & 3	7
1.1.1. UC Specifications - Clear Goal defined (level 2)	7
1.1.2. UC Specifications - Required datasets identified (level 2)	7
1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)	7
1.1.4. UC Specifications - Workflow designed (level 2)	8
1.1.5. UC Specifications - Visualisation of outputs designed (level 2)	8
1.1.6. User assessment/fit-for-purpose (level 2)	8
2. Data ingestion (WP5)	9
2.1. LEVEL 1	9
2.2. LEVEL 2 & 3	9
2.2.1. Data validation checks (level 2)	9
2.2.2. Statistical Comparison of Characteristics with Source Dataset (level 2)	10
2.2.3. Characteristics Comparison with Source Dataset (level 2)	10
3. Processing and ML applications	11
3.1. LEVEL 1	11
3.2. LEVEL 2 & 3	11
3.2.1. Algorithm implementation validation (level 2)	11
3.2.2. Benchmarking (level 2)	12
3.2.3. Comprehensive documentation (level 2)	12
3.2.4. Machine learning validation (level 2)	12
4. FAIRiCUBE Hub and services	13
4.1. LEVEL 1	13
4.2. LEVEL 2 & 3	14
4.2.1. Hub (level 2)	14

4.2.2. Documentation (level 2)	14
4.2.3. Knowledge Base (level 2)	15
4.2.4. Metadata Editors (level 2)	15
4.2.5. Data Catalog (level 2)	15
4.2.6. WCS (level 2)	16
4.2.7. S3 (level 2)	16
5. Communication and dissemination	17
5.1. LEVEL 1	17
5.2. LEVEL 2 & 3	18
5.2.1. Communication and Dissemination (level 2)	18
6. AI ethics assessment	19
6.1. LEVEL 1	19
6.2. LEVEL 2 & 3	20
6.2.1. Ethics (Trustworthy AI) (level 2)	20
6.2.2. GDPR applicability (level 2)	20

Introduction

The **FAIRiCUBE project** empowers stakeholders beyond traditional Earth Observation (EO) communities by enabling them to provide, access, process, and share gridded data and algorithms in a **FAIR** (Findable, Accessible, Interoperable, Reusable) and TRUSTable manner. By leveraging Machine Learning (ML) on multi-thematic data cubes, FAIRiCUBE opens up new opportunities for governance and research institutions that have previously struggled to utilize EO resources effectively.

Central to this mission is the development of the **FAIRiCUBE Hub**—a crosscutting platform for data ingestion, analysis, processing, and dissemination across European data spaces. Through five innovative Use Cases that address EU Green Deal priorities at urban and regional scales, FAIRiCUBE demonstrates practical applications of this infrastructure:

- **UC1** – Urban adaptation to climate change
- **UC2** – Biodiversity and agriculture nexus
- **UC3** – Environmental Adaptation Genomics in *Drosophila*
- **UC4** – Spatial and temporal assessment of neighbourhood building stock
- **UC5** – Biodiversity occurrence cubes

This report presents the **validation process** established for each Use Case, focusing on both thematic consistency and the evaluation of fitness for purpose. The process involves two key components: (1) the formal specification and implementation of each Use Case, and (2) user-centered assessments to determine relevance, usability, and practical value.

By aligning technical outputs with specific user needs and the overarching goals of the EU Green Deal, this process supports the issuance of structured **validation certificates**. These certificates document the stakeholder's engagement and evaluation outcomes, ensuring traceability, transparency, and credibility in the co-creation of FAIR and TRUSTable data-driven solutions across environmental, biodiversity, and climate domains.

More information on **FAIRiCUBE validation** can be found here: <https://hub.fairicube.eu/validation.html>

Validation certificate

- **Name:** Marian Vittek
- **Date of response:** 13/05/2025 15:38:04
- **Level of validation:** UC2
- **Specific data set of process to validate:**
- **Stage of validation:** In progress stage

This validation certificate certifies the following items:

- General Use Case
- Data Ingestion
- Processing and ML Applications
- FAIRiCUBE Hub and Services
- Communication and Dissemination
- AI Ethics Assessment

1. General Use Case Validation

The **validation process** to be performed for each Use Case at a thematic level. Covers both the Use Case specification as well as wider User assessment & fit-for-purpose checks.

1.1. LEVEL 1

Table 1 highlights the key general checks on use case specifications and user assessment/fit-for-purposes.

Table 1: Validation Summary – General Checks

Check	Value
General Checks - Use Case Specifications	Clear Goal defined
General Checks - Use Case Specifications	Required datasets identified
General Checks - Use Case Specifications	Required ML/AI approaches identified
General Checks - Use Case Specifications	Workflow designed
General Checks - User assessment/fit-for-purpose	Support the users' work
General Checks - User assessment/fit-for-purpose	Reliability
General Checks - User assessment/fit-for-purpose	Applicability
General Checks - User assessment/fit-for-purpose	Data systems stability
General Checks - User assessment/fit-for-purpose	reliability
General Checks - User assessment/fit-for-purpose	
General Checks - User assessment/fit-for-purpose	interoperability

1.1. LEVEL 2 & 3

1.1.1. UC Specifications - Clear Goal defined (level 2)

The detailed aspects validated by the UC on the clear definition of goals are addressed in table 2

Table 2: Clear Goal Definition – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Specific - The objective is clearly defined and addresses a precise problem.	X				
Realistic - The objective is practical and relevant.	X				
Time-bound - The objective has a clear deadline.	X				

Level 3 Detail - (Additional stakeholder comments):

1.1.2. UC Specifications - Required datasets identified (level 2)

Table 3 summarises the aspects related to the required datasets identified.

Table 3: Required Datasets Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Relevant datasets	X				
Resource estimation	X				

Level 3 Detail - (Additional stakeholder comments):

1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)

Table 4 summarises the general validation of ML/AI approaches identification.

