Project title: European Federation for Cancer Images

Project acronym: EUCAIM

Grant Agreement: 101100633

Call identifier: DIGITAL-2022-CLOUD-AI-02

D2.5: The FAIR Report: Data Maturity Model Adoption on the EUCAIM data federation

Responsible Partner: HULAFE

Authors: Silvia Flor (HULAFE), David Rodríguez (CSIC-IFCA), Inés Victoria Rodríguez (CSIC-IFCA), Santiago Frid (HCB), Ignacio Blanquer (UPV), Valia Kalokyri (FORTH), Mahlet Birhanu (ERASMUS), Esther Bron (ERASMUS), Irene Marín (HULAFE), Mirna El-Ghosh (LIMICS)

Contributors: Marco Aiello (SYNLAB), Kurt Majcen (BBMRI-ERIC)

Reviewers: Krystian Brzozowski (GUMED), Laura Portell-Silva (BSC)

Date of delivery: 23/12/2025

Version: 1

Table of contents

List of Abbreviations and Acronyms

- **API:** Application Programming Interface
- **CC-BY:** Creative Commons Attribution; Licence models for open content
- **CDM:** Common Data Model
- **CLI:** Command Line Interface
- **CSV:** Comma-Separated Values
- **DAML+OIL:** DARPA Agent Markup Language + Ontology Inference Layer
- **DCAT-AP:** Data Catalogue Vocabulary Application Profile
- **DH:** Data Holder
- **DICOM**: Digital Imaging and Communications in Medicine
- **DOI:** Digital Object Identifier
- **EOSC:** European Open Science Cloud
- **FAIR:** Findable, Accessible, Interoperable, Reusable
- **FAIR-EVA:** FAIR Evaluation and Validation Application
- **FDP:** FAIR Data Point
- **HTTP:** Hypertext Transfer Protocol
- **IANA:** Internet Assigned Numbers Authority
- **IRI**: International Resource Identifier
- **JSON:** JavaScript Object Notation
- **KPI:** Key Performance Indicators
- **NIfTI:** Neuroimaging Informatics Technology Initiative
- **OAI-PMH:** Open Archives Initiative Protocol for Metadata Harvesting
- **ODC:** Open Data Commons; Licence models for open content
- **ORCID:** Open Researcher and Contributor ID
- **OWL:** Web Ontology Language
- **PURL:** Persistent Uniform Resource Locator
- **RDA:** Research Data Alliance
- **RDF:** Resource Description Framework
- **RWD:** Real World Data
- **SHACL:** Shapes Constraint Language
- **TXT:** Text file format
- **UI:** User Interface
- **URI:** Uniform Resource Identifier
- **URL:** Uniform Resource Locator
- **UUID:** Universally Unique Identifier
- **WP**: Work Package
- **XSD:** XML Schema Definition

1. Introduction

1.1. Background and Context

The FAIR principles (Findable, Accessible, Interoperable and Reusable) serve as the foundation for the management and stewardship of scientific data in the digital era. They aim to maximize the value and reusability of data by ensuring that both humans and machines can efficiently discover, access, integrate, and reuse information.

These principles are essential to foster reproducibility, transparency and interoperability across diverse research infrastructures and datasets. In the context of EUCAIM, the adoption of FAIR principles is pivotal to achieving a unified, sustainable, and federated infrastructure for cancer imaging data in Europe.

In alignment with international guidelines and the FAIR principles defined by Wilkinson et al. (2016), the following dimensions guide the EUCAIM FAIR adoption framework:

- Findable (F): Data and metadata must be uniquely and persistently identifiable, described with rich and machine-readable metadata, and registered in searchable catalogues.
- Accessible (A): Data should be retrievable via standardized, open, and interoperable communication protocols, with clear access conditions and persistent metadata availability even if the data are withdrawn.
- Interoperable (I): Data and metadata should utilize standardized formats, shared vocabularies, and formal languages (e.g., RDF, JSON-LD, OWL) that facilitate integration and reuse across systems and domains.
- Reusable (R): Data must be richly described with accurate attributes, clear usage licenses, provenance information, and adherence to community standards, ensuring long-term reproducibility and value.

In the next section (1.2. FAIR Principles Overview), we will go over each of these principles with more detail, so that their scope can be clearly understood.

Task 2.4: FAIR Implementation Support, is designed to support Data Holders (DHs) in adopting FAIR principles through practical tools and guidance.

The main objectives and actions imply:

- Establishing a FAIR Implementation Support Team, composed of data scientists, engineers, and clinicians, to provide coordinated assistance to Data Holders during the FAIRification process (the process of making data FAIR).
- Developing a FAIR compliance checklist, based on the RDA FAIR Data Maturity Model and Data Object Assessment Metrics, to evaluate:
    - Metadata and ontology descriptions,
    - Data location and accessibility,
    - Access levels and authentication mechanisms,
    - Management of persistent and unique identifiers.
- Performing quantitative analyses to measure the adoption of FAIR principles across the EUCAIM federation, using these outcomes as Key Performance Indicators (KPIs).
- Providing structured feedback and corrective actions to Data Holders to improve metadata quality and dataset alignment with the EUCAIM Common Data Model (CDM): https://eucaim.gitbook.io/eucaim-common-data-model/1.-introduction
- Leveraging experience from EOSC-Synergy and other FAIR implementation initiatives, combining automated and manual evaluation approaches to ensure accuracy, scalability and reproducibility.

To operationalize the evaluation of FAIR compliance within the EUCAIM federation, Task 2.4 uses FAIR-EVA for EUCAIM, a customised version of the FAIR-EVA tool (FAIR-Evaluation and Validation Application), developed in EUCAIM's WP5 (see section 2.5).

FAIR-EVA is an automated assessment framework that evaluates dataset metadata against the RDA FAIR Data Maturity indicators (see Annex 1), providing both quantitative and qualitative FAIRness scores and qualitative feedback.

In EUCAIM, FAIR-EVA has been configured to automatically process the datasets listed in the EUCAIM Public Catalogue, generating detailed reports in both individual (.txt) and aggregated (.csv) formats.

These results help identify strengths, gaps and opportunities for improvement, allowing the FAIR Implementation Support team to guide Data Holders toward higher FAIR maturity levels for metadata and data.

In EUCAIM, FAIR maturity is not only assessed through indicator scores, but also operationalised through the tier-level system (see Figure 1) as mentioned in D4.4 Final Rules for Participation, section 2.2. Tier upgrades in the context of the EUCAIM EDIC.

The main dimension of value of the EUCAIM infrastructure is the provision of large, homogenised, standardised, and de-identified cancer imaging datasets together with curated, highly structured clinical data for secondary use. While it is recognised that datasets made accessible through EUCAIM will initially be predominantly Tier 1-compliant, the achievement of Tier 3 remains the ultimate objective for all data collections. Progression through the tiers inherently drives higher FAIRness (see Figure 1), as compliance with advanced Tier requirements directly increases adherence to the FAIR principles. This approach creates a structured incentive for Data Holders participating in European or publicly funded initiatives, where FAIR alignment is often a mandatory requirement. Tier improvement is therefore both a FAIR compliance mechanism and a sustainability strategy, particularly in view of the future EUCAIM EDIC, in which ongoing funding and collaboration will be sought to enable quality enhancement of EUCAIM collections.

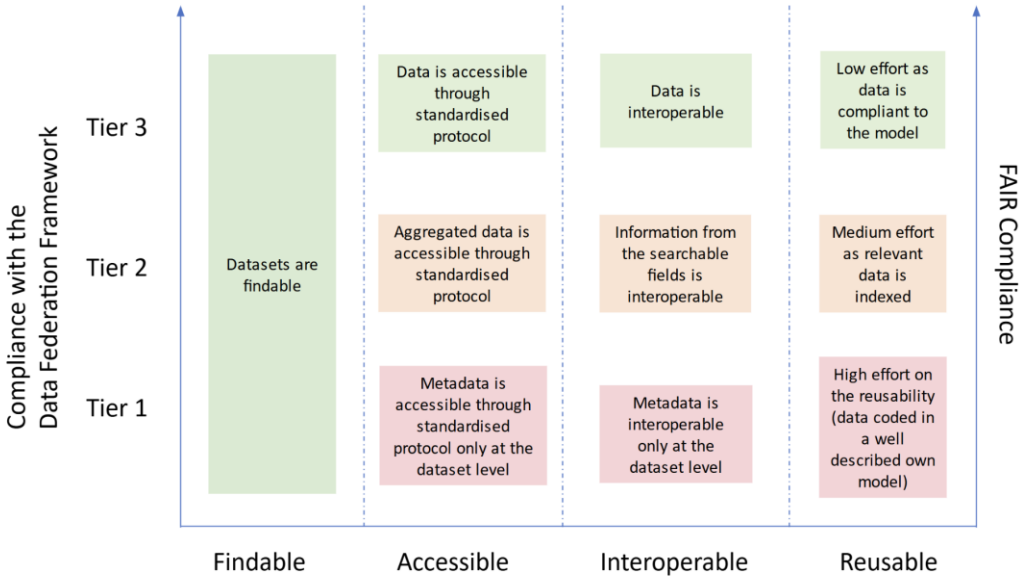


Figure 1. Tiers and FAIR compliance.

It is worth noting that this task has required substantial inter-work-package (WP) collaboration. Several synergies with other WPs are essential to implement the catalogue updates, which are coordinated through WP4 to ensure that metadata changes are consistently reflected in the EUCAIM Public Catalogue

and RDF endpoints. Regarding metadata alignment, the work follows WP5 guidelines, using the common standards and hyperontology definitions to harmonize dataset descriptions and enhance machine interpretability.

1.2. FAIR Principles Overview

This section presents the four components of the FAIR principles based on the Go-Fair initiative: https://www.go-fair.org/fair-principles/.

F: Findable

The initial phase of data utilization involves locating the necessary datasets. It's crucial that both metadata and data are readily discoverable by humans as well as by computer systems. The presence of metadata that can be read and interpreted by machines plays a key role in the automated identification of datasets and related services, marking a critical step in making data FAIR-compliant. There are four indicators included in this principle:

- **F1. (Meta)data are assigned a globally unique and persistent identifier**. This principle emphasizes the use of globally unique and persistent identifiers to eliminate ambiguity in data publication, facilitating precise data interpretation and integration. For example, an individual researcher can be uniquely identified by an ORCID link such as "https://orcid.org/0000-1111-2222-333X".
- **F2. Data are described with rich metadata (defined by R1 below)**. It mandates enriching data with extensive metadata, detailing everything from the data's creation context to its specific attributes, to enhance discoverability and usability without reliance on identifiers. For instance, metadata for an image file might include the DICOM information captured automatically by the imaging device, along with contextual details like the measurement protocol and device links, ensuring the data can be found and utilized for a variety of unforeseen research purposes.
- **F3. Metadata clearly and explicitly include the identifier of the data they describe**. F3 underlines the necessity for metadata to explicitly reference the unique and persistent identifier of the dataset they describe, ensuring a clear linkage between the dataset and its descriptive metadata. An example of this is using the "foaf:primaryTopic" predicate in RDF metadata to formally establish the connection between metadata and its dataset.
- **F4. (Meta)data are registered or indexed in a searchable resource**. (Meta)data must be made discoverable by registering or indexing them in searchable resources, ensuring datasets are easily found by researchers and machines alike. An example of this practice includes the indexing of FAIR Datasets published on FAIR Data Points, which are then searchable through specialized engines like the DTL Search Engine.

A: Accessible

After locating the desired data, users must understand the access process, which may encompass authentication and authorization requirements. This implies that:

- **A1. (Meta)data are retrievable by their identifier using a standardised communications protocol.** The protocols used must be open, free, and universally implementable, like HTTP, to maximize data reuse. Additionally, these protocols should support authentication and authorization procedures when necessary, ensuring secure access to data while maintaining the flexibility to accommodate protected and private data within the FAIR framework.
- **A2: Metadata should be accessible even when the data is no longer available**. This ensures that information about the data's context, creation, and usage can persist. Although data online has

costs, and data may be removed over time, the enduring presence of metadata allows researchers to understand and potentially replicate studies or trace the origins and applications of the data, preserving the value of the original research efforts.

I: Interoperable

Interoperability ensures that data can be seamlessly combined with other datasets and effectively work with various applications or workflows for analysis, storage, and processing, facilitating comprehensive research and operational efficiencies. There are three principles regarding interoperability:

- **I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation**. This implies that both humans and machines can interpret the language easily, without requiring specialized algorithms or translations. This language should have a well-defined syntax and grammar for precise communication, and be openly accessible. Examples include RDF, OWL, DAML+OIL, and JSON LD.
- **I2. (Meta)data use vocabularies that follow FAIR principles**. This means that the terms or concepts used must be well-documented and identifiable through unique, persistent identifiers, and that any user of the dataset can understand and utilize the vocabulary effectively.
- **I3. (Meta)data include qualified references to other (meta)data**. The aim is to weave a rich tapestry of interconnected data resources that enhance the understanding of the data's context and relationships, thereby augmenting the dataset's value. This approach involves clearly indicating when one dataset relies upon or complements another, ensuring that these scientific connections are well articulated and that all referenced datasets are properly cited with their unique identifiers.

R: Reusable

Reusability, the cornerstone of the FAIR principles, aims to maximize data's potential for reuse. This is achieved by ensuring that both data and metadata are thoroughly documented, allowing for their replication and integration across various contexts.

- **R1: (Meta)data are richly described with a plurality of accurate and relevant attributes**. Data usage licenses must be clearly specified to ensure legal interoperability and facilitate reuse. Furthermore, detailed provenance must be provided to establish the data's origin and processing history, enhancing trust and clarity on how to cite or acknowledge the data. Finally, (meta)data should adhere to domain-relevant community standards and best practices, ensuring consistency and ease of integration with existing datasets, while allowing for deviations when justified.

2. Methodology

2.1 FAIR-EVA for EUCAIM

FAIR-EVA: Evaluator, Validator & Advisor (https://github.com/IFCA-Advanced-Computing/FAIR_eva), originally developed within the EOSC Synergy project, is a tool designed to assess the FAIRness level of digital objects from different repositories or data portals. It requires the object identifier (preferably persistent and unique identifier) and the repository to check. FAIR-EVA offers a generic, repository-agnostic approach to checking digital objects. FAIRness scores depend on completeness and correctness of metadata.

In EUCAIM's task 5.3.5 a dedicated plugin, adapted to the project needs, has been developed. This plugin (**FAIR-EVA for EUCAIM,** https:/bio.tools/fair_eva_4_eucaim) implements a completely customised version of FAIR-EVA, in which the validation logic has been specifically adapted to the RDF structure used in the EUCAIM catalogue, including the namespaces, vocabularies and metadata fields actually present. Since different RDF models require different validation approaches, the tests are tailored to the way FAIR requirements are expressed in EUCAIM metadata, while avoiding as much as possible changes in the main code to avoid future compatibility issues. As the project advanced and EUCAIM's metadata was updated, the plugin has been further developed and updated to the requirements, while keeping track with the original FAIR-EVA updates.

2.1.1. Goals

The goals of this service are:

- Evaluate the FAIRness of the EUCAIM datasets in the EUCAIM Public Catalogue based on the RDA indicators.
- Provide feedback to the DHs to improve the FAIRness of their datasets.

2.1.2. Integration with EUCAIM

FAIR-EVA for EUCAIM implements a dedicated plugin for FAIR-EVA, and comes with a tailored configuration to work with EUCAIM's catalogue. On top of this, some minor changes have been made to the main program (FAIR-EVA) in order to cater for some of EUCAIM's particular requirements. These are being reported to FAIR-EVA main developers so there is no divergence in the future. Furthermore, along with the development of FAIR-EVA for EUCAIM, we have kept track of the development of the main package to avoid future maintainability problems, and we have provided feedback for future versions.

This plugin implements methods for the evaluation of the RDA indicators (see Annex 1) customised for EUCAIM use. Following FAIR-EVA implementation, it provides feedback for each indicator, but on top of this, it also offers a simplified evaluation stating the top EUCAIM tier the dataset reaches (see deliverable D5.6). The current version supports the currently available EUCAIM Public Catalogue, but also has a preview support for the planned changes in EUCAIM Public Catalogue schema.

Part of the evaluation is performed on the RDF graphs using SHACL shapes, both with existing SHACLs for DCAT-AP and HealthDCAT-AP, and with custom ones developed inside FAIR-EVA for EUCAIM to check EUCAIM's own parts of the schema. The latter are available for both versions of the catalogue mentioned above.

2.1.3. Components

FAIR-EVA is a web service that can be deployed as a stand-alone application or in a docker container. But the latter is preferred. On top of this, different clients are provided. The list of independently deployable components available for EUCAIM use of FAIR-EVA follows:

- **Web Service**: Core FAIR-EVA evaluator, modified, adapted and configured for EUCAIM use.
- **Web Client**: User-friendly interface that supports queries for individual Datasets.
- **Command Line Client**: Supports single dataset and batch evaluation. The latter produces output files, detailed below, and checks only new datasets, or the ones with metadata modified in the EUCAIM Public Catalogue.

For deploying these components, three Docker images are available:

- Web service: ***fair_eva_service***
- Web client: ***fair_eva_web***
- Command line client: ***fair_eva_cli***

These images have been optimised to reduce as much as possible their sizes. Both the Web Client and the Command Line client rely on the Web Service, which must be running before using any client.

The command line client admits both a query for a single dataset with the dataset ID as input, and a batch mode that checks all the datasets in the catalogue.

2.1.4. Getting started: Use of FAIR-EVA

Docker Web Service deployment

To deploy FAIR-EVA for EUCAIM Web Service in a Docker container, just run the latest image:

```
docker run --name=FAIR_EVA -dit --network host
harbor.eucaim.cancerimage.eu/processing-tools/fair_eva_service:latest
```

This will run as a service as long as it is not stopped, and will expose the Web Service in the default port 9090. To change this, use the usual Docker mechanism, e.g. “-p HOST_PORT:CONTAINER_PORT”.