Table 4: Required ML/AI Approaches Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Suitable approaches	X				
Processing resources estimation	X				

Level 3 Detail - (Additional stakeholder comments):

1.1.4. UC Specifications - Workflow designed (level 2)

In relation to workflow designed, table 5 summarises its validation level.

Table 5: Workflow Designed – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Workflow outline	X				
Documentation	X				

Level 3 Detail - (Additional stakeholder comments):

1.1.5. UC Specifications - Visualisation of outputs designed (level 2)

Table 6 lists the status of the visualisation of outputs designed.

Table 6: Visualisation of Outputs – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Visualisation of outputs designed			X		

Level 3 Detail - (Additional stakeholder comments):

1.1.6. User assessment/fit-for-purpose (level 2)

In relation to user assessment/fit-for purposes, table 7 provides level 2 validation summary for the UC.

Table 7: User Assessment – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Support the users' work - User assistance				X	
Service orientation - Transparent service chain				X	
Reliability - Method documentation		X			
Applicability - Fit-for-purpose	X				
Stable and interoperable systems		X			
Standardised feedback collection		X			

Level 3 Detail - (Additional stakeholder comments):

2. Data ingestion (WP5)

Validation process for **dataset ingested** to FAIRiCUBE. This covers data validation checks as well as statistical and spatial comparison of dataset characteristics with source dataset.

2.1. LEVEL 1

Table 8 covers the data validation checks as well as statistical and spatial comparison of dataset characteristics with source dataset.

Table 8: General Checks – Data Ingestion (WP5.2.6)

	Included
List of characteristics	X
Descriptive Statistics Calculation	
Spatial Validation	
Anomaly detection	
Error Labelling and Data Incorporation	X
Reporting and Logging	

2.2. LEVEL 2 & 3

2.2.1. Data validation checks (level 2)

Data validation checks performed, and their status are summarised in table 9.

Table 9: Data Validation Checks – WP5 (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Duplicates - Check for duplicate entries.	X				
Date overlap - Check for overlaps in the date column	X				
Date gaps - Check for missing dates between start and end date (only for timeseries).	X				
No data values - Verification of the correct use of no data.				X	
Value types - Check if data types are correct				X	
Value encoding - Check if the encoding of the data is correct				X	

Level 3 Detail - (Additional stakeholder comments):

2.2.2. Statistical Comparison of Characteristics with Source Dataset (level 2)

The proposed approach for data quality assessment and anomaly detection involves a systematic way to identify discrepancies in the ingested data. Table 10 summarises the characteristics validated.

Table 10: Statistical Comparison with Source Dataset – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Completeness	X				
Count of distinctive values				X	
Ratio of the most frequent value				X	
Maximum				X	
Minimum				X	
Standard deviation				X	
Number of records				X	
Date range				X	

Level 3 Detail - (Additional stakeholder comments):

2.2.3. Characteristics Comparison with Source Dataset (level 2)

Validating the spatial integrity of the ingested data involved checking whether the spatial information associated with the datasets has been preserved accurately during the ingestion process. Table 11 summarises the spatial validation checks performed.

Table 11: Characteristic Comparison with Source Dataset – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Grid boundaries	X				
Data completeness	X				
Projection/CRS	X				
Pixel size				X	
Number of bands				X	
Number of attributes				X	
Datatype				X	
Data format	X				
Centre coordinates and total area				X	

Level 3 Detail - (Additional stakeholder comments):

3. Processing and ML applications

Validation process on **processing and ML applications** to be performed for Use Case at an analysis and processing level. Covers data processing validation including algorithm implementation validation, comprehensiveness of documentation and benchmarking; machine learning validation; ethical and bias validation.

3.1. LEVEL 1

Table 12 highlights the key general checks on processing and ML applications.

Table 12: General Checks – Use Case Specifications

	Included
Algorithm implementation validation	X
Benchmarking	X
Comprehensive documentation	X
Machine learning validation	X

3.2. LEVEL 2 & 3

3.2.1. Algorithm implementation validation (level 2)

The detailed aspects validated on algorithm implementation are addressed in Table 13.

Table 13: Algorithm Implementation Validation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Technical Robustness and safety		X			
Assess the interactions		X			
End-to-end testing		X			
Cross-validation				X	

Level 3 Detail - (Additional stakeholder comments):

3.2.2. Benchmarking (level 2)

Table 14 lists the performance validation (benchmarking) aspects.

Table 14: Benchmarking – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Monitor compute resources	X				
Re-run and compare		X			

Level 3 Detail - (Additional stakeholder comments): NA

3.2.3. Comprehensive documentation (level 2)

Documentation and transparency aspects are included in table 15

Table 15: Comprehensive Documentation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Documentation and transparency	X				
Meta-data		X			

Level 3 Detail - (Additional stakeholder comments): NA

3.2.4. Machine learning validation (level 2)

Validation of Machine learning features are summarised in table 16.

Table 16: Machine Learning Validation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Dataset preparation for training	X				
Define appropriate validation metrics			X		
Prevent/Test overfitting and underfitting		X			
Statistical bias validation				X	
Human agency and oversight				X	

Level 3 Detail - (Additional stakeholder comments): NA

4. FAIRiCUBE Hub and services

Validation process to ensure that the components of **FAIRiCUBE Hub on information, data, processing and portrayal** each function individually, as well as jointly.

4.1. LEVEL 1

Table 17 lists the processes validated within the FAIRiCUBE Hub and services.

Table 17: General Checks – FAIRiCUBE Hub and Services

	Included
Hub	X
Documentation	X
Knowledge Base	X
Metadata Editors	X
Data Catalog	X
WCS	
S3	X
Sentinel Hub	
GitHub	X
Lab	X

4.2. LEVEL 2 & 3

4.2.1. Hub (level 2)

Related to FC Hub validation, table 18 lists the status of different aspects related the FC Hub.

Table 18: FAIRiCUBE Hub Validation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Findability of components			X		
Authentication across components			X		
Interplay - Accessibility of datasets					
Interplay - Accessibility of analysis/processing resources		X			
Interplay - Accessibility of linked datasets and a/p resources				X	
Interplay - Storage and accessibility of processing results				X	
Interplay - Visualization of processing results				X	
User support				X	

Level 3 Detail - (Additional stakeholder comments):

4.2.2. Documentation (level 2)

Table 19 illustrates the status of the different aspects in relation to documentation

Table 19: Documentation – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Findability of required information				X	
Correctness of available information				X	
Completeness of available information				X	
Usefulness of available information				X	
Ease of provision of documentation				X	

Level 3 Detail - (Additional stakeholder comments):

4.2.3. Knowledge Base (level 2)

Table 20 lists the status of the validation criteria for the knowledge base.

Table 20: Knowledge Base – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Findability of required information				X	
Correctness of available information				X	
Completeness of available information				X	
Usefulness of available information				X	
Ease of provision of information for the KB			X		

Level 3 Detail - (Additional stakeholder comments):

4.2.4. Metadata Editors (level 2)

FC should ensure that Metadata Editors are effective tools for sharing and validating metadata, thereby enhancing the overall quality and usability of shared data. The validation criteria for FC metadata is listed in table 21

Table 21: Metadata Editors – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Persistence	X				
Alignment	X				
Findability			X		
Labeling			X		
Errors			X		

Level 3 Detail - (Additional stakeholder comments):

4.2.5. Data Catalog (level 2)

The Data Catalog aims to be user-friendly and efficient. Table 22 lists the status of the validated criteria.

Level 3 Detail - (Additional stakeholder comments):

Table 22: Data Catalog – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Findability			X		
Simple Search	X				
Advanced Search		X			
Navigation		X			

4.2.6. WCS (level 2)

Data access via WCS is important, the status of the validation criteria are summarised in table 23, below.

Table 23: WCS – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Correct Response					
Measure Types					
XML Validity					
Extent/Bbox					
Subset Accuracy					
Format					

Level 3 Detail - (Additional stakeholder comments):**4.2.7. S3 (level 2)**

Data access to AWS S3 bucket should be user-friendly, efficient and secure. The status of the validation criteria is summarised in table 24, below.

Table 24: S3 – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Correct Response					
Measure Types					
XML Validity					
Extent/Bbox					
Subset Accuracy					
Format					

Level 3 Detail - (Additional stakeholder comments):

5. Communication and dissemination

For **communication and dissemination** , a number of KPI's are used for validation measures.

5.1. LEVEL 1

The validation checks included in the validation for communication and dissemination are summarised in table 25.

Table 25: General Checks – Communication and Dissemination

	Included
EU policy	X
Domain specific events	X
Newsletter	X
Website	X
Social media	X
Press releases and articles published in national/regional/European online media	X
Innovation workshops	X
Trainings and master workshop	X
Scientific papers	X

5.2. LEVEL 2 & 3

5.2.1. Communication and Dissemination (level 2)

The achievement status of various areas are listed in table 26.

Table 26: Communication and Dissemination – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
EU policy					
Domain specific events					
Newsletter					
Website					
Social media					
Press releases and articles published in national/regional/European online media					
Innovation workshops					
Trainings and master workshop					
Scientific papers					

Level 3 Detail - (Additional stakeholder comments):

6. AI ethics assessment

Validation process to be performed for the Use Case to assure **relevant considerations pertaining to ethical use of AI are met**.

6.1. LEVEL 1

The validation process in relation to ethics (Trustworthy AI) and GDPR applicability includes a checklist which requires validation. Table 27 lists the AI ethics assessment checklist included in the validation.

Table 27: General Checks – AI Ethics Assessment

	Included
Fundamental rights	X
Privacy and data protection	X
Transparency rights	X
Accessibility	X
Education and tutorials	X
Data management	X
Security	X
Ease to deactivate/remove	X
Ease to acces services without using the AI system	
Open-source code	X
Ownership	X
Openness about Data governance	X
Legislation and Policy	X
Design Impact Assessment and Open Development Process	X
Right to contest/liability	X
List of data	X
Personal data	X

6.2. LEVEL 2 & 3

6.2.1. Ethics (Trustworthy AI) (level 2)

In relation to Ethics (Trustworthy AI), the status of the different aspects validated are summarised in table 28.

Table 28: Ethics – Trustworthy AI Assessment (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Fundamental rights	X				
Privacy and data protection		X			
Transparency rights				X	
Accessibility	X				
Education and tutorials				X	
Data management	X				
Security	X				
Ease to deactivate/remove	X				
Ease to access services without using the AI system				X	
Open-source code			X		
Ownership			X		
Openness about Data governance	X				
Legislation and Policy				X	
Design Impact Assessment and Open Development Process				X	
Right to contest/liability	X				

Level 3 Detail - (Additional stakeholder comments):

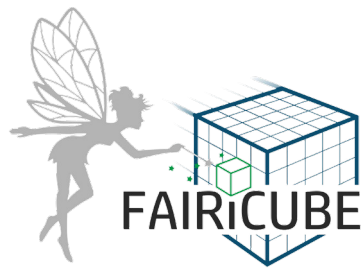
6.2.2. GDPR applicability (level 2)

In relation to GDPR applicability, the status of the different aspects validated are summarised in table 29.