2.1.4.1. Docker Web Client

For using the web interface, first run the following command:

```
docker run -dit -p 5000:5000 harbor.eucaim.cancerimage.eu/processing-
tools/fair_eva_web:latest
```

Open the browser (see Figure 2):

http://[ip_address]:5000/en

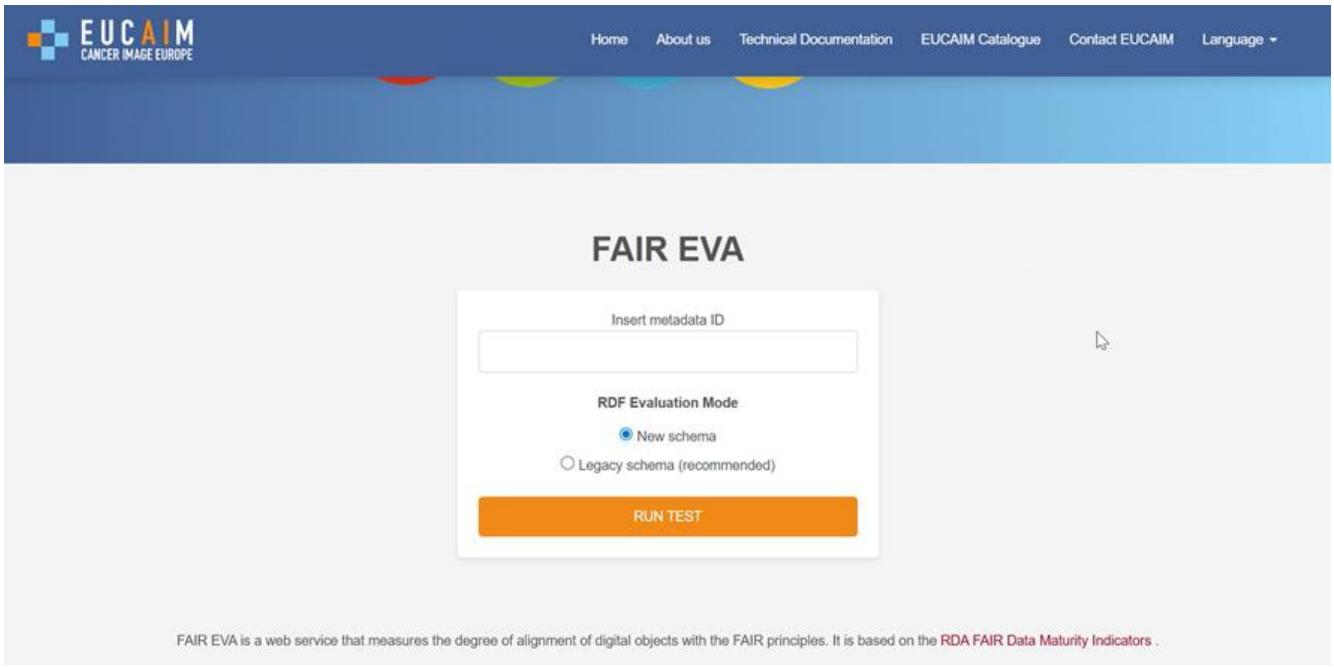


Figure 2. FAIR-EVA's main page using the web interface.

Choose a *dataset id* from the EUCAIM Public Catalogue (see Figure 3)
https://catalogue.eucaim.cancerimage.eu/Eucaim/eucaim-ui/#/catalogue

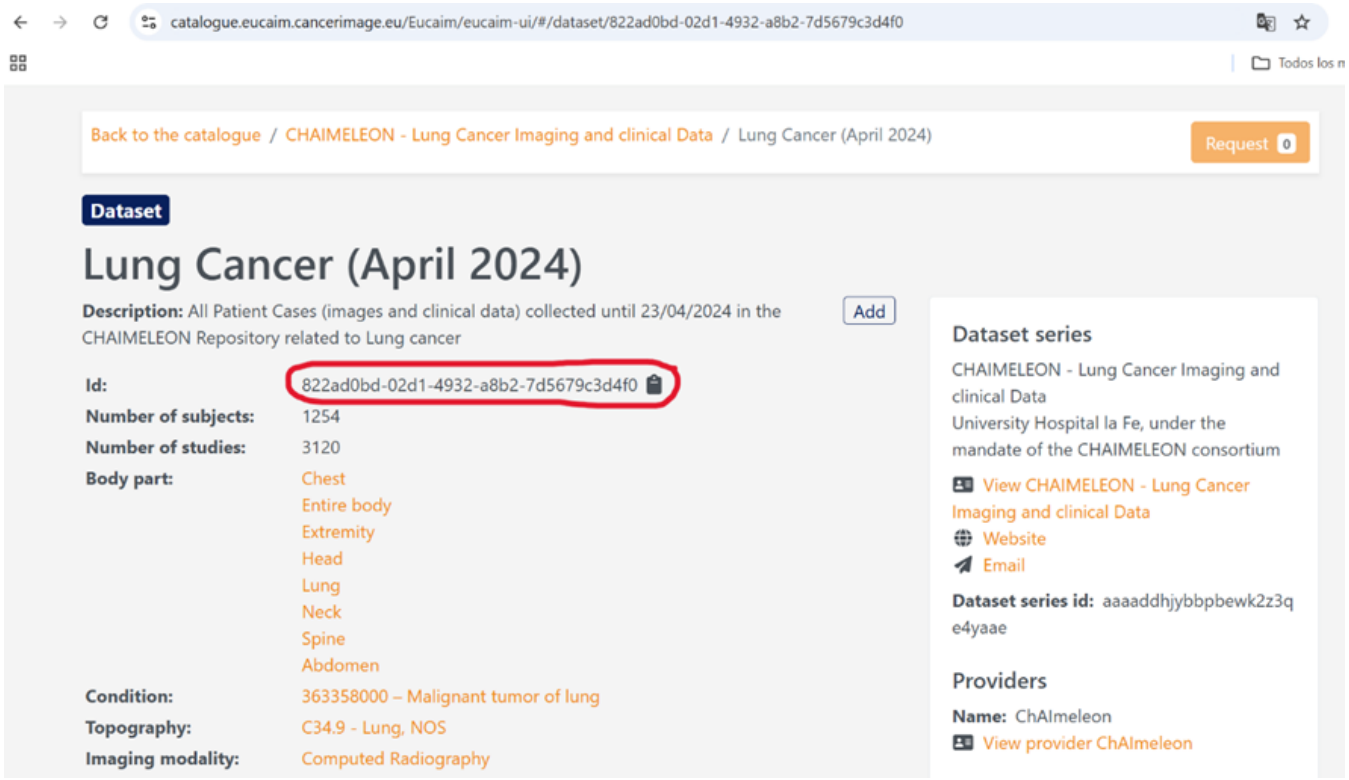


Figure 3. Where to find the dataset id in the catalogue.

And choose the RDF Evaluation mode:

- **Legacy Schema** — this matches the current structure of the RDF exposed in the EUCAIM Public Catalogue.

- **New Schema** — this follows the new EUCAIM's catalogue vocabulary (EUCAIM-DCAT-AP (max) v1.0.xlsx), but it is not yet used in the catalogue.

Wait for the results and if you want, download the results pressing on the "Export PDF" button (see Figure 4).

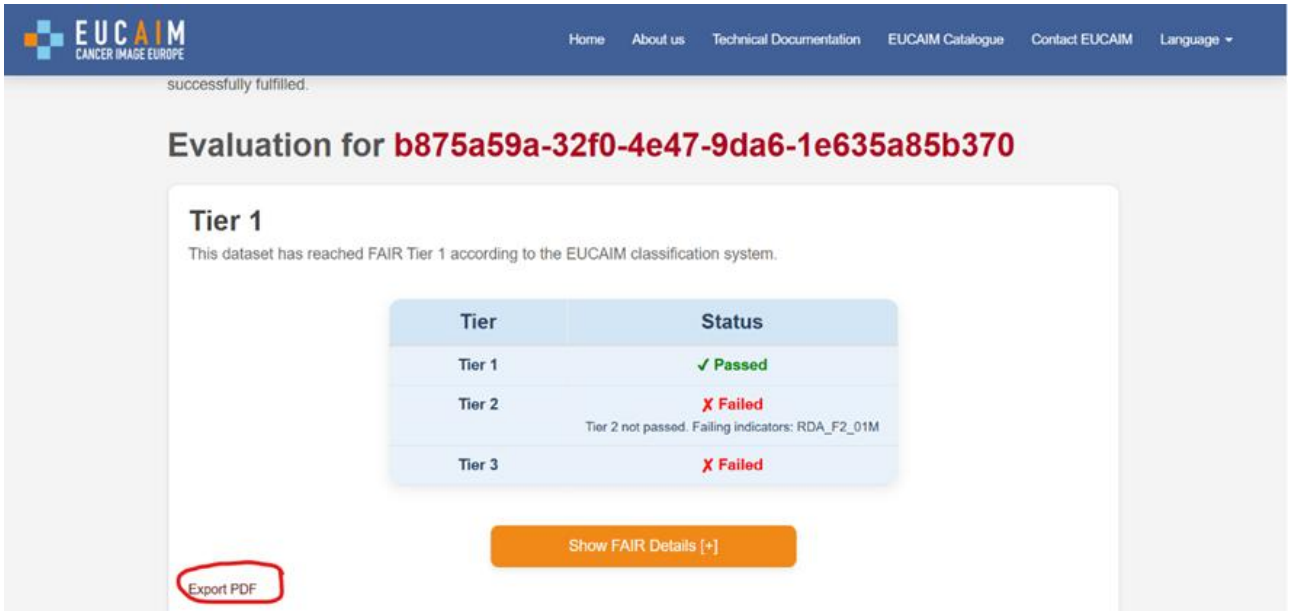


Figure 4. Location of the "Export PDF" button in the results page.

2.1.4.2. Docker Command Line Client

First we need to pull the Command Line Client image:

```
docker pull harbor.eucaim.cancerimage.eu/processing-tools/fair_eva_cli:latest
```

1. Batch Mode — Evaluate ALL datasets

Runs *batch_runner.py* inside the container.

If using Windows

```
docker run --rm -t --network host -v
C:/path_to_project_faireva_out:/FAIR_eva/faireva_out
harbor.eucaim.cancerimage.eu/processing-tools/fair_eva_cli:latest  -b --eucaim_old
```

If using Linux

```
docker run --rm  -t --network host -v OUTDIR:/FAIR_eva/faireva_out
harbor.eucaim.cancerimage.eu/processing-tools/fair_eva_cli:latest -b  --eucaim_old
```

Additional Available Flags

`--api_root API_ROOT` EUCAIM Public Catalogue RDF root URL

`--out OUT`	Output directory inside container (default: faireva_out)
`--eucaim_old`	Use old EUCAIM SHACL shapes
`--host HOST`	Hostname where FAIR-EVA API is running
`--limit LIMIT`	Process only first N datasets

(Most users only need `-b --eucaim_old`.)

Text files will be created for each dataset in the output folder, additionally a csv file for all the datasets will also be created.

2. Single Dataset Mode — Evaluate ONE dataset

Runs *fair-eva.py* inside the container.

Command Example

```
docker run --rm --network host harbor.eucaim.cancerimage.eu/processing-tools/fair_eva_cli:latest -i 822ad0bd-02d1-4932-a8b2-7d5679c3d4f0 -p eucaim_catalogue -j --eucaim_old
```

Available Flags

`-d`	Debug mode
`-i IDENTIFIER`	Dataset ID
`-p PLUGIN`	FAIR-EVA plugin
`-r URL`	Repository endpoint
`--api-endpoint URL`	API endpoint for evaluation
`-j`	Print JSON output
`--eucaim_old`	Use EUCAIM-OLD SHACL shapes

2.1.4.3. Optional: using docker-compose

Using Docker compose you can easily deploy both the web service and the web frontend at the same time, exposing both ports or only the one corresponding to the frontend. Example *docker-compose.yaml* file:

```
name: FAIR_EVA
services:
  fair_eva:
    image: harbor.eucaim.cancerimage.eu/processing-tools/fair_eva_service
    container_name: web_service
    ports:
      - "9090:9090"

  web:
```

```
image: harbor.eucaim.cancerimage.eu/processing-tools/fair_eva_web
container_name: web_client
command: ["--api-endpoint", "http://web_service:9090/v1.0/rda/rda_all"]
ports:
 - "5000:5000"
depends_on:
 -    fair_eva
```

To start both the evaluation service and the web interface:

```
docker compose up -d
```

2.1.4.4. Configuration

The *config.ini* file contains all the configuration parameters. They are distributed in different sections. To customize your FAIR evaluator deployment. In principle, there is no need for the users to adjust any of the values, as they have already been customised for EUCAIM use.

2.1.5. Limitations

- **Scope**: Currently limited to EUCAIM Public Catalogue datasets.
  - If Data Holders provide their own FDPs, any divergence with EUCAIM Public Catalogue schema will limit the validity of the results.
- **Performance**: Batch evaluation may require significant resources for large catalogues.
  - Previous results from batch execution need to be kept to avoid unnecessary checks.
- **Future Changes**: Adaptation needed for major schema updates or FAIR-EVA upstream changes.

2.2 Checklist design

The FAIR compliance checklist provides Data Holders with a practical instrument to self-evaluate the FAIR maturity of their (meta)data and associated infrastructure. The checklist items correspond directly to the FAIR-EVA indicators implemented for EUCAIM; the detailed description of each indicator and the metadata fields evaluated are provided in Annex 1.

1. Findable (F)

1.1 Persistent Identifiers

- A **persistent and globally unique identifier** (UUID / DOI / Handle) is assigned to the metadata. *(RDA_F1_01M, RDA_F1_02M)*
- A **persistent and globally unique identifier** (UUID / DOI / Handle) is assigned to the dataset. *(RDA_F1_01D, RDA_F1_02D)*
- The dataset identifier is explicitly included in the metadata *(RDA_F3_01M)*

N.B. A Metadata identifier, distinct from the Dataset identifier, is assigned to each dataset upon its inclusion in the EUCAIM catalogue.

1.2 Metadata richness

- Metadata include the minimum discovery fields:
  - title
  - description
  - identifier.
- The EUCAIM findability elements are present:
  - hasCondition
  - topography (only present in old version)
  - hasImageBodyPart
  - hasEquipmentManufacturer/hasImageVendor
  - hasImageModality
  - collectionMethod/collection_method
  - numberofUniqueIndividuals/number_of_subjects
  - hasBirthSex/hasAssociatedSex
  - interoperabilityLevel
  - image_access_type
  - provider
  - spatial/geographical_coverage

(RDA_F2_01M)

1.3 Indexing and harvesting

- Metadata are accessible via HTTP(S) as RDF.
- Metadata can be harvested through OAI-PMH or API REST endpoints.

(RDA_F4_01M) (Handled by EUCAIM's catalogue)

2. Accessible (A)

2.1 Access documentation

- Metadata clearly describe how the data can be accessed (image_access_type, image_access_description, image_access_uri).
  (RDA_A1_01M)

2.2 Manual access

- The data access link resolves correctly to a valid resource or landing page. *(RDA_A1_02D)*

2.3 Automatic access

- Data can be accessed automatically through a resolvable access URL (publication_uri) that supports machine-driven retrieval. *The access URI must return a valid, reachable endpoint without requiring manual intervention.*
  (RDA_A1_05D)

2.4 Identifier resolution

- The DOI or publication URI resolves correctly to the digital object or its landing page. Data identifier resolves correctly to the digital object.
  (RDA_A1_03D)

2.5 Protocols

- Data are retrievable through a standardized protocol (HTTP/HTTPS/FTP).
  (RDA_A1_04D)
- The protocol is open and free to use.
  (RDA_A1.1_01M, RDA_A1.1_01D)

2.6 Authentication

- Authentication and authorization mechanisms are documented in the metadata.
  (RDA_A1_2_01D)

3. Interoperable (I)

3.1 Standardised metadata formats

- Metadata comply with DCAT-AP.
- Metadata comply with HealthDCAT-AP.
- Metadata comply with the current EUCAIM profile.
  (RDA_I1_01M)

3.2 Machine-readable metadata

- Metadata are expressed as a valid, non-empty RDF graph.
  (RDA_I1_02M)

3.3 Standardised data formats

- Data are provided in IANA-registered formats such as DICOM, DICOM-SEG, or NIfTI.
  (RDA_I1_01D, RDA_I1_02D)

3.4 FAIR vocabularies

- Values for controlled fields use resolvable URIs rather than literals.
  - hasCondition
  - hasImageBodyPart
  - hasEquipmentManufacturer/hasImageVendor
  - hasBirthSex/hasAssociatedSex
  - type
  - theme

    (RDA_I2_01D)

3.5 Qualified references

- Metadata include qualified references to:
  - publisher
  - contactPoint
  - spatial / temporal_coverage
  - provenance
  - relation
  - hasQualityAnnotation

    (RDA_I3_01M, RDA_I3_03M)

4. Reusable (R)

4.1 Reuse-supporting metadata

- Metadata include fields relevant for reuse
    - spatial/geographical_coverage
    - theme
    - type
    - sample
    - version
    - collectionMethod/collection_method
    - hasBirthSex/hasAssociatedSex
    - hasCondition
    - hasEquipmentManufacturer/hasImageVendor
    - hasImageModality
    - hasImageBodyPart
    - interoperabilityLevel
    - image_year_range
    - hasQualityAnnotation
    - minTypicalAge
    - maxTypicalAge
    - numberofRecords
    - numberofUniqueIndividuals/number_of_subjects
    - number_of_studies
    - hasPurpose

    (RDA_R1_01M)

4.2 Licensing

- A clear reuse license is provided in rights.
  (RDA_R1.1_01M)
- The license corresponds to a standard framework (e.g., CC-BY, ODC).
  (RDA_R1.1_02M)
- The license is provided as a resolvable URI (machine-readable).
  (RDA_R1.1_03M)

4.3 Provenance

- Provenance information is provided in provenance.
  (RDA_R1.2_01M)
- Provenance is expressed using a URI or typed literal (machine-readable).
  (RDA_R1.2_02M)

4.4 Community standards

- Metadata comply with HealthDCAT-AP in a machine-readable form.
  (RDA_R1.3_01M, RDA_R1.3_02M)
- Data comply with community standards (DICOM / NIfTI).
  (RDA_R1.3_01D, RDA_R1.3_02D)

2.3 Dataset cohorts

The FAIR-EVA assessment in EUCAIM was performed using all datasets currently available in the EUCAIM Public Catalogue: https://catalogue.eucaim.cancerimage.eu/Eucaim/eucaim-ui/#/catalogue (See Figure 5).