Table 29: GDPR Applicability – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
List of data		X			
Personal data				X	

Level 3 Detail - (Additional stakeholder comments):



FAIRiCUBE - Validation certificate

Validation of UC3 - Initial stage

Project Horizon Europe No 101059238

FAIRiCUBE - Validation certificate

Validation of UC3 - Initial stage

Jaume Targa, María Colina & Jaume Archilés

© 4sfera Innova
Carrer Cor de Maria 2, àtic 1a
17002 Girona
www.4sfera.com
contact@4sfera.com

Published on: 16/07/2025

Contents

Introduction	3
Validation certificate	4
1. General Use Case Validation	5
1.1. LEVEL 1	5
1.1. LEVEL 2 & 3	6
1.1.1. UC Specifications - Clear Goal defined (level 2)	6
1.1.2. UC Specifications - Required datasets identified (level 2)	6
1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)	6
1.1.4. UC Specifications - Workflow designed (level 2)	6
1.1.5. UC Specifications - Visualisation of outputs designed (level 2)	7
1.1.6. User assessment/fit-for-purpose (level 2)	7
4. FAIRiCUBE Hub and services	8
4.1. LEVEL 1	8
4.2. LEVEL 2 & 3	9
4.2.1. Hub (level 2)	9
4.2.2. Documentation (level 2)	9
4.2.3. Knowledge Base (level 2)	10
4.2.4. Metadata Editors (level 2)	10
4.2.5. Data Catalog (level 2)	10
4.2.6. WCS (level 2)	11
4.2.7. S3 (level 2)	11
5. Communication and dissemination	12
5.1. LEVEL 1	12
5.2. LEVEL 2 & 3	13
5.2.1. Communication and Dissemination (level 2)	13

Introduction

The **FAIRiCUBE project** empowers stakeholders beyond traditional Earth Observation (EO) communities by enabling them to provide, access, process, and share gridded data and algorithms in a **FAIR** (Findable, Accessible, Interoperable, Reusable) and TRUSTable manner. By leveraging Machine Learning (ML) on multi-thematic data cubes, FAIRiCUBE opens up new opportunities for governance and research institutions that have previously struggled to utilize EO resources effectively.

Central to this mission is the development of the **FAIRiCUBE Hub**—a crosscutting platform for data ingestion, analysis, processing, and dissemination across European data spaces. Through five innovative Use Cases that address EU Green Deal priorities at urban and regional scales, FAIRiCUBE demonstrates practical applications of this infrastructure:

- **UC1** – Urban adaptation to climate change
- **UC2** – Biodiversity and agriculture nexus
- **UC3** – Environmental Adaptation Genomics in *Drosophila*
- **UC4** – Spatial and temporal assessment of neighbourhood building stock
- **UC5** – Biodiversity occurrence cubes

This report presents the **validation process** established for each Use Case, focusing on both thematic consistency and the evaluation of fitness for purpose. The process involves two key components: (1) the formal specification and implementation of each Use Case, and (2) user-centered assessments to determine relevance, usability, and practical value.

By aligning technical outputs with specific user needs and the overarching goals of the EU Green Deal, this process supports the issuance of structured **validation certificates**. These certificates document the stakeholder's engagement and evaluation outcomes, ensuring traceability, transparency, and credibility in the co-creation of FAIR and TRUSTable data-driven solutions across environmental, biodiversity, and climate domains.

More information on **FAIRiCUBE validation** can be found here: <https://hub.fairicube.eu/validation.html>

Validation certificate

- **Name:** Sonja Steindl
- **Date of response:** 21/05/2025 11:44:50
- **Level of validation:** UC3
- **Specific data set of process to validate:**
- **Stage of validation:** Initial stage

This validation certificate certifies the following items:

- General Use Case
- FAIRiCUBE Hub and Services
- Communication and Dissemination

1. General Use Case Validation

The **validation process** to be performed for each Use Case at a thematic level. Covers both the Use Case specification as well as wider User assessment & fit-for-purpose checks.

1.1. LEVEL 1

Table 1 highlights the key general checks on use case specifications and user assessment/fit-for-purposes.

Table 1: Validation Summary – General Checks

Check	Value
General Checks - Use Case Specifications	Clear Goal defined
General Checks - Use Case Specifications	Required datasets identified
General Checks - Use Case Specifications	Required ML/AI approaches identified
General Checks - Use Case Specifications	Workflow designed
General Checks - User assessment/fit-for-purpose	Support the users' work
General Checks - User assessment/fit-for-purpose	Applicability

1.1. LEVEL 2 & 3

1.1.1. UC Specifications - Clear Goal defined (level 2)

The detailed aspects validated by the UC on the clear definition of goals are addressed in table 2

Table 2: Clear Goal Definition – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Specific - The objective is clearly defined and addresses a precise problem.	X				
Realistic - The objective is practical and relevant.	X				
Time-bound - The objective has a clear deadline.		X			

Level 3 Detail - (Additional stakeholder comments): Some analysis that is done in the scope of the project will trigger further investigations that will not be completed within the project time.

1.1.2. UC Specifications - Required datasets identified (level 2)

Table 3 summarises the aspects related to the required datasets identified.

Table 3: Required Datasets Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Relevant datasets	X				
Resource estimation			X		

Level 3 Detail - (Additional stakeholder comments): Resource estimation is complicated due to local computation.

1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)

Table 4 summarises the general validation of ML/AI approaches identification.

Level 3 Detail - (Additional stakeholder comments):

1.1.4. UC Specifications - Workflow designed (level 2)

In relation to workflow designed, table 5 summarises its validation level.

Table 4: Required ML/AI Approaches Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Suitable approaches	X				
Processing resources estimation			X		

Table 5: Workflow Designed – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Workflow outline	X				
Documentation		X			

Level 3 Detail - (Additional stakeholder comments): Some details and links in the documentation might still change- also result section.

1.1.5. UC Specifications - Visualisation of outputs designed (level 2)

Table 6 lists the status of the visualisation of outputs designed.

Table 6: Visualisation of Outputs – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Visualisation of outputs designed		X			

Level 3 Detail - (Additional stakeholder comments): Grafana Dashboard needs to be public

1.1.6. User assessment/fit-for-purpose (level 2)

In relation to user assessment/fit-for purposes, table 7 provides level 2 validation summary for the UC.

Level 3 Detail - (Additional stakeholder comments):

Table 7: User Assessment – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Support the users' work - User assistance	X				
Service orientation - Transparent service chain	X				
Reliability - Method documentation	X				
Applicability - Fit-for-purpose		X			
Stable and interoperable systems		X			
Standardised feedback collection		X			

4. FAIRiCUBE Hub and services

Validation process to ensure that the components of **FAIRiCUBE Hub on information, data, processing and portrayal** each function individually, as well as jointly.

4.1. LEVEL 1

Table 8 lists the processes validated within the FAIRiCUBE Hub and services.

Table 8: General Checks – FAIRiCUBE Hub and Services

	Included
Hub	
Documentation	
Knowledge Base	X
Metadata Editors	X
Data Catalog	X
WCS	
S3	X
Sentinel Hub	X
GitHub	X
Lab	

4.2. LEVEL 2 & 3

4.2.1. Hub (level 2)

Related to FC Hub validation, table 9 lists the status of different aspects related the FC Hub.

Table 9: FAIRiCUBE Hub Validation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Findability of components		X			
Authentication across components	X				
Interplay - Accessibility of datasets		X			
Interplay - Accessibility of analysis/processing resources	X				
Interplay - Accessibility of linked datasets and a/p resources		X			
Interplay - Storage and accessibility of processing results			X		
Interplay - Visualization of processing results			X		
User support		X			

Level 3 Detail - (Additional stakeholder comments):

4.2.2. Documentation (level 2)

Table 10 illustrates the status of the different aspects in relation to documentation

Table 10: Documentation – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Findability of required information	X				
Correctness of available information	X				
Completeness of available information	X				
Usefulness of available information	X				
Ease of provision of documentation	X				

Level 3 Detail - (Additional stakeholder comments):

4.2.3. Knowledge Base (level 2)

Table 11 lists the status of the validation criteria for the knowledge base.

Table 11: Knowledge Base – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Findability of required information	X				
Correctness of available information	X				
Completeness of available information	X				
Usefulness of available information	X				
Ease of provision of information for the KB	X				

Level 3 Detail - (Additional stakeholder comments):

4.2.4. Metadata Editors (level 2)

FC should ensure that Metadata Editors are effective tools for sharing and validating metadata, thereby enhancing the overall quality and usability of shared data. The validation criteria for FC metadata is listed in table 12

Table 12: Metadata Editors – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Persistence	X				
Alignment		X			
Findability		X			
Labeling			X		
Errors		X			

Level 3 Detail - (Additional stakeholder comments):

4.2.5. Data Catalog (level 2)

The Data Catalog aims to be user-friendly and efficient. Table 13 lists the status of the validated criteria.

Level 3 Detail - (Additional stakeholder comments):

Table 13: Data Catalog – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Findability		X			
Simple Search		X			
Advanced Search		X			
Navigation	X				

4.2.6. WCS (level 2)

Data access via WCS is important, the status of the validation criteria are summarised in table 14 , below.

Table 14: WCS – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Correct Response					
Measure Types					
XML Validity					
Extent/Bbox					
Subset Accuracy					
Format					

Level 3 Detail - (Additional stakeholder comments):**4.2.7. S3 (level 2)**

Data access to AWS S3 bucket should be user-friendly, efficient and secure. The status of the validation criteria is summarised in table 15, below.

Table 15: S3 – FAIRiCUBE Hub (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Correct Response					
Measure Types					
XML Validity					
Extent/Bbox					
Subset Accuracy					
Format					

Level 3 Detail - (Additional stakeholder comments):

5. Communication and dissemination

For **communication and dissemination** , a number of KPI's are used for validation measures.

5.1. LEVEL 1

The validation checks included in the validation for communication and dissemination are summarised in table 16.

Table 16: General Checks – Communication and Dissemination

	Included
EU policy	X
Domain specific events	
Newsletter	
Website	X
Social media	X
Press releases and articles published in national/regional/European online media	X
Innovation workshops	
Trainings and master workshop	
Scientific papers	

5.2. LEVEL 2 & 3

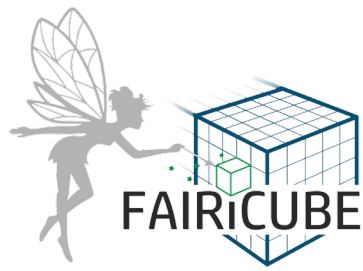
5.2.1. Communication and Dissemination (level 2)

The achievement status of various areas are listed in table 17.

Table 17: Communication and Dissemination – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
EU policy					
Domain specific events					
Newsletter					
Website					
Social media					
Press releases and articles published in national/regional/European online media					
Innovation workshops					
Trainings and master workshop					
Scientific papers					

Level 3 Detail - (Additional stakeholder comments):



FAIRiCUBE - Validation certificate

Validation of UC4 - Final Stage

Project Horizon Europe No 101059238

FAIRiCUBE - Validation certificate

Validation of UC4 - Final Stage

Jaume Targa, María Colina & Jaume Archilés

© 4sfera Innova
Carrer Cor de Maria 2, àtic 1a
17002 Girona
www.4sfera.com
contact@4sfera.com

Published on: 16/07/2025

Contents

Introduction	3
Validation certificate	4
1. General Use Case Validation	5
1.1. LEVEL 1	5
1.1. LEVEL 2 & 3	6
1.1.1. UC Specifications - Clear Goal defined (level 2)	6
1.1.2. UC Specifications - Required datasets identified (level 2)	6
1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)	6
1.1.4. UC Specifications - Workflow designed (level 2)	7
1.1.5. UC Specifications - Visualisation of outputs designed (level 2)	7
1.1.6. User assessment/fit-for-purpose (level 2)	7
3. Processing and ML applications	8
3.1. LEVEL 1	8
3.2. LEVEL 2 & 3	8
3.2.1. Algorithm implementation validation (level 2)	8
3.2.2. Benchmarking (level 2)	9
3.2.3. Comprehensive documentation (level 2)	9
3.2.4. Machine learning validation (level 2)	9
5. Communication and dissemination	10
5.1. LEVEL 1	10
5.2. LEVEL 2 & 3	11
5.2.1. Communication and Dissemination (level 2)	11
6. AI ethics assessment	12
6.1. LEVEL 1	12
6.2. LEVEL 2 & 3	13
6.2.1. Ethics (Trustworthy AI) (level 2)	13
6.2.2. GDPR applicability (level 2)	13