82 datasets, from 13 different Data Holders, represent cohorts collected across different EU projects (e.g., CHAIMELEON, PROCANCER-I, INCISIVE, EUCANIMAGE) as well as institutional contributions from hospital and research infrastructures (e.g., APHP, SAS, HULAFE, UOA, etc). (See Table 1).

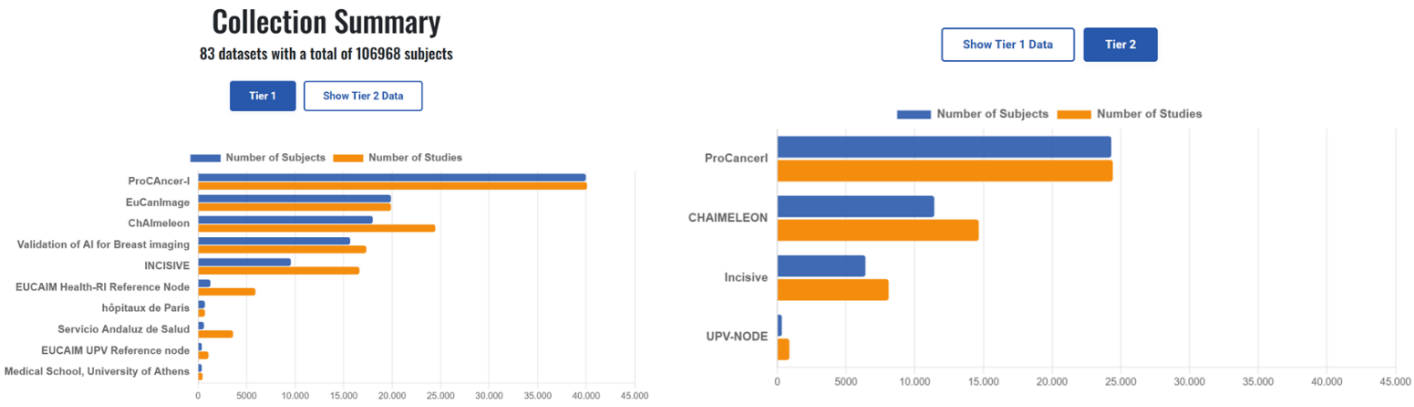


Figure 5. Collection Summary graphs shown in the EUCAIM Public Catalogue.

Table 1. Data Holders, number of datasets and cancer type in the EUCAIM Public Catalogue.

Type of provider	Data Holder	Number of datasets	Cancer type
AI4HI PROJECTS	CHAIMELEON	29	Lung Cancer, Breast Cancer, Colon Cancer, Prostate Cases, Rectum Cancer
	PROCANCER-I	11	Prostate Cancer
	INCISIVE	4	Breast, Colorectal, Lung, and Prostate Cancer
	EUCANIMAGE	23	Breast Cancer, Colon and Rectum Cancer, Liver Cancer

Transferring to UPV node	HULAFE	3	Neuroblastoma, DIPG & Glioblastoma multiforme
Transferring to Health-RI node	ERASMUS	1	Glioma
	AUMC	1	Glioma
Sharing as local nodes	APHP	2	Prostate Cancer, Hepatocellular carcinoma
	KI	1	Breast Cancer
	SCIENSANO	1	Various Cancer Types
	AUTH	2	Neoplasm of breast, Metastatic malignant neoplasm to thyroid gland
	UOA	3	Brain, Prostate and Soft Tissue Cancer
	SAS	1	Lung Cancer

* The FAIR_EVA tool had also analysed the entry “Build an observational study with RWD”, which appears under the provider “EUCAIM Observational Studies”. However, this is not an actual provider, nor is this an actual dataset. Instead, it is a placeholder entry designed to allow users to request the creation of new datasets for observational studies using real-world data (RWD). Therefore, this entry is excluded from the statistics.

3. Results

3.1. Overview of Dataset Evaluation

The 82 datasets were evaluated using the FAIR-EVA assessment tool, based on the Research Data Alliance (RDA) indicator framework (see Annex 1). Each dataset was assessed against 41 indicators covering the four FAIR principles: Findable, Accessible, Interoperable, and Reusable. The evaluation

employed a dual approach, separately assessing metadata (M) and data (D) components for each indicator, resulting in 3,362 individual assessments across the dataset collection. This comprehensive evaluation provides both quantitative scores and qualitative insights into FAIR compliance at both the metadata and data levels.

The FAIR-EVA tool that is automatically executed across the entire EUCAIM Public Catalogue has been configured to generate, for each dataset, a .txt file name [id]_dateandtime and a combined .csv file for all datasets (see Annex 2 for an example of table, extracted from the combined csv, for the HULAFE's PerProGlio's dataset), both include:

- Scores for each RDA indicator.
- Short explanatory notes on potential issues detected, to allow corrections if needed.
- The Tier assigned to the dataset and the missing RDA indicators required to reach the next Tier.

Also, the web interface for the FAIR-EVA tool allows to obtain pdf documents for each dataset as a visual FAIR report for Data Holders distribution (see Annex 3 for an example of HULAFE's PerProGlio's FAIR report).

3.2. EUCAIM Tier Classification

The tier classification analysis reveals that 48 datasets (58.5%) achieve the minimum FAIR standards required for Tier 1, with 4 datasets (4.9%) reaching the Tier 2 classification, and none achieving Tier 3. Furthermore, 30 datasets (36.6%) do not meet Tier 1 requirements, indicating significant room for improvement in foundational FAIR compliance (see Figure 6).

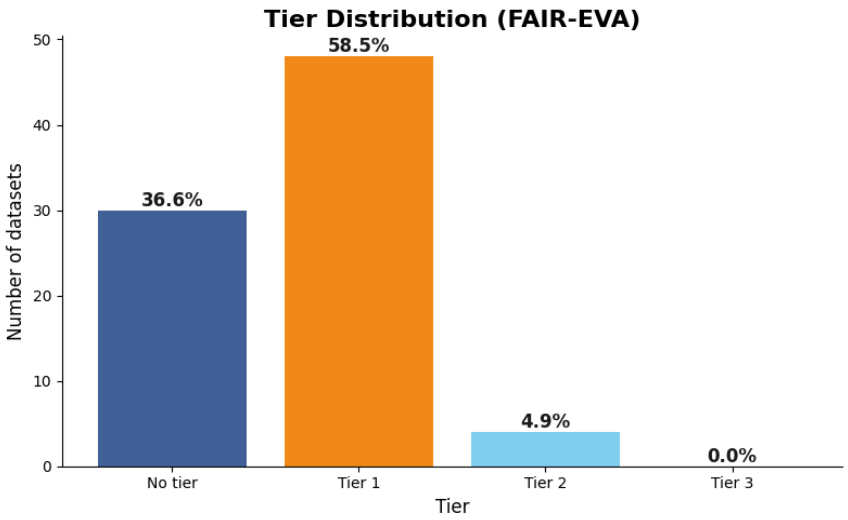


Figure 6. Tier distribution.

The 30 datasets failing to achieve Tier 1 exhibit a consistent pattern of deficiencies. All 30 datasets fail the persistent identifier requirements (RDA_F1_01M and RDA_F1_01D), using local or non-standardized identifiers rather than globally unique, persistent identifiers. Additional universal failures among this group include absence of DOI assignment for data (RDA_A1_03D), lack of authentication or authorization protocols (RDA_A1_2_01D), and deficiencies across the complete suite of data interoperability indicators. The root cause traces to the use of local identifiers (e.g., project codes, internal naming schemes) that do not conform to persistent identifier standards such as UUIDs or DOIs. Addressing this single issue—migration to persistent, globally unique identifiers—would immediately elevate all 30 datasets to Tier 1 status. As shown in *Table 2*, in the "Main failure points" column, these datasets cannot reach Tier 1 because the required indicators (RDA_F1_01M, RDA_F1_01D and RDA_F4_01M) are not met.

To aid interpretation of Table 2, it is important to note that the rightmost column lists the main failure points preventing datasets from reaching the next Tier level. Therefore, Tier 2 displays the requirements that are not yet met for achieving Tier 3 status (e.g., RDA_A1_03D and RDA_A1_04D). Since Tier 3 is the highest possible classification, there is no higher tier to ascend to; consequently, there are no associated failure points for Tier 3.

Table 2: EUCAIM Tier Distribution and Requirements. Failure points refer to unmet requirements for ascending to the next Tier; thereforeTier 3 has none.

Tier Level	Datasets (n)	Percentage	Core requirements met	Main failure points
No tier	30	36.6%	None—fails basic requirements	Non-persistent identifiers (RDA_F1, 100%), Missing DOI (RDA_A1_03D, 100%), No authentication (RDA_A1_2_01D, 100%)
Tier 1	48	58.5%	Persistent unique identifier and metadata can be gathered programmatically (RDA_F1_01M, RDA_F1_01D, RDA_F4_01M)	RDA_F2_01m missing findability terms present in current catalogue
Tier 2	4	4.9%	Globally unique identifier, findability terms present, metadata includes identifier for the data, metadata record can be retrieved from metadata id and standarised protocol to access metadata (RDA_F1_02M, RDA_F1_02D, RDA_F2_01M, RDA_F3_01M, RDA_A1_03M, RDA_A1_04M)	RDA_A1_03D, RDA_A1_04D
Tier 3	0	0%	DOI and standardised protocol to access data (RDA_A1_03D, RDA_A1_04D)	-

3.3. FAIR Principles Assessment

Overall FAIR compliance varies substantially across the four principles, ranging from 78.0% for Findable to 56.5% for Interoperability. A systematic metadata-data gap emerges across most principles, with metadata consistently outperforming data-level indicators.

Findable (F) - 78.0%

The Findable principle demonstrates relatively strong performance, with metadata scoring 83.8% and data 63.4%. Critical infrastructure for findability is universally implemented: all datasets achieve 100% compliance for metadata harvesting capability (RDA_F4_01M) via HTTP/API REST protocols and for inclusion of data identifiers within metadata (RDA_F3_01M).

The primary weakness lies in persistent identifier implementation, where 36.6% of datasets fail to use persistent, globally unique identifiers—directly corresponding to the No Tier population. A secondary gap appears in metadata richness (RDA_F2_01M), necessary to achieve Tier 2, where the average score is 92%. This indicator is only considered fulfilled when all findability elements used in the catalogue are present, only 34.1% of the datasets meet this requirement. The fields usually missing are interoperabilityLevel (36), topography (32), hasImageVendor (13), hasImageBodyPart (6), and hasCondition (5), (see Figure 7).

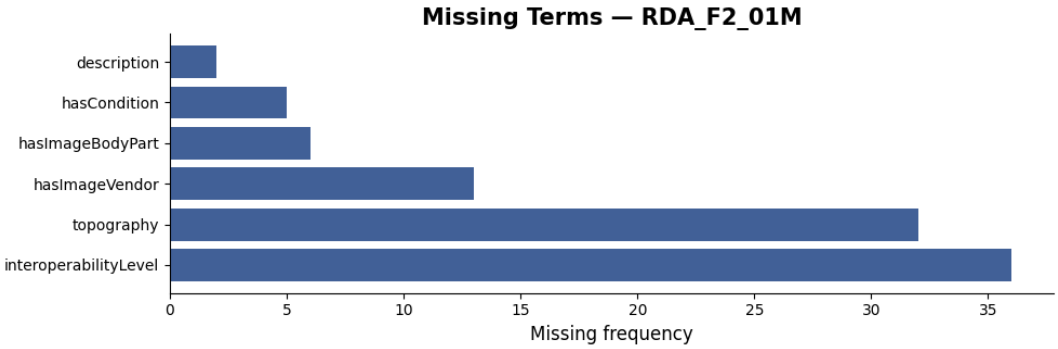


Figure 7. Most missing metadata terms in the RDA_F2_01M.

Accessible (A) - 69.0%

Accessibility reveals a stark metadata-data divide, with metadata achieving 90.1% versus data at 48.0%. Metadata accessibility benefits from universal implementation of key infrastructure: documentation for manual access (RDA_A1_02M, 100%), HTTPS standardized protocols (RDA_A1_04M, 100%), free/open protocols (RDA_A1_1_01M, 100%), and formalized preservation policies (RDA_A2_01M, 100%). Data accessibility faces significant challenges: only 14 datasets (19.5%) possess a resolvable DOI assignment (A1_03D), approximately 30% of data access URLs return resolution errors (404s) (RDA_A1_05D), and critically, zero datasets report authentication or authorization protocols for accessing the data in their metadata (RDA_A1_2_01D, 0%), representing a systemic gap in access control infrastructure.

Interoperable (I) - 84.7%

Interoperability emerges as the strongest FAIR dimension, particularly for data (98.9% versus metadata at 76.2%). Metadata demonstrates solid technical foundations: all datasets provide valid RDF representations (RDA_I1_02M, 100%) with 73-124 triples per dataset, and all referenced namespaces successfully resolve (RDA_I2_01M, 100%). The largest decrease in the total score comes from the group of indicators associated with qualified references (I3), where there is a consistent issue: publisher and provenance fields are currently expressed as literals, when they should instead be URIs or typed literals (RDA_I3_01M and RDA_I3_03M). In addition, indicators (RDA_I3_01D/02D) were excluded from the scoring because they require inspecting references within the data. Indicators (RDA_I3_02M/04M) were also excluded, as they assess references from metadata to other data.

Reusable (R) - 58.0%

The Reusable principle exhibits that data significantly outperforms metadata (98.8% versus 47.8%). This reversal stems from near-universal adoption of community-standard data formats, with approximately 100% of datasets using DICOM, DICOM-SEG, or NIfTI formats (RDA_R1_3_01D/02D). Metadata faces two universal failures affecting reusability: not a single dataset employs community-standard licensing frameworks (RDA_R1_1_02M, 0%), instead relying on internal policies or lacking explicit licenses entirely, and provenance information, while present in most datasets, is provided as plain text rather than machine-readable formats using standards such as W3C PROV-O (RDA_R1_2_02M, 0%). The absence of standardized, machine-interpretable licensing and provenance severely limits automated reuse workflows despite strong data format compliance.

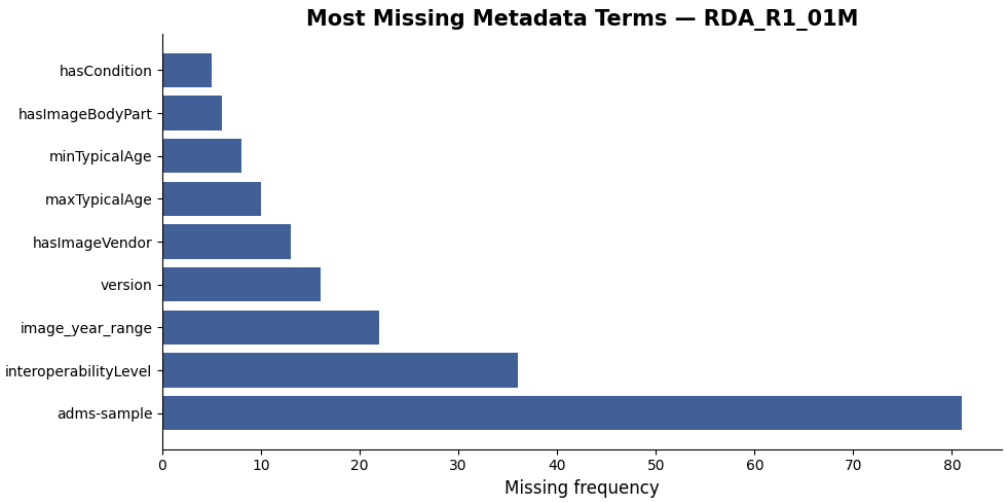


Figure 8. Most missing metadata terms in the RDA_R1_01M.

Table 3: FAIR Principles Summary Scores.