Introduction

The **FAIRiCUBE project** empowers stakeholders beyond traditional Earth Observation (EO) communities by enabling them to provide, access, process, and share gridded data and algorithms in a **FAIR** (Findable, Accessible, Interoperable, Reusable) and TRUSTable manner. By leveraging Machine Learning (ML) on multi-thematic data cubes, FAIRiCUBE opens up new opportunities for governance and research institutions that have previously struggled to utilize EO resources effectively.

Central to this mission is the development of the **FAIRiCUBE Hub**—a crosscutting platform for data ingestion, analysis, processing, and dissemination across European data spaces. Through five innovative Use Cases that address EU Green Deal priorities at urban and regional scales, FAIRiCUBE demonstrates practical applications of this infrastructure:

- **UC1** – Urban adaptation to climate change
- **UC2** – Biodiversity and agriculture nexus
- **UC3** – Environmental Adaptation Genomics in *Drosophila*
- **UC4** – Spatial and temporal assessment of neighbourhood building stock
- **UC5** – Biodiversity occurrence cubes

This report presents the **validation process** established for each Use Case, focusing on both thematic consistency and the evaluation of fitness for purpose. The process involves two key components: (1) the formal specification and implementation of each Use Case, and (2) user-centered assessments to determine relevance, usability, and practical value.

By aligning technical outputs with specific user needs and the overarching goals of the EU Green Deal, this process supports the issuance of structured **validation certificates**. These certificates document the stakeholder's engagement and evaluation outcomes, ensuring traceability, transparency, and credibility in the co-creation of FAIR and TRUSTable data-driven solutions across environmental, biodiversity, and climate domains.

More information on **FAIRiCUBE validation** can be found here: <https://hub.fairicube.eu/validation.html>

Validation certificate

- **Name:** Daniel Moran
- **Date of response:** 13/05/2025 13:18:59
- **Level of validation:** UC4
- **Specific data set of process to validate:**
- **Stage of validation:** Final stage

This validation certificate certifies the following items:

- General Use Case
- Processing and ML Applications
- Communication and Dissemination
- AI Ethics Assessment

1. General Use Case Validation

The **validation process** to be performed for each Use Case at a thematic level. Covers both the Use Case specification as well as wider User assessment & fit-for-purpose checks.

1.1. LEVEL 1

Table 1 highlights the key general checks on use case specifications and user assessment/fit-for-purposes.

Table 1: Validation Summary – General Checks

Check	Value
General Checks - Use Case Specifications	Clear Goal defined
General Checks - Use Case Specifications	Required datasets identified
General Checks - Use Case Specifications	Required ML/AI approaches identified
General Checks - Use Case Specifications	Workflow designed
General Checks - User assessment/fit-for-purpose	Service orientation

1.1. LEVEL 2 & 3

1.1.1. UC Specifications - Clear Goal defined (level 2)

The detailed aspects validated by the UC on the clear definition of goals are addressed in table 2

Table 2: Clear Goal Definition – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Specific - The objective is clearly defined and addresses a precise problem.	X				
Realistic - The objective is practical and relevant.	X				
Time-bound - The objective has a clear deadline.		X			

Level 3 Detail - (Additional stakeholder comments):

1.1.2. UC Specifications - Required datasets identified (level 2)

Table 3 summarises the aspects related to the required datasets identified.

Table 3: Required Datasets Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Relevant datasets		X			
Resource estimation		X			

Level 3 Detail - (Additional stakeholder comments): data selection felt opportunistic rather than structured

1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)

Table 4 summarises the general validation of ML/AI approaches identification.

Table 4: Required ML/AI Approaches Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Suitable approaches		X			
Processing resources estimation		X			

Level 3 Detail - (Additional stakeholder comments): again, ML/AI approaches felt opportunistic rather than structured.

1.1.4. UC Specifications - Workflow designed (level 2)

In relation to workflow designed, table 5 summarises its validation level.

Table 5: Workflow Designed – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Workflow outline	X				
Documentation		X			

Level 3 Detail - (Additional stakeholder comments):

1.1.5. UC Specifications - Visualisation of outputs designed (level 2)

Table 6 lists the status of the visualitaion of outputs designed.

Table 6: Visualisation of Outputs – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Visualisation of outputs designed			X		

Level 3 Detail - (Additional stakeholder comments):

1.1.6. User assessment/fit-for-purpose (level 2)

In relation to user assessment/fit-for purposes, table 7 provides level 2 validation summary for the UC.

Level 3 Detail - (Additional stakeholder comments):

Table 7: User Assessment – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Support the users' work - User assistance			X		
Service orientation - Transparent service chain			X		
Reliability - Method documentation		X			
Applicability - Fit-for-purpose	X				
Stable and interoperable systems	X				
Standardised feedback collection				X	

3. Processing and ML applications

Validation process on **processing and ML applications** to be performed for Use Case at an analysis and processing level. Covers data processing validation including algorithm implementation validation, comprehensiveness of documentation and benchmarking; machine learning validation; ethical and bias validation.

3.1. LEVEL 1

Table 8 highlights the key general checks on processing and ML applications.

Table 8: General Checks – Use Case Specifications

	Included
Algorithm implementation validation	X
Benchmarking	
Comprehensive documentation	
Machine learning validation	X

3.2. LEVEL 2 & 3

3.2.1. Algorithm implementation validation (level 2)

The detailed aspects validated on algorithm implementation are addressed in Table 9.