Principle	Average overall score	Metadata (M)	Data (D)	M-D gap	Representative strength (100%)	Critical gap (0%)
Findable	78.0%	83.8%	63.4%	+20.4%	Harvesting capability (F4)	-
Accessible	69.0%	90.1%	48.0%	+42.1%	Preservation policy (A2)	A1_2_01D
Interoperable	84.7%	76.2%	98.9%	-22.7%	Namespace resolution (I2)	EUCAIM schema properties (I1_01M)
Reusable	58.0%	47.8%	98.8%	-51.0%	Standard data formats (R1_3_01D)	Community-standard licenses (R1_1_02M)

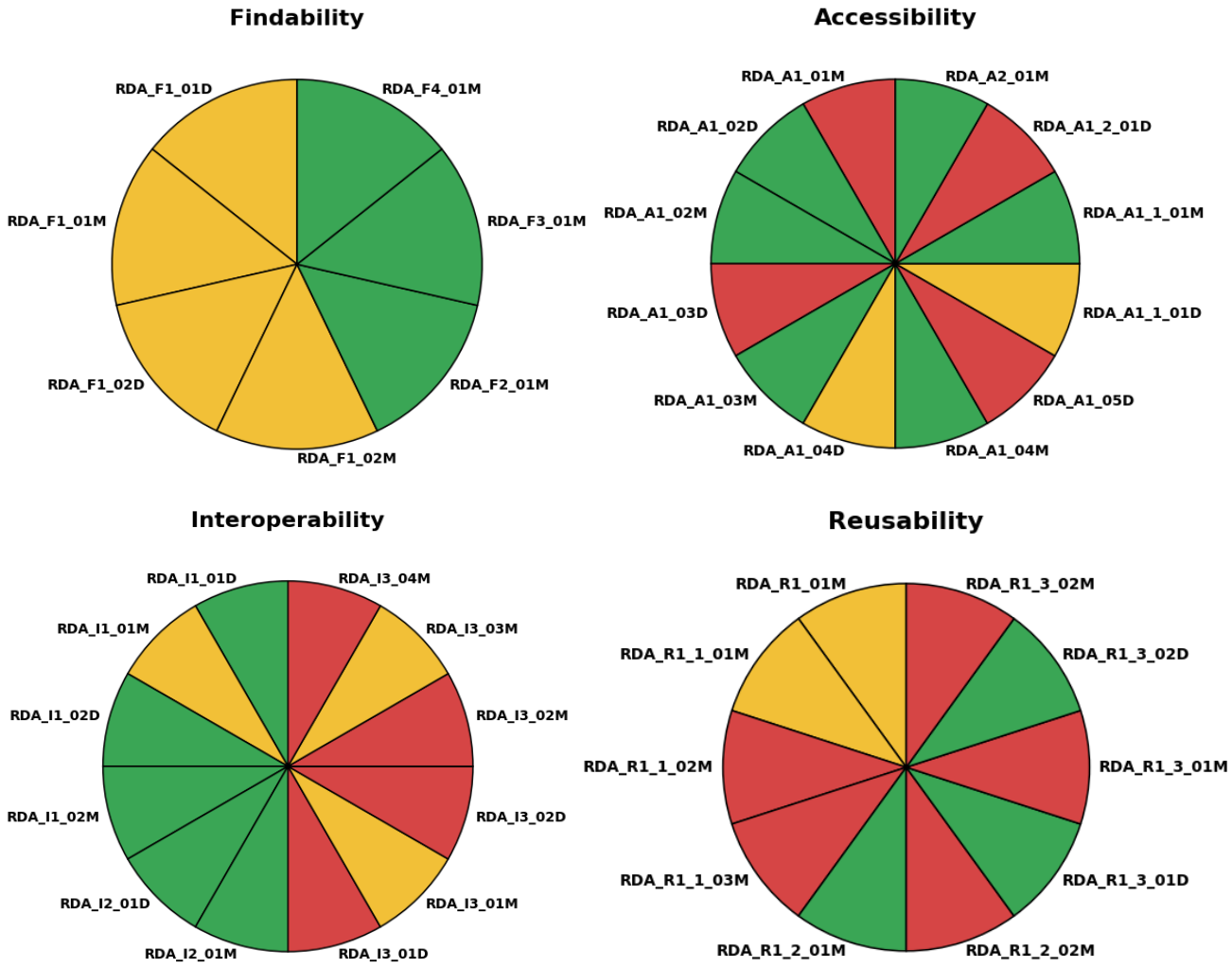


Figure 9. Pie charts showing the distribution of RDA indicators within each FAIR category (Findability, Accessibility, Interoperability and Reusability), coloured according to their mean score: green for ≥90, yellow for 50–90 and red for <50.

3.4. Cross-Cutting Findings

Analysis across all indicators reveals seven indicators with 0% compliance across the entire dataset collection, representing systemic implementation gaps: authentication/authorization for data access (RDA_A1_2_01D), the complete suite of qualified reference indicators (RDA_I3_01D, I3_02D, I3_02M, I3_04M), community-standard licensing adoption (RDA_R1_1_02M), and machine-readable provenance (RDA_R1_2_02M). These failures cluster thematically around access control mechanisms, and adoption of community standards for licensing and provenance.

Conversely, eleven indicators achieve 100% compliance, demonstrating robust foundational infrastructure: metadata documentation and retrieval (RDA_A1_02M, A1_03M), standardized metadata protocols (RDA_A1_04M, A1_1_01M), preservation policies (RDA_A2_01M), metadata identifier inclusion (RDA_F3_01M), harvesting capability (RDA_F4_01M), valid RDF metadata (RDA_I1_02M), namespace resolution (RDA_I2_01M), and data access documentation (RDA_A1_02D). This pattern indicates that while basic metadata infrastructure and documentation are well-established, advanced features requiring semantic standardization, access control, and community standard adoption remain unimplemented. The infrastructure exists for metadata management and discovery, but integration with broader FAIR

ecosystem standards (licensing frameworks, vocabulary services, authentication systems) has not been realized.

3.5. FAIR report analysis example: HULAFE's PerProGlio dataset

In order to thoroughly understand the reasons for the scoring obtained by the FAIR-EVA tool, we proceed to particularly analyze one of the FAIR reports obtained: HULAFE's PerProGlio dataset.

It is important to highlight that the FAIR-EVA tool has been run twice on the catalogue in order to address those low scores that could be easily improved in the short term. This analysis corresponds to the results of the first execution, during which these necessary improvements were identified. After applying them, the tool was executed again (with the updated results described throughout the remaining sections of this deliverable, and presented in Annex 2 and Annex 3). Therefore, only this section reflects an earlier execution of the tool and not the execution used for the final presentation of results.

It is also important to clarify that while the details of all indicators can be useful for improvement, it is not necessary to achieve 100% or meet a threshold in all of them.

The results obtained from the first execution of the tool can be classified into two groups: actionable indicators and non-actionable indicators, as shown in Table 4. Actionable indicators refer to checks whose scoring can be directly used to provide feedback to DH and non-actionable indicators, instead, correspond to tests whose scoring must not be interpreted as definitive, either because they require manual verification or because of their nature.

After inspecting the catalogue FAIR Data Point (https://catalogue.eucaim.cancerimage.eu/Eucaim/api/ttl), we found that many indicators with low scores do not depend on the metadata quality provided by the Data Holder directly, but rather on the way the metadata have been exposed. In other words, many of these low scores are not due to missing values, but instead reflect misalignments between the RDF format and standard vocabularies expected by the FAIR assessment, and the outdated version of the current FAIR Data Point implementation of the metadata catalogue.

Table 4. Actionable indicators.

Indicator	Actionable	Implementation problem
rda_f1_01m	No	We need to distinguish between metadata and data id
rda_f1_01d	No	We need to distinguish between metadata and data id
rda_f1_02m	No	We need to distinguish between metadata and data id
rda_f1_02d	No	We need to distinguish between metadata and data id
rda_f2_01m	Yes	No
rda_f3_01m	No	We need to distinguish between metadata and data id
rda_f4_01m	Yes	No
rda_a1_01m	Yes	No
rda_a1_02m	No	We need to distinguish between metadata and data id
rda_a1_02d	No	We need to distinguish between metadata and data id
rda_a1_03m	No	We need to distinguish between metadata and data id
rda_a1_03d	No	We need to distinguish between metadata and data id
rda_a1_04m	No	We need to distinguish between metadata and data id
rda_a1_04d	No	We need to distinguish between metadata and data id
rda_a1_05d	No	We need to distinguish between metadata and data id
rda_a1_1_01m	No	We need to distinguish between metadata and data id
rda_a1_1_01d	No	We need to distinguish between metadata and data id
rda_a1_2_01d	No	Metadata needs to include authentication/authorization protocol. Image_access_type?
rda_a2_01m	Yes	No
rda_i1_01m	No	Two versions available at the moment. One with how the RDFs are structured right now, another one with how it is specified in https://docs.google.com/spreadsheets/d/1_-iD5JDQuUF-OeX3_2MpQxVgWmi_n2GZ/edit?usp=sharing&ouid=117854846179287583875&rtpof=true&sd=true
rda_i1_01d	Yes	We do not check if data is really the format the metadata claims (dicom, dicom-seg, nifti)
rda_i1_02m	Yes	No
rda_i1_02d	Yes	No
rda_i2_01m	Yes	No
rda_i2_01d	Yes	No
rda_i3_01m	Yes	Publisher, provenance, contact point, spatial, relation, sample and quality annotation must be resolvable links or typed literals not literals. I don't think there are typed literals for provenance and publisher, so those two would need to be resolvable links
rda_i3_01d	Yes	No
rda_i3_02m	No	We need to distinguish between metadata and data id. It's about references to other (different) data in the metadata
rda_i3_02d	No	We need to distinguish between metadata and data id. This test implies checking the presence of qualified references within the content of the data.
rda_i3_03m	Yes	No
rda_i3_04m	No	No. It's about how metadata is connected to other data.
rda_r1_01m	No	Also dependent on the rdf structure
rda_r1_1_01m	Yes	No
rda_r1_1_02m	Yes	No
rda_r1_1_03m	Yes	No
rda_r1_2_01m	Yes	No
rda_r1_2_02m	Yes	No
rda_r1_3_01m	Yes	No
rda_r1_3_01d	Yes	No
rda_r1_3_02m	Yes	No
rda_r1_3_02d	Yes	No

A summary of the key findings from the actionable indicators that did not reach 100% for this dataset is presented below:

- **RDA-F2-01M - There is sufficiently rich metadata to enable resource discovery (57%)**:

  As described in Table 5, this indicator related to principle F2 is not detecting 6 of the 14 terms (Level Findability terms missing: 6/14). However, these terms do exist in the FAIR Data Point. Therefore, the issue arises from mismatches in the RDF properties rather than from missing values. As it can be seen in Annex 3, this issue has been corrected for the second execution, obtaining a score of 100%.

  Table 5. RDA-F2-01M.

Indicator Level	**Essential**
Indicator Assessment	This indicator is linked to the following principle F2: Data are described with rich metadata The indicator concerns not only the presence of metadata but also the amount provided how such metadata supports resource discovery.
Technical Implementation	Compares the metadata elements present with a configured list considered essential for discoverability. Calculates the proportion of elements found relative to those expected to assign a score.
Technical feedback	Findability terms present (8/14): • http://purl.org/dc/terms/title • http://purl.org/dc/terms/description • http://purl.org/dc/terms/identifier • http://purl.org/dc/terms/spatial • http://www.healthdcatap.org/numberofUniqueIndividuals http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_type• http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/provider • http://www.w3.org/ns/adms#interoperabilityLevel Findability terms missing (6/14): • https://www.eucaim.org/hasCondition • https://www.eucaim.org/hasImageBodyPart • https://www.eucaim.org/collectionMethod • https://www.eucaim.org/hasBirthSex • https://www.eucaim.org/hasImageModality • https://www.eucaim.org/hasEquipmentManufacturer
Tips	Add all fields related to the findability of the dataset in the catalogue: title, description, identifier, hasCondition, topography, ImageBodyPart, geographical coverage, collection method, number of subjects, birth sex, image modality, vendor/manufacturer, image access type, provider and interoperability level.

Findability terms missing (6/14):

- ***https://www.eucaim.org/hasCondition***

```
<https:/www.eucaim.org/hasCondition>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Disease_types?name=SNOMED-CT%3A393563007>;
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/condition>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Disease_types?name=SNOMED-CT%3A393563007>;
```

- *https://www.eucaim.org/hasImageBodyPart*

```
<https:/www.eucaim.org/hasImageBodyPart>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Body_parts?name=SNOMED-CT%3A69536005>;
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/body_part_examined>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Body_parts?name=SNOMED-CT%3A69536005>;
```

- *https://www.eucaim.org/collectionMethod*

```
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/collection_method>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collection_types?name=DISEASE_SPECIFIC>,
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collection_types?name=COHORT>;
```

- *https://www.eucaim.org/hasBirthSex*

```
<https:/www.eucaim.org/hasAssociatedSex>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Sex?name=8507>,
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Sex?name=8532>;
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/sex>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Sex?name=8507>,
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Sex?name=8532>;
```

- *https://www.eucaim.org/hasImageModality*

```
<https:/www.eucaim.org/hasImageModality>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Imaging_modality?name=MR>;
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/imaging_modality>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Imaging_modality?name=MR>;
```

- *https://www.eucaim.org/hasEquipmentManufacturer*

```
<https:/www.eucaim.org/hasImageVendor>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Vendor?name=philips>,
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Vendor?name=siemens>,
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Vendor?name=ge>;
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/vendor>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Vendor?name=philips>,
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Vendor?name=siemens>,
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Vendor?name=ge>;
```

- **RDA-A1-01M - Metadata contain information on how the user can access the data (33%)**: As described in Table 6, this indicator linked to principle A1 is not detecting the `image_access_uri` and the `publication_uri`. In this particular case, this is because the preparation of the dataset in the UPV reference node is still in progress, and therefore the dataset has not yet been published. Consequently, these terms are currently missing values that will be completed as soon as they become available.

Table 6. RDA-A1-01M.

Indicator Level	**Recommendable**
Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator refers to the information necessary to allow the requester to access the digital object. This includes (i) whether there are restrictions to accessing the data (i.e., whether access to the data is open, restricted, or closed), (ii) the actions an interested person must take to access the data, particularly if the data are not published on the Web, and (iii) specify that the resources are available via eduGAIN or via specialized solutions as proposed for EPOS.
Technical Implementation	Checks whether the metadata include the necessary information to enable data access. It verifies the presence of key access-related terms such as image_access_type, image_access_description, image_access_uri and publication_uri. It also ensures that certain elements, like URIs, are resolvable.
Technical feedback	1/3 access elements satisfied. Missing: image_access_uri, publication_uri.

Tips	Include the terms: image_access_type, image_access_description, image_access_uri and publication_uri making sure the URIs are resolvable.

- **RDA-I2-01M - Metadata use FAIR vocabularies (75%)**:

As described in Table 7, this indicator relates to principle I2, which requires (meta)data to use globally unique and resolvable identifiers. As a result, it was found that two external RDF namespaces are not resolving correctly (e.g., cube#, dc/terms/), both returning HTTP 503 errors. These failures were most likely caused by the high number of requests performed within a short time interval during the batch execution. As it can be seen in Annex 3, this issue has also been corrected in the second execution, obtaining a score of 100%.

Table 7. RDA-I2-01M.

Indicator Level	**Recommendable**
Indicator Assessment	This indicator is linked to the following principle I2: (meta)data use vocabularies that follow FAIR principles. The indicator requires that the vocabulary used in the metadata adheres to FAIR principles and is at least documented and resolvable via globally persistent and unique identifiers. The documentation must be easy to access and find.
Technical Implementation	Extracts HTTP(S) namespace URIs used in the RDF graph and checks if they are resolvable via HTTP requests. A score is assigned proportional to the percentage of vocabularies accessible online.
Technical feedback	Namespaces resolving (6/8): • http://www.w3.org/ns/dcat# (200 text/turtle) • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/ (200 text/turtle) • https://w3id.org/fdp/fdp-o# (200 text/plain) • http://www.w3.org/2001/XMLSchema# (200 text/html) • http://www.w3.org/1999/02/22-rdf-syntax-ns# (200 text/turtle) • http://www.w3.org/2000/01/rdf-schema# (200 text/turtle) Namespaces not resolving (2/8): • http://purl.org/linked-data/cube# (503 text/html) • http://purl.org/dc/terms/ (503 text/html)
Tips	Ensure vocabularies used in RDF (namespace prefixes like dcterms, dcat, foaf, etc.) are publicly accessible via the Internet.

- **RDA-I3-01M - Metadata include references to other metadata (50%), RDA-I3-03M - Metadata include qualified references to other metadata (50%) & RDA-R1.2-02M - Metadata include provenance using a community-aligned, cross-domain format (0%):**

These indicators relate to principle I3 and principle R1, and the issue arises because, in some cases, the metadata is using literals instead of URIs. This is explained in more detail in Section *4.3 Ensure machine-readable (missing IRIs or typed literals).*

Table 8. RDA-I3-01M.

Indicator Level	**Recommendable**
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. The indicator concerns how metadata connect to other metadata—for example, via links to information about organizations, people, places, projects, or time periods related to the

	digital object described by the metadata.
Technical Implementation	Checks whether terms like relation or contributor are present in the RDF and contain qualified references, such as resolvable URIs, blank nodes, or dictionary-type structures. The score is based on the percentage of terms meeting these conditions.
Technical feedback	Relation terms with qualified links: 2/4 Terms present but **not** linked: • http://purl.org/dc/terms/publisher • http://purl.org/dc/terms/provenance
Tips	Ensure the RDF includes qualified references in the following fields: 'http://purl.org/dc/terms/publisher', 'http://www.w3.org/ns/dcat#contactPoint', 'http://purl.org/dc/terms/spatial', 'http://purl.org/dc/terms/temporal', 'http://purl.org/dc/terms/provenance', 'http://purl.org/dc/terms/relation', 'http://purl.org/dc/terms/isPartOf', 'http://www.w3.org/ns/adms#sample', 'http://www.w3.org/ns/dcat#accessURL', 'http://www.w3.org/ns/dqv#hasQualityAnnotation'. These references must be expressed as valid URIs, blank nodes, or structured dictionaries to count as qualified.

Table 9. RDA-I3-03M.

Indicator Level	**Recommendable**
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. This indicator concerns how metadata connect with other metadata—for example, with descriptions of related resources that provide additional context to the data.
Technical Implementation	This indicator reuses the logic of rda_i3_01m.
Technical feedback	Relation terms with qualified links: 2/4 Terms present but **not** linked: • http://purl.org/dc/terms/publisher • http://purl.org/dc/terms/provenance
Tips	Ensure the RDF includes qualified links in the following fields: 'http://purl.org/dc/terms/publisher', 'http://www.w3.org/ns/dcat#contactPoint', 'http://purl.org/dc/terms/spatial', 'http://purl.org/dc/terms/temporal', 'http://purl.org/dc/terms/provenance', 'http://purl.org/dc/terms/relation', 'http://purl.org/dc/terms/isPartOf', 'http://www.w3.org/ns/adms#sample', 'http://www.w3.org/ns/dcat#accessURL', 'http://www.w3.org/ns/dqv#hasQualityAnnotation'. These references should clearly indicate the relationship type (e.g., creator, editor, or geographic coverage) and be expressed as valid URIs, blank nodes, or structured dictionaries.

Table 10 (RDA-R1.2-02M).

Indicator Level	**Optional**
Indicator Assessment	This indicator is linked to the following principle R1.2: (Meta)data are associated with detailed provenance. It requires that provenance information in metadata follow a cross-domain machine-understandable format.
Technical Implementation	Checks whether the RDF includes provenance information using a machineunderstandable value (IRI or typed literal). Evaluates the value of

	the configured provenance field; if it's a URI or typed literal, 100 points are awarded.
Technical feedback	Provenance is present but its value is not machine-understandable (must be an IRI or a typed literal).
Tips	Ensure that 'http://purl.org/dc/terms/provenance' contains either a URI (e.g., DOI or URL) or a typed literal (e.g., date with 'xsd:date'). This allows automatic processing by systems.

- **RDA-I3-01D - Data include references to other data (0%)**:

As described in Table 11, this indicator relates to principle I3 and shows there are *"No references to other data".* In this case, it is unlikely that any references would exist, but the indicator has been included for completeness, as the FAIR-EVA tool follows the FAIR Data Maturity Model from the RDA. This indicator evaluates the interconnection of the dataset with other datasets, either via identifiers of the data itself or of the metadata describing it. Such interconnections add value, but not all datasets are expected to have them. Examples where this could apply include derived data pointing to the original data, a longitudinal study where a time point references previous time points, or the use of external data for processing (e.g., an atlas like MNI). However, this is not mandatory, and the RDA itself classifies this indicator as *Useful*, the lowest level of importance assigned.

Table 11. RDA-I3-01D.

Indicator Level	**Optional**
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. This indicator concerns how the data connect to other data—for example, linking to prior or related research data that provide additional context to the data.
Technical Implementation	This test implements the logic used in RDA-I3-02M.
Technical feedback	No references to other data.
Tips	rda_i3_01d.tips

- **RDA-R1.1-02M - Metadata include references to a standard license under which the data are published for reuse (0%)**:

As described in Table 12, this indicator requires the use of a standard license to ensure clear terms of use. However, in the current model, three values are recorded for the `dcterms:rights` field: `IN_SITU, REMOTE, and DOWNLOAD.`

```
dcterms:rights
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Rights?name=IN_SITU>;
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/rights>
<http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Rights?name=IN_SITU>;
```

Table 12. RDA-R1.1-02M.

Indicator Level	**Recommendable**
Indicator Assessment	This indicator requires that the reference to reuse conditions be a standard

	license, rather than a locally defined one. The indicator can be evaluated by checking that the license is indeed a standard license. Examples include Creative Commons and Open Data Commons licenses.
Technical Implementation	Checks the results of rda_r1_1_01m to identify any standard license, whether provided as a link or literal value.
Technical feedback	Community-standard requirement not satisfied. Consider using a standard license.
Tips	Include the resource standard license in the dcterms:rights metadata field if available.

- **RDA-R1.3-01M - Metadata comply with community standards (46%) & RDA-R1.3-02M - Metadata are machine-readable and comply with community standards (46%)**

  These indicators show that the current implementation of the FAIR Data Point for the metadata catalogue is not compliant with the HealthDCAT-AP profile, as noted previously.

  Table 13 (RDA-R1.3-01M)

Indicator Level	**Essential**
Indicator Assessment	This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires metadata to comply with standards established in their domain.
Technical Implementation	Checks whether metadata conform to the 'healthdcat-ap' profile, evaluated via the result from rda_i1_01m.
Technical feedback	Community-standard requirement not satisfied using Health-DCAT-AP < 50 points.
Tips	Ensure the RDF complies with the HealthDCAT-AP profile

  Table 14 (RDA-R1.3-02M)

Indicator Level	**Essential**
Indicator Assessment	This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires that metadata comply with a community standard in a machine-understandable format.
Technical Implementation	Evaluates whether metadata comply with a machine-understandable community standard using the score obtained for the HealthDCAT-AP profile in rda_i1_01m.
Technical feedback	Metadata is not expressed in any machine-understandable community standard. Please check the Health-DCAT-AP. Health-DCAT-AP < 50 points.
Tips	Ensure the RDF is structured according to the HealthDCAT-AP profile.

In conclusion, the analysis of these indicators provides a comprehensive overview of the current status of the metadata catalogue in terms of FAIR compliance. While several indicators show high adherence to FAIR principles (see the complete report in Annex 3 - FAIR report example: PerProGlio metadata), the

evaluation also highlights specific aspects requiring improvement, such as missing values, RDF property mismatches, unresolved namespaces, and the standardization of licensing fields.

4. Corrective actions to be done

After analysing the results of the FAIR-EVA tool, we made efforts to determine why some datasets received low scores, even though all available information had been filled in the catalogue metadata. This was found to be caused partly due to the version of the Molgenis software used (https://molgenis.org/). Molgenis is the software of the catalogue. In particular, in the current version, URIs of properties are not displayed as expected in the RDF endpoint of the catalogue. This functionality is fixed in the newest released version (pending updating, see 2 below).

Hence, the corrective actions we plan to execute are summarized below.

1- Align the current metadata model with the latest hyper-ontology model of EUCAIM

We have observed many discrepancies between the defined hyper-ontology and the WP5 definition of the metadata model based on Health DCAT-AP, and the current version of the metadata model exposed in the EUCAIM Public Catalogue using the Molgenis software. Many property values are empty or do not contain the URI value from the corresponding vocabulary as defined in the hyperontology documentation. Our goal is to keep track of both the catalogue metadata model and hyper-ontology definition and keep aligning them recurrently. The frequency of update/alignment is still under discussion.

2- Update the version of Molgenis software

Molgenis is an open-source software continuously changing and expanding to include more functionalities. The initial version of the EUCAIM public catalogue was built by using the BBMRI-ERIC Directory application as a reference. Since the launch of the catalogue, there have been new major version releases of Molgenis with new functionalities. Upgrading should automatically solve the RDF output, which is why we are already working on upgrading and testing changes in a local environment. Due to the fact that this is a major version upgrade, there are breaking changes, e.g., the front-end and back-end now use a new NodeJS and Java version respectively, resulting in additional development and required testing. Once testing is ready, we plan on deploying the new version in the development and production environment.

3- Replace literals with object and URI property values where required

In the current FAIR Data Point implementation of the metadata catalogue, there is a general modelling issue concerning the use of literals where the underlying ontology expects links to objects (resources). Certain properties are defined in the EUCAIM DCAT-AP and Health DCAT-AP with ranges such as “Agent”, or other classes, meaning that they are intended to point to RDF resources that can themselves be further described (i.e., object properties). In practice, however, these properties are sometimes populated with plain string literals instead of resource identifiers. A typical example is a property such as dct:publisher. In the current implementation one can encounter a triple of the form: :dataset1 dct:publisher "Example Hospital". Here the value of dct:publisher is a literal, which RDF treats as a terminal node: it cannot have its own properties, identifiers, links or multilingual labels. The intention of EUCAIM DCAT-AP and Health DCAT-AP however, is that such a property points to an object that represents the publisher as an Agent, for instance:

:dataset1 dct:publisher :example-hospital .
:example-hospital a foaf:Agent ;

skos:prefLabel "Example Hospital"@en ;
dct:identifier "123456789" .

This pattern allows the catalogue to attach further information to the publisher, such as persistent identifiers, contact details, alternative names, language-specific labels and links to external registries or files. Validation frameworks such as SHACL which FAIR-Eva uses, flags these cases as non-conformant, and consuming applications that assume an Agent resource will either fail or need special-case handling.

To address this issue, the implementation will be revised so that properties whose range is an Agent (or other non-literal class) are consistently modelled with objects rather than literals. This will be done by restructuring the internal database tables in the Molgenis software. For the example mentioned above, a table with 'Agents' can be added, in which we can create the ':example-hospital' resource. The property 'dct:publisher' of the Dataset, can then point to this resource. That resource will carry human-readable labels, identifiers and any additional metadata required. The example above would then be represented with dct:publisher pointing to :example-hospital, which is explicitly typed as an Agent and annotated with a preferred label and identifier. After indicating in Molgenis that all objects from this 'Agents' table are 'foaf:Agent'-s, these resources will be presented in the RDF endpoint as such. The current implementation already uses this functionality, but alignment with the designed metadata model is needed.

4- Ensure there are not missing values

It has been identified that some datasets contain missing values in mandatory indicators. The fields most frequently missing are interoperabilityLevel (36 datasets), topography (32), hasImageVendor (13), hasImageBodyPart (6), and hasCondition (5). The corresponding Data Holders will need to review these metadata entries with missing values and provide the required information so that the catalogue can be updated accordingly.

5. Improvements implemented

To support the integration of FAIR-EVA within the EUCAIM ecosystem, a substantial number of improvements and extensions have been developed during this reporting period. These enhancements focus on increasing the accuracy of the FAIR assessments, ensuring full compatibility with EUCAIM metadata standards, and improving the usability and deployability of the tool across different environments. The main developments carried out are summarised below:

- Development of a dedicated EUCAIM-specific FAIR-EVA plugin, including 41 custom tests aligned with EUCAIM metadata requirements and controlled vocabularies.
- Integration of HealthDCAT-AP and DCAT-AP validation profiles into the evaluation workflow.
- Design and implementation of EUCAIM-specific SHACL shapes, including controlled vocabulary checks.
- Development of a full web client for FAIR-EVA, including page design, multilingual support, and interactive reporting.
- Inclusion of Tier computation and visualisation, integrating the FAIR-EVA tiering model into both the CLI and web outputs.
- Implementation of a batch runner for automated nightly evaluations, including generation of .txt and .csv outputs for all datasets in the EUCAIM Public Catalogue.
- Support for multiple RDF evaluation schemas, including the legacy schema to ensure backward compatibility with existing metadata.
- Improved and modular Dockerisation, separating the FAIR-EVA service, CLI and web clients into independent containers for easier deployment and maintenance.
- Implementation of PDF report generation.

- Ensure persistent, globally unique identifiers: EUCAIM uses UUID v4 identifiers for the datasets. The dataset identifiers are critical to ensure interoperability among the different applications (catalogue, federated search, negotiator and Reference Nodes) which must be coherent along the different applications. By definition, the chances of having a collision between two UUID identifiers are extremely low, and they could be detected when registering the dataset in the catalogue. The procedure starts asking the data holder in federated nodes if they already have a UUID and reusing it, or generating a new one otherwise.

  In order to make the UUIDs universally unique PIDs (Persistent Identifiers), we have created a prefix that redirects (doi-like) the PIDs to the catalogue pages (facilitating dataset referencing) so the ids can be considered totally unique. The redirection is https://pid.eucaim.cancerimage.eu/<<<UID>>, so for example, for the dataset whose UUID is 822ad0bd-02d1-4932-a8b2-7d5679c3d4f0 ("Lung Cancer (April 2024)" in the CHAIMELEON Repository), the PID https://pid.eucaim.cancerimage.eu/822ad0bd-02d1-4932-a8b2-7d5679c3d4f0 will provide a unique and discoverable reference.

6. Impact and Sustainability

The results of the FAIR assessment provide the first quantitative baseline of FAIR data maturity in the EUCAIM federation. By analysing 82 datasets from 13 providers using the RDA indicator framework. The project now has objective evidence of where improvements are needed and which corrective measures will have the strongest impact.

Three systemic aspects have immediate strategic implications:

1. **Identifier:**

The absence of persistent and globally unique identifiers is the single largest barrier to achieving Tier 1 FAIR compliance. This indicates the need for a cross-project workflow for DOI or UUID assignment at dataset onboarding.

2. **Metadata standardisation:**

While the metadata infrastructure is operational, standardized licensing, provenance, and controlled vocabularies are not yet implemented at scale.

Addressing these gaps will improve automated discovery, access control, and interoperability with research infrastructures.

3. **Catalogue alignment and semantic governance:**

The evaluation confirms the importance of maintaining consistent metadata and RDF schema definitions across catalogue versions.

Catalogue changes are coordinated through the WP4 and metadata alignment uses WP5 and hyper-ontology definitions to ensure continuity.

From an operational perspective, FAIR monitoring will support:

- Providing measurable indicators of metadata quality.
- WP2 engagement actions by identifying concrete improvements for Data Holders.
- Sustainability by establishing reproducible FAIR scoring processes.

Finally, this FAIR baseline creates a repeatable evaluation model that can be applied to newly onboarded datasets. Tracking improvements over time will demonstrate measurable progress in the FAIR maturity of EUCAIM datasets and reduce integration challenges.

Annex 1 - FAIR-EVA for EUCAIM: Implementation of the RDA indicators

Author: Inés Victoria Rodríguez (IFCA, CSIC)

Version: 1.0

Date: 2nd of October 2025

This annex summarises the RDA FAIR indicators implemented. Indicators are classified by relevance:

- **Essential** indicators correspond to requirements that are fundamental for achieving FAIRness. If these criteria are not met, compliance with the FAIR principles would be severely compromised.
- **Recommendable** indicators refer to aspects that may not be strictly mandatory, but whose fulfilment leads to significant improvement in the overall level of FAIRness.
- **Useful** indicators capture complementary or desirable features that enhance FAIRness but are not essential.

Summary table:

Indicator Code	EN Title	Level	Weight
RDA-F1-01M	Metadata is identified by a persistent identifier	Essential	3
RDA-F1-01D D	Data is identified by a persistent identifier	Essential	3
RDA-F1-02M	Metadata is identified by a globally unique identifier	Essential	3
RDA-F1-02D	Data is identified by a globally unique identifier	Essential	3
RDA-F2-01M	Rich metadata is provided to allow discovery	Essential	3
RDA-F3-01M	Metadata includes the identifier for the data	Essential	3
RDA-F4-01M	Metadata is offered in such a way that it can be harvested and indexed	Essential	3

RDA-A1-01M	Metadata contains information to enable the user to get access to the data	Recommendable	2
RDA-A1-02M	Metadata can be accessed manually (i.e. with human intervention)	Essential	3
RDA-A1-02D	Data can be accessed manually (i.e. with human intervention)	Essential	3
RDA-A1-03M	Metadata identifier resolves to a metadata record	Essential	3
RDA-A1-03D	Data identifier resolves to a digital object	Essential	3
RDA-A1-04M	Metadata is accessed through standardised protocol	Essential	3
RDA-A1-04D	Data is accessible through standardised protocol	Essential	3
RDA-A1-05D	Data can be accessed automatically (i.e. by a computer program)	Recommendable	2
RDA-A1.1-01M	Metadata is accessible through a free access protocol	Recommendable	2
RDA-A1.1-01D	Data is accessible through a free access protocol	Essential	3
RDA-A1.2-01D	Data is accessible through an access protocol that supports	Optional	1

	authentication and authorisation		
RDA-A2-01M	Metadata is guaranteed to remain available after data is no longer available	Essential	3
RDA-I1-01M	Metadata uses knowledge representation expressed in standardised format	Recommendable	2
RDA-I1-01D	Data uses knowledge representation expressed in standardised format	Recommendable	2
RDA-I1-02M	Metadata uses machine-understandable knowledge representation	Recommendable	2
RDA-I1-02D	Data uses machine-understandable knowledge representation	Recommendable	2
RDA-I2-01M	Metadata uses FAIR-compliant vocabularies	Optional	1
RDA-I2-01D	Data uses FAIR-compliant vocabularies	Recommendable	2
RDA-I3-01M	Metadata includes references to other metadata	Optional	1
RDA-I3-01D	Data includes references to other data	Recommendable	2

RDA-I3-02M	Metadata includes references to other data	Optional	1
RDA-I3-02D	Data includes qualified references to other data	Optional	1
RDA-I3-03M	Metadata includes qualified references to other metadata	Recommendable	2
RDA-I3-04M	Metadata include qualified references to other data	Optional	1
RDA-R1-01M	Plurality of accurate and relevant attributes are provided to allow reuse	Essential	3
RDA-R1.1-01M	Metadata includes information about the licence under which the data can be reused	Essential	3
RDA-R1.1-02M	Metadata refers to a standard reuse licence	Recommendable	2
RDA-R1.1-03M	Metadata refers to a machine-understandable reuse licence	Recommendable	2
RDA-R1.2-01M	Metadata includes provenance information according to community-specific standards	Recommendable	2
RDA-R1.2-02M	Metadata includes provenance information according to a crosscommunity language	Optional	1
RDA-R1.3-01M	Metadata complies with a community standard	Essential	3

RDA-R1.3-01D	Data complies with a community standard	Essential	3
RDA-R1.3-02M	Metadata is expressed in compliance with a machine understandable community standard	Recommendable	2
RDA-R1.3-02D	Data is expressed in compliance with a machine-understandable	Essential	3

RDA-F1-01M

Metadata are identified by a persistent identifier.

- Indicator:

This indicator is linked to the following principle: F1 — (meta)data are assigned a globally unique and eternally persistent identifier. This indicator evaluates whether metadata are identified by a persistent identifier. A persistent identifier ensures that the metadata can be found over time and reduces the risk of broken links.

- Technical Description:

Searches through a predefined list of possible metadata terms (dcterms:identifier) to determine whether any such information is available in the metadata.

- Tips:

Metadata need a unique identifier, defined in dcterms:identifier

RDA-F1-01D

Data are identified by a persistent identifier.

- Indicator:

This indicator is linked to the following principle: F1 — (meta)data are assigned a globally unique and eternally persistent identifier. This indicator evaluates whether the data are identified by a persistent identifier. A persistent identifier guarantees that the data can be found over time and reduces the risk of broken links.

- Technical Description:

Right now, metadata id and data id are the same id. Searches through a predefined list of possible metadata terms (dcterms:identifier) to determine whether any such information is available in the metadata.

- Tips:

If a DOI has not yet been generated, include it in the metadata record to indicate the external site where the data files are located.

RDA-F1-02M

Metadata are identified by a globally persistent identifier.

- Indicator:

This indicator is linked to the following principle: F1 — (meta)data are assigned a globally unique and eternally persistent identifier. The indicator evaluates whether the metadata's identifier is globally unique, i.e., that no two identical identifiers identify different metadata records.

- Technical Description:

Checks the type of identifiers using the idutils package. Identifies valid identifiers. If it is not a persistent identifier (e.g. a UUID), the test fails.

- Tips:

dcterms:identifier needs to be a globally persistent identifier (handle, DOI, ARK, PURL…).

RDA-F1-02D

Data are identified by a globally persistent identifier.

- Indicator:

This indicator is linked to the following principle: F1 — (meta)data are assigned a globally unique and eternally persistent identifier. The indicator evaluates whether the data's identifier is globally unique, that is, that no two datasets use the same identifier.

- Technical Description:

Checks that the data identifier is globally unique and eternally persistent (UUID or accepted by idutils).

RDA-F2-01M

There is sufficiently rich metadata to enable resource discovery

- Indicator:

This indicator is linked to the following principle F2: Data are described with rich metadata. The indicator concerns not only the presence of metadata but also the amount provided and how such metadata supports resource discovery.

- Technical Description:

Compares the metadata elements present with a configured list considered essential for discoverability. Calculates the proportion of elements found relative to those expected to assign a score.