Level 3 Detail - (Additional stakeholder comments):

Table 9: Algorithm Implementation Validation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Technical Robustness and safety	X				
Assess the interactions		X			
End-to-end testing	X				
Cross-validation				X	

3.2.2. Benchmarking (level 2)

Table 10 lists the performance validation (benchmarking) aspects.

Table 10: Benchmarking – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Monitor compute resources	X				
Re-run and compare		X			

Level 3 Detail - (Additional stakeholder comments): NA

3.2.3. Comprehensive documentation (level 2)

Documentation and transparency aspects are included in table 11

Table 11: Comprehensive Documentation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Documentation and transparency		X			
Meta-data			X		

Level 3 Detail - (Additional stakeholder comments): NA

3.2.4. Machine learning validation (level 2)

Validation of Machine learning features are summarised in table 12.

Level 3 Detail - (Additional stakeholder comments): NA

Table 12: Machine Learning Validation – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Dataset preparation for training	X				
Define appropriate validation metrics	X				
Prevent/Test overfitting and underfitting	X				
Statistical bias validation	X				
Human agency and oversight	X				

5. Communication and dissemination

For **communication and dissemination**, a number of KPI's are used for validation measures.

5.1. LEVEL 1

The validation checks included in the validation for communication and dissemination are summarised in table 13.

Table 13: General Checks – Communication and Dissemination

	Included
EU policy	X
Domain specific events	
Newsletter	
Website	
Social media	
Press releases and articles published in national/regional/European online media	
Innovation workshops	
Trainings and master workshop	
Scientific papers	

5.2. LEVEL 2 & 3

5.2.1. Communication and Dissemination (level 2)

The achievement status of various areas are listed in table 14.

Table 14: Communication and Dissemination – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
EU policy					
Domain specific events					
Newsletter					
Website					
Social media					
Press releases and articles published in national/regional/European online media					
Innovation workshops					
Trainings and master workshop					
Scientific papers					

Level 3 Detail - (Additional stakeholder comments):

6. AI ethics assessment

Validation process to be performed for the Use Case to assure **relevant considerations pertaining to ethical use of AI are met**.

6.1. LEVEL 1

The validation process in relation to ethics (Trustworthy AI) and GDPR applicability includes a checklist which requires validation. Table 15 lists the AI ethics assessment checklist included in the validation.

Table 15: General Checks – AI Ethics Assessment

	Included
Fundamental rights	
Privacy and data protection	
Transparency rights	
Accessibility	
Education and tutorials	
Data management	
Security	
Ease to deactivate/remove	
Ease to access services without using the AI system	
Open-source code	
Ownership	
Openness about Data governance	
Legislation and Policy	
Design Impact Assessment and Open Development Process	
Right to contest/liability	
List of data	
Personal data	

6.2. LEVEL 2 & 3

6.2.1. Ethics (Trustworthy AI) (level 2)

In relation to Ethics (Trustworthy AI), the status of the different aspects validated are summarised in table 16.

Table 16: Ethics – Trustworthy AI Assessment (Level 2)

	Done	Partially Done	In Progress	Not Done	Could not be Done
Fundamental rights	X				
Privacy and data protection	X				
Transparency rights	X				
Accessibility	X				
Education and tutorials					X
Data management				X	
Security	X				
Ease to deactivate/remove					X
Ease to access services without using the AI system					X
Open-source code					X
Ownership	X				
Openness about Data governance	X				
Legislation and Policy	X				
Design Impact Assessment and Open Development Process	X				
Right to contest/liability	X				

Level 3 Detail - (Additional stakeholder comments):

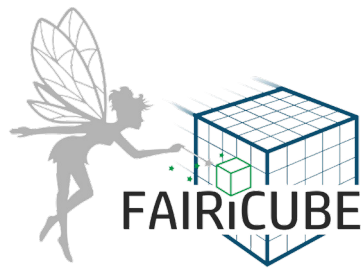
6.2.2. GDPR applicability (level 2)

In relation to GDPR applicability, the status of the different aspects validated are summarised in table 17.

Table 17: GDPR Applicability – Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
List of data				X	
Personal data					X

Level 3 Detail - (Additional stakeholder comments):



FAIRiCUBE - Validation certificate

Validation of UC5 - Final stage

Project Horizon Europe No 101059238

FAIRiCUBE - Validation certificate

Validation of UC5 - Final stage

Jaume Archilés, Eduard Lama, Jaume Targa,
María Colina & Lorena Banyuls

© 4sfera Innova
Carrer Cor de Maria 2, àtic 1a
17002 Girona
www.4sfera.com
contact@4sfera.com

Published on: 22/07/2025

Contents

Introduction	3
Validation certificate	4
1. General Use Case Validation	5
1.1. LEVEL 1	5
1.1. LEVEL 2 & 3	6
1.1.1. UC Specifications - Clear Goal defined (level 2)	6
1.1.2. UC Specifications - Required datasets identified (level 2)	6
1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)	7
1.1.4. UC Specifications - Workflow designed (level 2)	7
1.1.5. UC Specifications - Visualisation of outputs designed (level 2)	7
1.1.6. User assessment/fit-for-purpose (level 2)	8

Introduction

The **FAIRiCUBE project** empowers stakeholders beyond traditional Earth Observation (EO) communities by enabling them to provide, access, process, and share gridded data and algorithms in a **FAIR** (Findable, Accessible, Interoperable, Reusable) and TRUSTable manner. By leveraging Machine Learning (ML) on multi-thematic data cubes, FAIRiCUBE opens up new opportunities for governance and research institutions that have previously struggled to utilize EO resources effectively.