- Tips:

Add all fields related to the findability of the dataset in the catalogue: title, description, identifier, hasCondition, topography, ImageBodyPart, geographical coverage, collection method, number of subjects, birth sex, image modality, vendor/manufacturer, image access type, provider and interoperability level.

RDA-F3-01M

Metadata include the identifier of the data.

- Indicator:

This indicator is linked to the following principle F3: Metadata clearly and explicitly include the identifier of the data they describe. The indicator concerns the inclusion of the reference (i.e., the identifier) of the digital object in the metadata so that the digital object can be accessed.

- Technical Description:

Verifies the metadata element where the object is identified dcterms:identifier

RDA-F4-01M

Metadata are provided in such a way that they can be harvested and indexed.

- Indicator:

This indicator is linked to the following principle: F4 — (meta)data are registered or indexed in a searchable resource. The indicator tests whether the metadata are offered in a way that allows indexing. In some cases, metadata may be provided alongside the data to a local institutional repository or to a regional or domain-specific portal, or be included on a landing page where a search engine can harvest them. The indicator remains purposely broad so as not to limit how and by whom the data may be collected and indexed.

- Technical Description:

Checks whether the metadata can be harvested.

RDA-A1-01M

Metadata contain information on how the user can access the data.

- Indicator:

This indicator is linked to the following principle: A1 — (Meta)data are retrievable by their identifier using a standardized communication protocol. It evaluates whether the metadata include information that allows users to directly access the underlying data (e.g., through access URIs or persistent publication links).

- Technical Description:

Checks whether the metadata include the necessary information to enable data access. It verifies the presence of key access-related terms such as image_access_type, image_access_description, image_access_uri and publication_uri. It also ensures that certain elements, like URIs, are resolvable.

- Tips:

Provide an http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_uri pointing to a reliable, standardized access endpoint and a publication DOI in http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/publication_uri that resolves correctly. Make sure the links are active and accessible. Also include image_access_type and image_access_description.

RDA-A1-02M

Metadata are accessible manually.

- Indicator:

This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator refers to any human interaction necessary if the requester wishes to access the metadata. FAIR primarily addresses automated interactions where a machine can access the metadata, but metadata may also require human interactions. This may be important where the metadata contain confidential information. Human interaction may involve sending an email to the metadata owner or calling by phone to receive instructions.

- Technical Description:

Checks if there is an accessible reference (e.g., a URL) allowing manual retrieval of the metadata. If the link is defined and functional, 100 points are awarded.

RDA-A1-02D

Data are accessible manually.

- Indicator:

This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator refers to any human interaction required if the requester wishes to access the digital object. FAIR primarily addresses automated interactions where a machine can access the object, but there may also be digital objects that require human interactions, such as clicking a link on a landing page, sending an email to the data owner, or even calling by phone. The indicator can be evaluated by searching metadata for information describing how the digital object can be obtained through human intervention.

- Technical Description:

Checks if there is an accessible reference (e.g., a URL) allowing manual retrieval of the data. If the link is defined and functional, 100 points are awarded.

RDA-A1-03M

The metadata identifier resolves to a metadata record.

- Indicator:

This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. This indicator concerns the resolution of the metadata identifier. The identifier assigned to the metadata must be associated with a resolution service that allows access to the metadata record.

- Technical Description:

Verifies whether metadata could be extracted from its identifier.

RDA-A1-03D

The data identifier resolves to a digital object.

- Indicator:

This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. This indicator concerns the resolution of the identifier that identifies the digital object. The identifier assigned to the data must be associated with a formally defined retrieval/resolution mechanism that allows access to the digital object, or provides access instructions in cases of human-mediated access. The FAIR principle and this indicator say nothing about the mutability or immutability of the digital object identified; this is governed by the Data Holder's persistence policy.

- Technical Description:

Detects if a publication_uri is present. It checks if the publication_uri is typed as a DOI and if the URI is resolvable (checked via ut.check_link). If the DOI is present but not resolvable, points are set to 0. If no DOI is found, points are also 0.

- Tips:

Ensure the dataset has a DOI registered and included in the metadata as http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/publication_uri . Verify that the DOI is resolvable through https://doi.org/.... Broken or inactive links will yield zero points. If a DOI is not possible, consider alternative persistent resolvable identifiers (e.g., Handle, ARK), but ensure they follow a standardized resolution mechanism.

RDA-A1-04M

Metadata are accessible via a standard protocol.

- Indicator:

This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. More information about this principle can be found here. The indicator concerns the protocol used to access the metadata and requires that the protocol be defined in a standard.

- Technical Description:

Identifies the protocol scheme (e.g., http, https) used to access the metadata, and compares it with a list of protocols accepted by standards.

RDA-A1-04D

Data are accessible via a standard protocol.

- Indicator:

This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator concerns the protocol used to access the digital object and requires that the protocol be defined in a standard.

- Technical Description:

Searches metadata for an element corresponding to data access. Extracts the protocol (such as http, https) from the provided link and compares it with a list of standardized protocols. If at least one valid protocol is found, full score is awarded.

RDA-A1-05D

Data are accessible automatically.

- Indicator:

This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator refers to automated interactions between machines to access digital objects. The indicator will evaluate how machines interact and grant access to the digital object.

- Technical Description:

Evaluates whether the data are accessible automatically via a URL, checking whether it is resolvable and present in the metadata.

RDA-A1.1-01M

Metadata are accessible via an open, free protocol.

- Indicator:

This indicator is linked to the following principle A1.1: The protocol is open, free, and universally applicable. The indicator tests that the protocol allowing access to metadata can be used freely. Free usage of the protocol enhances data reuse.

- Technical Description:

Same scoring as rda_a1_04m.

RDA-A1.1-01D

Data are accessible via a free protocol.

- Indicator:

This indicator is linked to the following principle A1.1: The protocol is open, free, and universally applicable. The indicator requires that the protocol can be used free of charge, facilitating unrestricted access.

- Technical Description:

Reuses the result from indicator RDA-A1-04D to check whether data access is via a free and open protocol (such as HTTP or HTTPS). If at least one is detected, 100 points are assigned.

RDA-A1.2-01D

Data are accessible via a protocol that supports authentication and authorization.

- Indicator:

This indicator is linked to the following principle A1.2: The protocol allows authentication and authorization when needed. The indicator requires how one can authenticate and authorize access to the digital object, and that access to the data is properly described and documented.

- Technical Description:

Checks whether metadata include information on authentication and authorization mechanisms required to access the data.

RDA-A2-01M

Metadata remain accessible even if the data are no longer available.

- Indicator:

This indicator is linked to the following principle A2: Metadata must remain accessible even when the data are no longer available. The indicator is intended to verify that information about a digital object is still available after the object has been removed or lost. Where possible, the metadata that remain available should also indicate why the object is no longer available.

- Technical Description:

Metadata are always subject to the EUCAIM Metadata Preservation Policy.

RDA-I1-01M

Metadata use some standardized form of knowledge representation

- Indicator:

This indicator is linked to the following principle I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. The indicator determines that an appropriate standard is used to express knowledge, e.g., controlled vocabularies for subject classifications.

- Technical Description:

Validates the metadata RDF graph against three SHACL profiles (DCAT-AP, Health-DCAT-AP, and EUCAIM). Calculates how many primary active NodeShapes comply with their constraints, recursively evaluating child properties. The score is the average of the three profiles.

- Tips:

It is recommended to describe the digital object using the controlled vocabulary Health-DCAT-AP and EUCAIM-specific terms.

RDA-I1-01D

Data use some standardized form of knowledge representation.

- Indicator:

This indicator is linked to the following principle I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. The indicator determines that an appropriate standard is used to express knowledge, particularly the data model and format.

- Technical Description:

Extracts data formats from the metadata ('availableFormats') and compares them with those officially registered as Internet Media Types by IANA. If at least one matches, 100 points are awarded.

- Tips:

Use IANA media types, for example DICOM.

RDA-I1-02M

Metadata use a machine-processable form of knowledge representation.

- Indicator:

This indicator is linked to the following principle I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. This indicator focuses on the machine-understandability of metadata. It means metadata must be legible—and thus interoperable—by machines without requirements like translators or specific mappings.

- Technical Description:

Verifies that metadata are represented as a valid, non-empty RDF graph. If the RDF graph contains at least one triple, the metadata are considered machine-understandable.

RDA-I1-02D

Data use a machine-processable form of knowledge representation.

- Indicator:

This indicator is linked to the following principle I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. This indicator focuses on machine-understandability of the data. It means the data must be legible—and thus interoperable—by machines without human-mediated translation or mappings.

- Technical Description:

Checks if the data format is a machine-processable form of knowledge representation.

- Tips:

Even though many formats are technically valid, it is recommended to use authorised formats such as DICOM, DICOM-SEG, and NIfTI.

RDA-I2-01M

Metadata use FAIR vocabularies.

- Indicator:

This indicator is linked to the following principle I2: (meta)data use vocabularies that follow FAIR principles. The indicator requires that the vocabulary used in the metadata adheres to FAIR principles and is at least

documented and resolvable via globally persistent and unique identifiers. The documentation must be easy to access and find.

- Technical Description:

Extracts HTTP(S) namespace URIs used in the RDF graph and checks if they are resolvable via HTTP requests. A score is assigned proportional to the percentage of vocabularies accessible online.

- Tips:

Ensure vocabularies used in RDF (namespace prefixes like dcterms, dcat, foaf, etc.) are publicly accessible via the Internet.

RDA-I2-01D

Data use FAIR vocabularies.

- Indicator:

This indicator is linked to the following principle: I2: (Meta)data use vocabularies that follow the FAIR principles. The indicator requires the controlled vocabulary used for the data to conform to the FAIR principles, and at least be documented and resolvable using globally unique.

- Technical Description:

Checks the configured vocabulary predicates and verifies that their objects are URIs or typed literals (not literals) and that these URIs are dereferenceable with acceptable content types.

- Tips:

Use resolvable HTTP(S) URIs from well-governed vocabularies (avoid free-text literals) in the following fields: hasImaeBodyPart, hasCondition, hasImageModality, hasImageVendor, hasAssociatedSex, type, theme, healthCategory and spatial.

RDA-I3-01D

Data use FAIR vocabularies.

- Indicator:

This indicator is linked to the following principle: I3: (Meta)data include qualified references to other (meta)data. More information about that principle can be found here. This indicator is about the way data is connected to other data, for example linking to previous or related research data that provides additional context to the data.

- Technical Description:

-

RDA-I3-02D

Data use FAIR vocabularies.

- Indicator:

This indicator is linked to the following principle: I3: (Meta)data include qualified references to other (meta)data. More information about that principle can be found here. The indicator is about the way that metadata is connected to other metadata, for example through links to information about organizations, people, places, projects or time periods that are related to the digital object that the metadata describes.

- Technical Description:

This test implies checking the presence of qualified references within the content of the data. As it is defined, its implementation is too costly.

RDA-I3-01M

Metadata include references to other metadata.

- Indicator:

This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. The indicator concerns how metadata connect to other metadata—for example, via links to information about organizations, people, places, projects, or time periods related to the digital object described by the metadata.

- Technical Description:

Checks whether terms like relation or contributor are present in the RDF and contain qualified references, such as resolvable URIs, blank nodes, or dictionary-type structures. The score is based on the percentage of terms meeting these conditions.

- Tips:

Ensure the RDF includes qualified references in the following fields: 'http://purl.org/dc/terms/publisher', 'http://www.w3.org/ns/dcat#contactPoint', 'http://purl.org/dc/terms/spatial', 'http://purl.org/dc/terms/temporal', 'http://purl.org/dc/terms/provenance', 'http://purl.org/dc/terms/relation', 'http://purl.org/dc/terms/isPartOf', 'http://www.w3.org/ns/adms#sample', 'http://www.w3.org/ns/dcat#accessURL', 'http://www.w3.org/ns/dqv#hasQualityAnnotation'. These references must be expressed as valid URIs, blank nodes, or structured dictionaries to count as qualified.

RDA-I3-02M

Data use FAIR vocabularies.

- Indicator:

This indicator is linked to the following principle: I3: (Meta)data include qualified references to other (meta)data. More information about that principle can be found here. This indicator is about the way metadata is connected to other data, for example linking to previous or related research data that provides additional context to the data. Please note that this is not about the link from the metadata to the data it describes.

- Technical Description:

-

RDA-I3-03M

Metadata include references to other metadata.

- Indicator:

This indicator is linked to the following principle: I3: (Meta)data include qualified references to other (meta)data. More information about that principle can be found here. The indicator evaluates how the metadata links to other metadata resources, such as organizations, people, projects, or datasets, **with a qualified relation** i.e., where the type of relation is explicit (e.g., creator, publisher, spatial coverage).

- Technical Description:

Reuses rda_i3_01 implementation

- Tips:

Ensure the RDF includes qualified references in the following fields: 'http://purl.org/dc/terms/publisher', 'http://www.w3.org/ns/dcat#contactPoint', 'http://purl.org/dc/terms/spatial', 'http://purl.org/dc/terms/temporal', 'http://purl.org/dc/terms/provenance', 'http://purl.org/dc/terms/relation', 'http://purl.org/dc/terms/isPartOf', 'http://www.w3.org/ns/adms#sample', 'http://www.w3.org/ns/dcat#accessURL', 'http://www.w3.org/ns/dqv#hasQualityAnnotation'. These references must be expressed as valid URIs, blank nodes, or structured dictionaries to count as qualified.

RDA-I3-04M

Data use FAIR vocabularies.

- Indicator:

This indicator is linked to the following principle: I3: (Meta)data include qualified references to other data. This indicator is about the way metadata is connected to other data. The references need to be qualified which means that the relationship role of the related resource is specified.

- Technical Description:

-

RDA-R1-01M

Certain attributes relevant for reuse are included.

- Indicator:

This indicator is linked to the following principle R1: (Meta)data are richly described with a plurality of accurate and relevant attributes. The indicator refers to both the quantity and quality of metadata provided to enhance data reuse.

- Technical Description:

Compares the RDF metadata with a configured list of terms considered relevant for reuse. The score is proportional to the number of elements found in the RDF relative to the expected total.

- Tips:

Ensure the RDF includes the following reuse-enhancing fields: 'http://purl.org/dc/terms/spatial', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/image_year_range', 'http://purl.org/dc/terms/type', 'http://www.w3.org/ns/dcat#theme', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/sample', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/quality_annotation', 'http://www.healthdcatap.org/minTypicalAge', 'http://www.healthdcatap.org/maxTypicalAge', 'http://www.healthdcatap.org/numberofRecords', 'http://www.healthdcatap.org/numberofUniqueIndividuals', 'https://www.eucaim.org/hasAssociatedSex', 'https:/www.eucaim.org/hasCondition', 'https:/www.eucaim.org/hasImageModality', 'https:/www.eucaim.org/hasImageBodyPart', 'https:/www.eucaim.org/hasImageVendor', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/collection_method', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/access_rights', 'https://w3id.org/dpv/dpv-skos#hasPurpose', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/version', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/interoperability_level'. These fields must be present with suitable values to consider the resource reusable.

RDA-R1.1-01M

Metadata include information about the license under which the data are published for reuse.

- Indicator:

This indicator is linked to the following principle R1.1: (Meta)data are released with a clear and accessible data usage license. This indicator concerns the information provided in the metadata about the conditions (e.g., obligations, restrictions) under which the data can be reused.

- Technical Description:

This indicator checks whether the metadata provide an explicit, clear, and accessible usage license (preferably standardized and resolvable), using dcterms:rights as the primary source. If no rights/license is present, it inspects image_access_type as contextual information (restricted/non-public access does not replace a license)

- Tips:

The metadata must include the usage license for the resource.

RDA-R1.1-02M

Metadata include references to a standard license under which the data are published for reuse.

- Indicator:

This indicator requires that the reference to reuse conditions be a standard license, rather than a locally defined one. The indicator can be evaluated by checking that the license is indeed a standard license. Examples include Creative Commons and Open Data Commons licenses.

- Technical Description:

Checks the results of rda_r1_1_01m to identify any standard license, whether provided as a link or literal value.

RDA-R1.1-03M

Metadata include machine-readable references to a standard license under which the data are published for reuse.

- Indicator:

This indicator is linked to the following principle R1.1: (Meta)data are released with a clear and accessible data usage license. This indicator concerns how the reuse license is expressed. Instead of a human-readable text, the license should be expressed in a machine-processable way, e.g., for automated searches.

- Technical Description:

Checks the results of rda_r1_1_01m to identify if the license is resolvable (must be a link).

RDA-R1.2-01M

Metadata include provenance information according to community standards.

- Indicator:

This indicator is linked to the following principle R1.2: (Meta)data are associated with detailed provenance. It requires that metadata include information about the data's origin, history, or workflow in a way that conforms to standards used in the producing community.

- Technical Description:

Checks whether the RDF contains at least one entry under the configured provenance term. If provenance information is found, 100 points are awarded.

- Tips:

Ensure your RDF includes the field 'http://purl.org/dc/terms/provenance' with information on the origin, history, or data generation process following community standards.

RDA-R1.2-02M

Metadata include provenance using a community-aligned, cross-domain format.

- Indicator:

This indicator is linked to the following principle R1.2: (Meta)data are associated with detailed provenance. It requires that provenance information in metadata follow a cross-domain machine-understandable format.

- Technical Description:

Checks whether the RDF includes provenance information using a machine-understandable value (IRI or typed literal). Evaluates the value of the configured provenance field; if it's a URI or typed literal, 100 points are awarded.

- Tips:

Ensure that 'http://purl.org/dc/terms/provenance' contains either a URI (e.g., DOI or URL) or a typed literal (e.g., date with 'xsd:date'). This allows automatic processing by systems.

RDA-R1.3-01M

Metadata comply with community standards.

- Indicator:

This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires metadata to comply with standards established in their domain.

- Technical Description:

Checks whether metadata conform to the 'healthdcat-ap' profile, evaluated via the result from rda_i1_01m.

- Tips:

Ensure the RDF complies with the HealthDCAT-AP profile.

RDA-R1.3-01D

Data comply with community standards.

- Indicator:

This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires the data to comply with relevant community standards.

- Technical Description:

Checks whether the field configured as 'terms_community_standard' contains any recognized community formats like DICOM, DICOM-SEG, or NIfTI. If any are found, 100 points are awarded.

- Tips:

Ensure the RDF includes 'http://purl.org/dc/terms/format' with one of the following: 'DICOM', 'DICOM-SEG', or 'NIfTI'.

RDA-R1.3-02M

Metadata are machine-readable and comply with community standards.

- Indicator:

This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires that metadata comply with a community standard in a machine-understandable format.

- Technical Description:

Evaluates whether metadata comply with a machine-understandable community standard using the score obtained for the HealthDCAT-AP profile in rda_i1_01m.

- Tips:

Ensure the RDF is structured according to the HealthDCAT-AP profile.

Annex 2 - Results .csv example: HULAFE's PerProGlio's dataset

dataset_id	dataset_name	provider	table_kind	indicator_or_tier	principle	score	output
aaaadcw3tn2ydbsux2urluqaae	PerProGlio Cohort	PerProGlio	indicators	RDA_F1_01D	findable	0	Identifier is not persistent for the data: aaaadcw3tn2ydbsux2urluqaae

aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_F1_01M	findable	0	Identifier is not persistent for the metadata: aaaadcw3tn2ydbsux2urluqaae
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_F1_02D	findable	0	Identifier found for the data is not globally unique: aaaadcw3tn2ydbsux2urluqaae
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_F1_02M	findable	0	Identifier found for the metadata is not globally unique: aaaadcw3tn2ydbsux2urluqaae
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_F2_01M	findable	100	Findability terms present (15/15): • http://purl.org/dc/terms/title • http://purl.org/dc/terms/description • http://purl.org/dc/terms/identifier • https:/www.eucaim.org/hasCondition • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/topography • https:/www.eucaim.org/hasImageBodyPart • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/geographical_coverage • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/collection_method • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/number_of_subjects • https:/www.eucaim.org/hasAssociatedSex • https:/www.eucaim.org/hasImageModality • https:/www.eucaim.org/hasImageVendor • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_type • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/provider • http://www.w3.org/ns/adms#interoperabilityLevel Findability terms missing (0/15): • —
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_F3_01M	findable	100	Metadata includes identifier/s for the data: aaaadcw3tn2ydbsux2urluqaae
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_F4_01M	findable	100	Metadata is gathered programmatically through HTTP (API REST), thus can be harvested and indexed.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_01M	accessible	25	1/4 access elements satisfied. Missing: image_access_description, image_access_uri, publication_uri.

aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_02D	accessible	100	Documentation for the manual obtention of the data can be found in https://catalogue.eucaim.cancerimage.eu/
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_02M	accessible	100	Documentation for the manual obtention of the metadata can be found in https://catalogue.eucaim.cancerimage.eu/
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_03D	accessible	0	No DOI found for the data
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_03M	accessible	100	Metadata record could be retrieved from metadata identifier: aaaadcw3tn2ydbsux2urluqaae
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_04D	accessible	0	Could not check data access protocol: metadata element http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_uri
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_04M	accessible	100	Found a standarised protocol to access the metadata record: https
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_05D	accessible	0	Could not check data access protocol: metadata element http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_uri not found or not accessible
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_1_01D	accessible	0	Could not check data access protocol: metadata element http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_uri
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_1_01M	accessible	100	Found a free protocol to access the metadata record: https
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A1_2_01D	accessible	0	At the time, does not provide authentication or authorisation protocols
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_A2_01M	accessible	100	EUCAIM Metadata preservation policy
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I1_01D	interoperable	100	Data uses a correct way to present information present in https://www.iana.org/assignments/media-types/media-types.xhtml

aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I1_01M	interoperable	80,2	dcat-ap: 94.44 pts - failing properties 2/36 [dct:provenance, dct:publisher] healthdcat-ap: 46.15 pts - failing properties 7/26 [dcat:distribution, dcatap:applicableLegislation, dct:accessRights, dct:publisher, dct:publisher, healthdcatap:hdab, healthdcatap:healthCategory] eucaim: 100.00 pts - failing properties 0/15 [-]
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I1_02D	interoperable	100	1/1 machine-readable data model values found.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I1_02M	interoperable	100	Metadata is valid RDF with 87 triples.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I2_01D	interoperable	100	Data vocabulary objects checked: 9 • Resolvable IRIs: 9 • Literals (0): • — • No resolvable IRIs (0): • —
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I2_01M	interoperable	87	Namespaces resolving (7/8): • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/ (200 text/turtle) • http://www.w3.org/2000/01/rdf-schema# (200 text/turtle) • http://purl.org/linked-data/cube# (200 text/plain) • http://www.w3.org/1999/02/22-rdf-syntax-ns# (200 text/turtle) • https://w3id.org/fdp/fdp-o# (200 text/plain) • http://www.w3.org/ns/dcat# (200 text/turtle) • http://www.w3.org/2001/XMLSchema# (200 text/html) Namespaces not resolving (1/8): • http://purl.org/dc/terms/ (500 text/html)
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I3_01D	interoperable	0	No references to other data.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I3_01M	interoperable	50	Relation terms with qualified links: 2/4 Terms present but **not** linked: • http://purl.org/dc/terms/publisher • http://purl.org/dc/terms/provenance
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I3_02D	interoperable	0	This test implies checking the presence of qualified references within the content of the data. As it is defined, its implementation is too costly.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I3_02M	interoperable	0	No references to other data.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I3_03M	interoperable	50	Relation terms with qualified links: 2/4 Terms present but **not** linked: • http://purl.org/dc/terms/publisher • http://purl.org/dc/terms/provenance

aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_I3_04M	interoperable	0	No references to other data.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_R1_01M	reusable	90	Found 19/21 metadata elements that enhance reusability. Missing: ['http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/quality_annotation', 'https://www.w3.org/TR/vocab-adms/#adms-sample']
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_R1_1_01M	reusable	100	License/rights detected in dcterms:rights. Non-standard but resolvable: http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Rights?name=IN_SITU (Internal policy detected; consider mapping to a community license). Access context: image_access_type=http://catalogue.eucaim. cancerimage.eu/Eucaim/api/rdf/Image_access_type?name=NON_PUBLIC, http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Image_access_type?name=RESTRICTED_ACCESS.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_R1_1_02M	reusable	0	Community-standard requirement not satisfied. Consider using a standard license.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_R1_1_03M	reusable	100	Machine-understandable requirement satisfied. Non-standard but resolvable: http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Rights?name=IN_SITU Non-standard but resolvable: http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Image_access_type?name=NON_PUBLIC, Non-standard but resolvable: http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Image_access_type?name=RESTRICTED_ACCESS.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_R1_2_01M	reusable	100	Found provenance information in the metadata.
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_R1_2_02M	reusable	0	Provenance is present but its value is not machine-understandable (must be an IRI or a typed literal).
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_R1_3_01D	reusable	100	Dataset format complies with community standards (DICOM, DICOM-SEG, nifti).
aaaadcw3tn2ydbsux2urluqaae	PerProGlioCohort	PerProGlio	indicators	RDA_R1_3_01M	reusable	46	Community-standard requirement not satisfied using Health-DCAT-AP < 50 points.

aaaadcw3tn2ydbsux2urluqaae	PerProGlio Cohort	PerProGlio	indicators	RDA_R1_3_02D	reusable	100	Dataset format complies with machine-understandable community standard (DICOM, DICOM-SEG, nifti).
aaaadcw3tn2ydbsux2urluqaae	PerProGlio Cohort	PerProGlio	indicators	RDA_R1_3_02M	reusable	46	Metadata is not expressed in any machine-understandable community standard. Please check the Health-DCAT-AP. Health-DCAT-AP < 50 points.
aaaadcw3tn2ydbsux2urluqaae	PerProGlio Cohort	PerProGlio	tier	No Tier, since it did not meet the requirements for Tier 1			Tier 1 not passed. Failing indicators: RDA_F1_01M, RDA_F1_01D

Annex 3 - FAIR report example: PerProGlio metadata

FAIReva

FAIR-EVA Evaluation Report

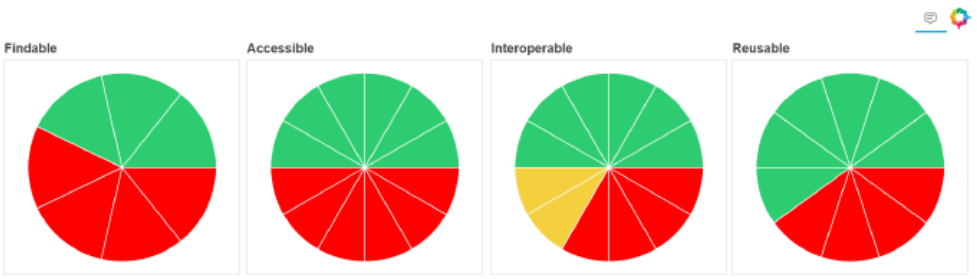
This FAIR assessment was generated as part of the **EUCAIM** infrastructure. The evaluation process ensures that data resources meet key FAIR principles — *Findable*, *Accessible*, *Interoperable*, and *Reusable* — and assigns a **Tier level** based on the set of RDA indicators successfully fulfilled.

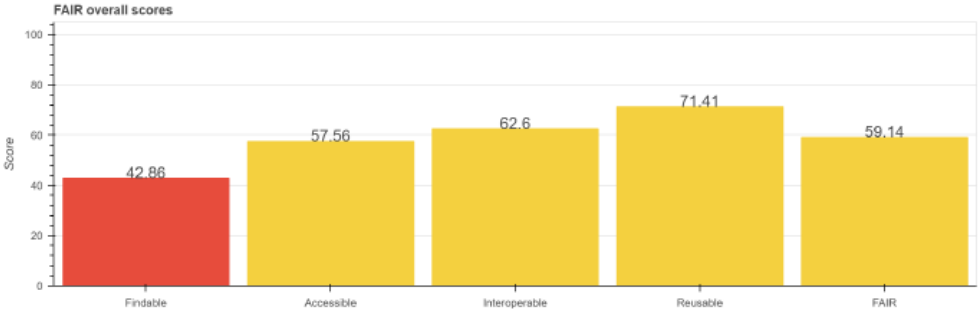
Evaluation for aaaadcw3tn2ydbsux2urluqaae (aaaadcw3tn2ydbsux2urluqaae)

No Tier, since it did not meet the requirements for Tier 1

This dataset has reached FAIR No Tier, since it did not meet the requirements for Tier 1 according to the EUCAIM classification system.

Tier	Status
Tier 1	✗ Failed Tier 1 not passed. Failing indicators: RDA_F1_01M, RDA_F1_01D
Tier 2	✗ Failed
Tier 3	✗ Failed





Findable

42.86%

RDA-F1-01D - Data are identified by a persistent identifier +

0%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle: F1 — (meta)data are assigned a globally unique and eternally persistent identifier. This indicator evaluates whether the data are identified by a persistent identifier. A persistent identifier guarantees that the data can be found over time and reduces the risk of broken links.
Technical Implementation	Searches through a predefined list of possible metadata terms (dcterms:identifier) to determine whether any such information is available in the metadata.
Technical feedback	Identifier is not persistent for the data: aaaadcw3tn2ydbsux2urluqaae
Tips	If a DOI has not yet been generated... If the resource had a DOI/other PID before deposit..., include it in the metadata record to indicate the external site where the data files are located.

RDA-F1-01M - Metadata are identified by a persistent identifier +

0%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle: F1 — (meta)data are assigned a globally unique and eternally persistent identifier. This indicator evaluates whether metadata are identified by a persistent identifier. A persistent identifier ensures that the metadata can be found over time and reduces the risk of broken links.
Technical Implementation	Searches through a predefined list of possible metadata terms (dcterms:identifier) to determine whether any such information is available in the metadata.
Technical feedback	Identifier is not persistent for the metadata: aaaadcw3tn2ydbsux2urluqaae
Tips	Metadata needs a unique identifier.

RDA-F1-02D - Data are identified by a globally persistent identifier +

0%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle: F1 — (meta)data are assigned a globally unique and eternally persistent identifier. The indicator evaluates whether the data's identifier is globally unique, that is, that no two datasets use the same identifier.
Technical Implementation	Checks that the data identifier is globally unique and eternally persistent (UUID or accepted by idutils).

Technical feedback	Identifier found for the data is not globally unique: aaaadcw3tn2ydbsux2urluqaae
Tips	Check whether the external site hosting the data files has generated a globally persistent identifier (handle, DOI, ARK, PURL...), and if so include it in the metadata.

RDA-F1-02M - Metadata are identified by a globally persistent identifier. +
0%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle: F1 — (meta)data are assigned a globally unique and eternally persistent identifier. The indicator evaluates whether the metadata's identifier is globally unique, i.e., that no two identical identifiers identify different metadata records.
Technical Implementation	Checks the type of identifiers using the idutils package. Identifies valid identifiers. If it is not a persistent identifier (e.g. a UUID), the test fails.
Technical feedback	Identifier found for the metadata is not globally unique: aaaadcw3tn2ydbsux2urluqaae
Tips	Metadata id must be globally unique and persistent (UUID).

RDA-F2-01M - There is sufficiently rich metadata to enable resource discovery +
100%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle F2: Data are described with rich metadata. The indicator concerns not only the presence of metadata but also the amount provided and how such metadata supports resource discovery.
Technical Implementation	Compares the metadata elements present with a configured list considered essential for discoverability. Calculates the proportion of elements found relative to those expected to assign a score.
Technical feedback	Findability terms present (15/15): • http://purl.org/dc/terms/title • http://purl.org/dc/terms/description • http://purl.org/dc/terms/identifier • https://www.eucaim.org/hasCondition • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/topography • https://www.eucaim.org/hasImageBodyPart • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/geographical_coverage • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/collection_method • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/number_of_subjects • https://www.eucaim.org/hasAssociatedSex • https://www.eucaim.org/hasImageModality • https://www.eucaim.org/hasImageVendor • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_type • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/provider • http://www.w3.org/ns/adms#interoperabilityLevel Findability terms missing (0/15): • —

RDA-F3-01M - Metadata include the identifier of the data +
100%

Indicator Level	Essential

Indicator Assessment	This indicator is linked to the following principle F3: Metadata clearly and explicitly include the identifier of the data they describe. The indicator concerns the inclusion of the reference (i.e., the identifier) of the digital object in the metadata so that the digital object can be accessed.
Technical Implementation	Verifies the metadata element where the object is identified dcterms:identifier.
Technical feedback	Metadata includes identifier/s for the data: aaaadcw3tn2ydbsux2urluqaae

RDA-F4-01M - Metadata are provided in such a way that they can be harvested and indexed

100%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle: F4 — (meta)data are registered or indexed in a searchable resource. The indicator tests whether the metadata are offered in a way that allows indexing. In some cases, metadata may be provided alongside the data to a local institutional repository or to a regional or domain-specific portal, or be included on a landing page where a search engine can harvest them. The indicator remains purposely broad so as not to limit how and by whom the data may be collected and indexed.
Technical Implementation	Checks whether the metadata can be harvested.
Technical feedback	Metadata is gathered programmatically through HTTP (API REST), thus can be harvested and indexed.

Accessible

57.56%

RDA-A1-01M - Metadata contain information on how the user can access the data

25%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator refers to the information necessary to allow the requester to access the digital object. This includes (i) whether there are restrictions to accessing the data (i.e., whether access to the data is open, restricted, or closed), (ii) the actions an interested person must take to access the data, particularly if the data are not published on the Web, and (iii) specify that the resources are available via eduGAIN or via specialized solutions as proposed for EPOS.
Technical Implementation	Checks whether the metadata include the necessary information to enable data access. It verifies the presence of key access-related terms such as Image_access_type, Image_access_description, Image_access_url and publication_url. It also ensures that certain elements, likes URIs, are resolvable.
Technical feedback	1/4 access elements satisfied. Missing: Image_access_description, Image_access_url, publication_url.

Tips	Include the terms: image_access_type, image_access_description, image_access_url and publication_url making sure the URIs are resolvable.

RDA-A1-02D - Data are accessible manually +

100%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator refers to any human interaction required if the requester wishes to access the digital object. FAIR primarily addresses automated interactions where a machine can access the object, but there may also be digital objects that require human interactions, such as clicking a link on a landing page, sending an email to the data owner, or even calling by phone. The indicator can be evaluated by searching metadata for information describing how the digital object can be obtained through human intervention.
Technical Implementation	Checks if there is an accessible reference (e.g., a URL) allowing manual retrieval of the data. If the link is defined and functional, 100 points are awarded.
Technical feedback	Documentation for the manual obtention of the data can be found in https://catalogue.eucaim.cancerimage.eu/

RDA-A1-02M - Metadata are accessible manually +

100%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator refers to any human interaction necessary if the requester wishes to access the metadata. FAIR primarily addresses automated interactions where a machine can access the metadata, but metadata may also require human interactions. This may be important where the metadata contain confidential information. Human interaction may involve sending an email to the metadata owner or calling by phone to receive instructions.
Technical Implementation	Checks if there is an accessible reference (e.g., a URL) allowing manual retrieval of the metadata.
Technical feedback	Documentation for the manual obtention of the metadata can be found in https://catalogue.eucaim.cancerimage.eu/

RDA-A1-03D - The data identifier resolves to a digital object +

0%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. This indicator concerns the resolution of the identifier that identifies the digital object. The identifier assigned to the data must be associated with a formally defined retrieval/resolution mechanism that allows access to the digital object, or provides access instructions in cases of human-mediated access. The FAIR principle and this indicator say nothing about the mutability or immutability of the digital object identified; this is governed by the data provider's persistence policy.

Technical Implementation	Searches for a publication identifier (such as a DOI) and checks whether it resolves correctly via an HTTP request. If the link is resolvable, 100 points are assigned.
Technical feedback	No DOI found for the data
Tips	Ensure the dataset has a DOI registered and included in the metadata as http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/publication_url. Verify that the DOI is resolvable through https://doi.org/.... Broken or inactive links will yield zero points. If a DOI is not possible, consider alternative persistent resolvable identifiers (e.g., Handle, ARRK), but ensure they follow a standardized resolution mechanism.