Central to this mission is the development of the **FAIRiCUBE Hub**—a crosscutting platform for data ingestion, analysis, processing, and dissemination across European data spaces. Through five innovative Use Cases that address EU Green Deal priorities at urban and regional scales, FAIRiCUBE demonstrates practical applications of this infrastructure:

- **UC1** – Urban adaptation to climate change
- **UC2** – Biodiversity and agriculture nexus
- **UC3** – Environmental Adaptation Genomics in *Drosophila*
- **UC4** – Spatial and temporal assessment of neighbourhood building stock
- **UC5** – Biodiversity occurrence cubes

This report presents the **validation process** established for each Use Case, focusing on both thematic consistency and the evaluation of fitness for purpose. The process involves two key components: (1) the formal specification and implementation of each Use Case, and (2) user-centered assessments to determine relevance, usability, and practical value.

By aligning technical outputs with specific user needs and the overarching goals of the EU Green Deal, this process supports the issuance of structured **validation certificates**. These certificates document the stakeholder's engagement and evaluation outcomes, ensuring traceability, transparency, and credibility in the co-creation of FAIR and TRUSTable data-driven solutions across environmental, biodiversity, and climate domains.

More information on **FAIRiCUBE validation** can be found here: <https://hub.fairicube.eu/validation.html>

Validation certificate

- **Name:** Susanna Ioni
- **Date of response:** 7/22/2025 12:37:52
- **Level of validation:** UC5
- **Specific data set of process to validate:**
- **Stage of validation:** Final stage

This validation certificate certifies the following items:

- General Use Case

1. General Use Case Validation

The **validation process** to be performed for each Use Case at a thematic level. Covers both the Use Case specification as well as wider User assessment & fit-for-purpose checks.

1.1. LEVEL 1

Table 1 highlights the key general checks on use case specifications and user assessment/fit-for-purposes.

Table 1: Validation Summary – General Checks

Check	Value
General Checks - Use Case Specifications	Clear Goal defined
General Checks - Use Case Specifications	Required datasets identified
General Checks - Use Case Specifications	Required ML/AI approaches identified
General Checks - Use Case Specifications	Workflow designed
General Checks - Use Case Specifications	Visualisation of outputs designed
General Checks - User assessment/fit-for-purpose	Support the users' work
General Checks - User assessment/fit-for-purpose	Service orientation
General Checks - User assessment/fit-for-purpose	Reliability
General Checks - User assessment/fit-for-purpose	Applicability
General Checks - User assessment/fit-for-purpose	Data systems stability
General Checks - User assessment/fit-for-purpose	reliability
General Checks - User assessment/fit-for-purpose	
General Checks - User assessment/fit-for-purpose	interoperability

1.1. LEVEL 2 & 3

1.1.1. UC Specifications - Clear Goal defined (level 2)

The detailed aspects validated by the UC on the clear definition of goals are addressed in table 2

Table 2: Clear Goal Definition – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Specific - The objective is clearly defined and addresses a precise problem.	X				
Realistic - The objective is practical and relevant.	X				
Time-bound - The objective has a clear deadline.	X				

Level 3 Detail - (Additional stakeholder comments): Use case 5 aims to refine and validate habitat prediction methods by incorporating species occurrence data from scientific, museum collections and citizen science sources, with environmental variables, through machine learning models. By integrating plant species occurrences from the Global Biodiversity Information Facility (GBIF) with climatic and topographic data from Earth Observation sources, UC5 seeks to improve habitat mapping accuracy within Europe and highlight the usefulness of data sources as GBIF. Our approach, built over the EUNIS habitat study case 'S22', focuses on comparing predicted species distributions with existing habitat prediction maps, particularly from the European Nature Information System (EUNIS), and exploring the effectiveness of data cubes in improving habitat classification accuracy. The ultimate goal is to improve our understanding of habitat dynamics and promote a more data-driven approach to habitat prediction, with applications in biodiversity monitoring, conservation planning, and museum collections management.

1.1.2. UC Specifications - Required datasets identified (level 2)

Table 3 summarises the aspects related to the required datasets identified.

Table 3: Required Datasets Identified – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Relevant datasets	X				
Resource estimation	X				

Level 3 Detail - (Additional stakeholder comments): Datasets relevant for the UC5 were retrieved (GBIF, environmental datasets, pseudoabsences datasets created) and resources were estimated during the process.

1.1.3. UC Specifications - Required ML/AI approaches identified (level 2)

Table 4 summarises the general validation of ML/AI approaches identification.

Table 4: Required ML/AI Approaches Identified - Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Suitable approaches	X				
Processing resources estimation	X				

Level 3 Detail - (Additional stakeholder comments): The suitable approaches were identified as ensemble model built on individual models and TSS scores, and processing resources were estimated during the process.

1.1.4. UC Specifications - Workflow designed (level 2)

In relation to workflow designed, table 5 summarises its validation level.

Table 5: Workflow Designed - Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Workflow outline	X				
Documentation	X				

Level 3 Detail - (Additional stakeholder comments): Documentation of the workflow and methods can be found in FAIRiCUBE website, GitHub repository of the UC5, and in the documentation published in Read the Docs.

1.1.5. UC Specifications - Visualisation of outputs designed (level 2)

Table 6 lists the status of the visualisation of outputs designed.

Table 6: Visualisation of Outputs – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Visualisation of outputs designed	X				

Level 3 Detail - (Additional stakeholder comments): Outputs of the UC5 were visualized as response plots, species distribution maps and visual comparison of EUNIS probability map against the UC5 prediction map of the Habitat S22.

1.1.6. User assessment/fit-for-purpose (level 2)

In relation to user assessment/fit-for purposes, table 7 provides level 2 validation summary for the UC.

Table 7: User Assessment – Validation Level 2

	Done	Partially Done	In Progress	Not Done	Could not be Done
Support the users' work - User assistance	X				
Service orientation - Transparent service chain	X				
Reliability - Method documentation	X				
Applicability - Fit-for-purpose	X				
Stable and interoperable systems		X			
Standardised feedback collection		X			

Level 3 Detail - (Additional stakeholder comments): The method is documented in different platforms, and a complete user-friendly script of UC5 method is available in GitHub