RDA-A1-03M - The metadata identifier resolves to a metadata record +
100%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. This indicator concerns the resolution of the metadata identifier. The identifier assigned to the metadata must be associated with a resolution service that allows access to the metadata record.
Technical Implementation	Verifies whether metadata could be extracted from its identifier.
Technical feedback	Metadata record could be retrieved from metadata identifier: aaaadcw3tn2ydbsux2urluqaae

RDA-A1-04D - Data are accessible via a standard protocol +
0%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator concerns the protocol used to access the digital object and requires that the protocol be defined in a standard.
Technical Implementation	Searches metadata for an element corresponding to data access. Extracts the protocol (such as http, https) from the provided link and compares it with a list of standardized protocols. If at least one valid protocol is found, full score is awarded.
Technical feedback	Could not check data access protocol: metadata element http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_url
Tips	rda_a1_04d.tips

RDA-A1-04M - Metadata are accessible via a standard protocol +
100%

Indicator Level	Essential

Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. More information about this principle can be found here. The indicator concerns the protocol used to access the metadata and requires that the protocol be defined in a standard.
Technical implementation	Identifies the protocol scheme (e.g., http, https) used to access the metadata, and compares it with a list of protocols accepted by standards.
Technical feedback	Found a standarised protocol to access the metadata record: https

RDA-A1-05D - Data are accessible automatically +

0%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle A1: (meta)data are retrievable by their identifier using a standardized communication protocol. The indicator refers to automated interactions between machines to access digital objects. The indicator will evaluate how machines interact and grant access to the digital object.
Technical implementation	Evaluates whether the data are accessible automatically via a URL, checking whether it is resolvable and present in the metadata.
Technical feedback	Could not check data access protocol: metadata element http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_url not found or not accessible
Tips	rda_a1_05d.tips

RDA-A1.1-01D - Data are accessible via a free protocol +

0%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle A1.1: The protocol is open, free, and universally applicable. The indicator requires that the protocol can be used free of charge, facilitating unrestricted access.
Technical implementation	Reuses the result from indicator RDA-A1-04D to check whether data access is via a free and open protocol (such as HTTP or HTTPS). If at least one is detected, 100 points are assigned.
Technical feedback	Could not check data access protocol: metadata element http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/image_access_url
Tips	rda_a1_1_01d.tips

RDA-A1.1-01M - Metadata are accessible via an open, free protocol +

100%

Indicator Level	Essential

Indicator Assessment	This indicator is linked to the following principle A1.1: The protocol is open, free, and universally applicable. The indicator tests that the protocol allowing access to metadata can be used freely. Free usage of the protocol enhances data reuse.
Technical implementation	Same scoring as rda_a1_04m.
Technical feedback	Found a free protocol to access the metadata record: https

RDA-A1.2-01D - Data are accessible via a protocol that supports authentication and authorization +

0%

Indicator Level	Optional
Indicator Assessment	This indicator is linked to the following principle A1.2: The protocol allows authentication and authorization when needed. The indicator requires how one can authenticate and authorize access to the digital object, and that access to the data is properly described and documented.
Technical implementation	Checks whether metadata include information on authentication and authorization mechanisms required to access the data.
Technical feedback	At the time, does not provide authentication or authorisation protocols
Tips	rda_a1_2_01d.tips

RDA-A2-01M - Metadata remain accessible even if the data are no longer available +

100%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle A2: Metadata must remain accessible even when the data are no longer available. The indicator is intended to verify that information about a digital object is still available after the object has been removed or lost. Where possible, the metadata that remain available should also indicate why the object is no longer available.
Technical implementation	Metadata are always subject to the EUCAIM Metadata Preservation Policy.
Technical feedback	EUCAIM Metadata preservation policy

Interoperable

62.6%

RDA-I1-01D - Data use some standardized form of knowledge representation +

100%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. The indicator determines that an appropriate standard is used to express knowledge, particularly the data model and format.
Technical Implementation	Extracts data formats from the metadata ('availableFormats') and compares them with those officially registered as Internet Media Types by IANA. If at least one matches, 100 points are awarded.
Technical feedback	Data uses a correct way to present information present in https://www.iana.org/assignments/media-types/media-types.xhtml

RDA-I1-01M - Metadata use some standardized form of knowledge representation

80.2%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. The indicator determines that an appropriate standard is used to express knowledge, e.g., controlled vocabularies for subject classifications.
Technical Implementation	Validates the metadata RDF graph against three SHACL profiles (DCAT-AP, Health-DCAT-AP, and EUCAIM). Calculates how many primary active NodeShapes comply with their constraints, recursively evaluating child properties. The score is the average of the three profiles.
Technical feedback	dcat-ap: 94.44 pts - failing properties 2/36 [dct:provenance, dct:publisher] healthdcat-ap: 46.15 pts - failing properties 7/26 [dcat:distribution, dcatap:applicableLegislation, dct:accessRights, dct:publisher, dct:publisher, healthdcatap:hdab, healthdcatap:healthCategory] eucaim: 100.00 pts - failing properties 0/15 [-]
Tips	It is recommended to describe the digital object using the controlled vocabulary Health-DCAT-AP and EUCAIM-specific terms.

RDA-I1-02D - Data use a machine-processable form of knowledge representation

100%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. This indicator focuses on machine-understandability of the data. It means the data must be legible—and thus interoperable—by machines without human-mediated translation or mappings.
Technical Implementation	Checks if the data format is a machine-processable form of knowledge representation.
Technical feedback	1/1 machine-readable data model values found.

RDA-I1-02M - Metadata use a machine-processable form of knowledge representation +

100%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. This indicator focuses on the machine-understandability of metadata. It means metadata must be legible—and thus interoperable—by machines without requirements like translators or specific mappings.
Technical Implementation	Verifies that metadata are represented as a valid, non-empty RDF graph. If the RDF graph contains at least one triple, the metadata are considered machine-understandable.
Technical feedback	Metadata is valid RDF with 87 triples.

RDA-I2-01D - Data use FAIR vocabularies +

100%

Indicator Level	Optional
Indicator Assessment	This indicator is linked to the following principle I2: (meta)data use vocabularies that follow FAIR principles. The indicator requires that the controlled vocabulary used for the data adheres to FAIR principles and is at least documented and resolvable using...
Technical Implementation	Checks the configured vocabulary predicates and verifies that their objects are URIs or typed literals (not literals) and that these URIs are dereferenceable with acceptable content types.
Technical feedback	Data vocabulary objects checked: 9 • Resolvable IRIs: 9 • Literals (0): • — • No resolvable IRIs (0): • —

RDA-I2-01M - Metadata use FAIR vocabularies +

100%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle I2: (meta)data use vocabularies that follow FAIR principles. The indicator requires that the vocabulary used in the metadata adheres to FAIR principles and is at least documented and resolvable via globally persistent and unique identifiers. The documentation must be easy to access and find.
Technical Implementation	Extracts HTTP(S) namespace URIs used in the RDF graph and checks if they are resolvable via HTTP requests. A score is assigned proportional to the percentage of vocabularies accessible online.
Technical feedback	Namespaces resolving (8/8): • http://purl.org/dc/terms/ (200 text/turtle) • http://purl.org/linked-data/cube# (200 text/plain) • http://www.w3.org/2000/01/rdf-schema# (200 text/turtle) • http://www.w3.org/2001/XMLSchema# (200 text/html) • http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/ (200 text/turtle) • http://www.w3.org/ns/dcat# (200 text/turtle) • http://www.w3.org/1999/02/22-rdf-syntax-ns# (200 text/turtle) • https://w3id.org/fdp/fdp-o# (200 text/plain) Namespaces not resolving (0/8): • —

RDA-I3-01D - Data include references to other data +

0%

Indicator Level	Optional
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. This indicator concerns how the data connect to other data—for example, linking to prior or related research data that provide additional context to the data.
Technical Implementation	This test implements the logic used in RDA-I3-02M.
Technical feedback	No references to other data.
Tips	rda_i3_01d.tips

RDA-I3-01M - Metadata include references to other metadata +

50%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. The indicator concerns how metadata connect to other metadata—for example, via links to information about organizations, people, places, projects, or time periods related to the digital object described by the metadata.
Technical Implementation	Checks whether terms like relation or contributor are present in the RDF and contain qualified references, such as resolvable URIs, blank nodes, or dictionary-type structures. The score is based on the percentage of terms meeting these conditions.
Technical feedback	Relation terms with qualified links: 2/4 Terms present but ""not"" linked: • http://purl.org/dc/terms/publisher • http://purl.org/dc/terms/provenance
Tips	Ensure the RDF includes qualified references in the following fields: 'http://purl.org/dc/terms/publisher', 'http://www.w3.org/ns/dcat#contactPoint', 'http://purl.org/dc/terms/spatial', 'http://purl.org/dc/terms/temporal', 'http://purl.org/dc/terms/provenance', 'http://purl.org/dc/terms/relation', 'http://purl.org/dc/terms/isPartOf', 'http://www.w3.org/ns/adms#sample', 'http://www.w3.org/ns/dcat#accessURL', 'http://www.w3.org/ns/dqv#hasQualityAnnotation'. These references must be expressed as valid URIs, blank nodes, or structured dictionaries to count as qualified.

RDA-I3-02D - Data include qualified references to other data +

0%

Indicator Level	Optional
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. This indicator concerns how data are connected to other data. References must be qualified, meaning that the role of the related resource is specified.
Technical Implementation	rda_i3_02d.technical
Technical feedback	This test implies checking the presence of qualified references within the content of the data. As it is defined, its implementation is too costly.

Tips	rda_I3_02d.tips

RDA-I3-02M - Metadata include references to other data
0%

Indicator Level	Optional
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. This indicator concerns how metadata are connected to other data—for example, linking to prior or related research data that provide additional context to the data. Note that this is not the link from metadata to the data being described; that link is covered by principle F3 and indicator RDA-F3-01M.
Technical Implementation	rda_I3_02m.technical
Technical feedback	No references to other data.
Tips	rda_I3_02m.tips

RDA-I3-03M - Metadata include qualified references to other metadata
50%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. This indicator concerns how metadata connect with other metadata—for example, with descriptions of related resources that provide additional context to the data.
Technical Implementation	This indicator reuses the logic of rda_I3_01m.
Technical feedback	Relation terms with qualified links: 2/4 Terms present but ""not"" linked: • http://purl.org/dc/terms/publisher • http://purl.org/dc/terms/provenance
Tips	Ensure the RDF includes qualified links in the following fields: 'http://purl.org/dc/terms/publisher', 'http://www.w3.org/ns/dcat#contactPoint', 'http://purl.org/dc/terms/spatial', 'http://purl.org/dc/terms/temporal', 'http://purl.org/dc/terms/provenance', 'http://purl.org/dc/terms/relation', 'http://purl.org/dc/terms/isPartOf', 'http://www.w3.org/ns/adms#sample', 'http://www.w3.org/ns/dcat#accessURL', 'http://www.w3.org/ns/dqv#hasQualityAnnotation'. These references should clearly indicate the relationship type (e.g., creator, editor, or geographic coverage) and be expressed as valid URIs, blank nodes, or structured dictionaries.

RDA-I3-04M - Metadata include qualified references to other data
0%

Indicator Level	Optional
Indicator Assessment	This indicator is linked to the following principle I3: (Meta)data include qualified references to other (meta)data. This indicator concerns how metadata connect to other data. References must be qualified, meaning that the relationship role of the related resource is specified, for example, dataset X is derived from dataset Y.

Technical implementation	This indicator reuses the logic of rda_i3_01d.
Technical feedback	No references to other data.
Tips	rda_i3_04m.tips

Reusable

71.41%

RDA-R1-01M - Certain attributes relevant for reuse are included +

90%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle R1: (Meta)data are richly described with a plurality of accurate and relevant attributes. The indicator refers to both the quantity and quality of metadata provided to enhance data reuse.
Technical implementation	Compares the RDF metadata with a configured list of terms considered relevant for reuse. The score is proportional to the number of elements found in the RDF relative to the expected total.
Technical feedback	Found 19/21 metadata elements that enhance reusability. Missing: ['http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Collections/column/quality_annotation', 'https://www.w3.org/TR/vocab-adms/#adms-sample']
Tips	Ensure the RDF includes the following reuse-enhancing fields: 'http://purl.org/dc/terms/spatial', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/image_year_range', 'http://purl.org/dc/terms/type', 'http://www.w3.org/ns/dcat#theme', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/sample', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/quality_annotation', 'http://www.healthdcatap.org/minTypicalAge', 'http://www.healthdcatap.org/maxTypicalAge', 'http://www.healthdcatap.org/numberofRecords', 'http://www.healthdcatap.org/numberofUniqueIndividuals', 'https://www.eucaim.org/hasAssociatedSex', 'https://www.eucaim.org/hasCondition', 'https://www.eucaim.org/hasImageModality', 'https://www.eucaim.org/hasImageBodyPart', 'https://www.eucaim.org/hasImageVendor', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/collection_method', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/access_rights', 'https://w3id.org/dpv/dpv-skos#hasPurpose', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/version', 'http://catalogue-eucaim.grycap.i3m.upv.es/Eucaim/api/rdf/Collections/column/interoperability_level'. These fields must be present with suitable values to consider the resource reusable.

RDA-R1.1-01M - Metadata include information about the license under which the data are published for reuse +

100%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle R1.1: (Meta)data are released with a clear and accessible data usage license. This indicator concerns the information provided in the metadata about the conditions (e.g., obligations, restrictions) under which the data can be reused.

Technical implementation	Checks if the metadata include at least one value under the term configured as in dcterms:rights. If one or more licenses are found, 100 points are awarded. I no rights/license is present, it inspects image_access_type as contextual information (restricted/-non-public access does not replace a license)
Technical feedback	License/rights detected in dcterms:rights. Non-standard but resolvable: http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Rights?name=IN_SITU (internal policy detected; consider mapping to a community license). Access context: image_access_type=http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Image_access_type?name=NON_PUBLIC, http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Image_access_type?name=RESTRICTED_ACCESS.

RDA-R1.1-02M - Metadata include references to a standard license under which the data are published for reuse

0%

Indicator Level	Recommendable
Indicator Assessment	This indicator requires that the reference to reuse conditions be a standard license, rather than a locally defined one. The indicator can be evaluated by checking that the license is indeed a standard license. Examples include Creative Commons and Open Data Commons licenses.
Technical implementation	Checks the results of rda_r1_1_01m to identify any standard license, whether provided as a link or literal value.
Technical feedback	Community-standard requirement not satisfied. Consider using a standard license.
Tips	Include the resource standard license in the dcterms:rights metadata field if available.

RDA-R1.1-03M - Metadata include machine-readable references to a standard license under which the data are published for reuse

100%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle R1.1: (Meta)data are released with a clear and accessible data usage license. This indicator concerns how the reuse license is expressed. Instead of a human-readable text, the license should be expressed in a machine-processable way, e.g., for automated searches.
Technical implementation	Checks the results of rda_r1_1_01m to identify if the license is resolvable (must be a link).
Technical feedback	Machine-understandable requirement satisfied. Non-standard but resolvable: http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Rights?name=IN_SITU Non-standard but resolvable: http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Image_access_type?name=NON_PUBLIC, Non-standard but resolvable: http://catalogue.eucaim.cancerimage.eu/Eucaim/api/rdf/Image_access_type?name=RESTRICTED_ACCESS.

RDA-R1.2-01M - Metadata include provenance information according to community standards

100%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle R1.2: (Meta)data are associated with detailed provenance. It requires that metadata include information about the data's origin, history, or workflow in a way that conforms to standards used in the producing community.
Technical Implementation	Checks whether the RDF contains at least one entry under the configured provenance term. If provenance information is found, 100 points are awarded.
Technical feedback	Found provenance information in the metadata.

RDA-R1.2-02M - Metadata include provenance using a community-aligned, cross-domain format +

0%

Indicator Level	Optional
Indicator Assessment	This indicator is linked to the following principle R1.2: (Meta)data are associated with detailed provenance. It requires that provenance information in metadata follow a cross-domain machine-understandable format.
Technical Implementation	Checks whether the RDF includes provenance information using a machine-understandable value (IRI or typed literal). Evaluates the value of the configured provenance field; if it's a URI or typed literal, 100 points are awarded.
Technical feedback	Provenance is present but its value is not machine-understandable (must be an IRI or a typed literal).
Tips	Ensure that 'http://purl.org/dc/terms/provenance' contains either a URI (e.g., DOI or URL) or a typed literal (e.g., date with 'xsd:date'). This allows automatic processing by systems.

RDA-R1.3-01D - Data comply with community standards +

100%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires the data to comply with relevant community standards.
Technical Implementation	Checks whether the field configured as 'terms_community_standard' contains any recognized community formats like DICOM, DICOM-SEG, or NIfTI. If any are found, 100 points are awarded.
Technical feedback	Dataset format complies with community standards (DICOM, DICOM-SEG, nifti).

RDA-R1.3-01M - Metadata comply with community standards +

46%

Indicator Level	Essential

Indicator Assessment	This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires metadata to comply with standards established in their domain.
Technical Implementation	Checks whether metadata conform to the 'healthdcat-ap' profile, evaluated via the result from rda_i1_01m.
Technical feedback	Community-standard requirement not satisfied using Health-DCAT-AP < 50 points.
Tips	Ensure the RDF compiles with the HealthDCAT-AP profile.

RDA-R1.3-02D - Data are machine-readable and comply with community standards

100%

Indicator Level	Recommendable
Indicator Assessment	This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires that data comply with a community standard in a machin
Technical Implementation	This indicator evaluates whether the data format is considered a machine-readable community standard, such as DICOM, DICOM-SEG, or NIfTI.
Technical feedback	Dataset format compiles with machine-understandable community standard (DICOM, DICOM-SEG, nifti).

RDA-R1.3-02M - Metadata are machine-readable and comply with community standards

46%

Indicator Level	Essential
Indicator Assessment	This indicator is linked to the following principle R1.3: (Meta)data meet domain-relevant community standards. It requires that metadata comply with a community standard in a machine-understandable format.
Technical Implementation	Evaluates whether metadata comply with a machine-understandable community standard using the score obtained for the HealthDCAT-AP profile in rda_i1_01m.
Technical feedback	Metadata is not expressed in any machine-understandable community standard. Please check the Health-DCAT-AP. Health-DCAT-AP < 50 points.
Tips	Ensure the RDF is structured according to the HealthDCAT-AP profile.