

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

WP4 Share

D4.2 Public Listing (Catalog) of FAIRiCUBE data
resources

Deliverable Lead: EOX

Deliverable due date: 30.06.2024

Version: 2.4

27-06-2024

Document Control Page

Document Control Page	
Title	D4.2 Public Listing (Catalog) of FAIRiCUBE data resources
Creator	EOX
Description	The Public Listing (Catalog) of FAIRiCUBE data resources details the processes and procedures as well as the software and services used for managing metadata of data resources within FAIRiCUBE
Publisher	"FAIRiCUBE – F.A.I.R. information cubes" Consortium
Contributors	EOX, all partners
Date of delivery	14-06-2023
Type	List
Language	EN-GB
Rights	Copyright "FAIRiCUBE – F.A.I.R. information cubes"
Audience	<input checked="" type="checkbox"/> Public <input type="checkbox"/> Confidential <input type="checkbox"/> Classified
Status	<input type="checkbox"/> In Progress <input type="checkbox"/> For Review <input checked="" type="checkbox"/> For Approval <input type="checkbox"/> Approved

Revision History			
Version	Date	Modified by	Comments
0.1	05-02-2023	Stephan Meißl, EOX	Initial Document
1.0	26-02-2023	Kathi Schleidt, EPS Cristina Carnerero, 4sfera Stephan Meißl, EOX	First release
1.1	28-02-2023	Jaume Targa, 4sfera	Revision
1.2	04-03-2023	Stephan Meißl, EOX	Release
1.3	05-03-2023	Stephan Meißl, EOX	Harmonize wording, final version for submission
1.4	14-06-2023	Christian Schiller, EOX	reworking of Tabel 1
1.5	17-11-2023	Christian Schiller, Mussab Abdalla, Stephan Meißl, EOX	reworking of chapter due to the big changes in the metadata requirements reworking of Tabel 1, according to new Inventory.xlsx updating/adding Figures 2-4 showing Data Ingestion Request WebGUI
1.6	29-11-2023	Christian Schiller	Updating Chapter 2 and 4
2.0	19-12-2023	Mussab Abdalla	Final Check
2.1	20-12-2023	Jaume Targa/Lorena Banyuls (4sfera)	Final Review
2.2	2024-05-31	Mussab Abdalla, Christian Schiller	Update STAC item description Update Table 1
2.3	2024-06-21	Jaume Targa	Final Review and forma checking
2.4	2024-06-27	Mussab Abdalla, Christian Schiller	Reworking of Table 2: extended List of available resources



D4.2 Public Listing (Catalog) of FAIRICUBE data resources

Disclaimer

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

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



Table of Contents

1 Introduction	6
2 Metadata Requirements for Data Resources	7
3 FAIRiCUBE Metadata for Data Resources	12
3.1 Metadata mapping	16
4 FAIRiCUBE Catalog for Data Resources	27
5 Data Resources	29
5.1 Resources made available in FAIRiCUBE	30
6 References	33

List of Figures

Figure 1: Data Ingestion Request Procedure	8
Figure 2: Data Ingestion Request Web GUI – Landing page	9
Figure 3: Data Ingestion Request Web GUI - Data entry Part-1	10
Figure 4: Data Ingestion Request Web GUI - Data entry Part-2	10
Figure 5: Dynamic Catalog based on STAC -fastapi.....	27
Table 6: Dynamic Catalog - Search Interface	27
Table 7: Dynamic Catalog – Data item description	28

List of Tables

Table 1: Mapping from metadata requirements to STAC	15
Table 2: Data requests by Use Cases as per Inventory Sheet (as of December 2023)	28



1 Introduction

The FAIRiCUBE Catalog, integrated into the FAIRiCUBE Hub, provides human access to the resource metadata, enabling potential users to identify and access available resources. Metadata required for the description of data resources are described within this deliverable D4.2 whereas metadata for processing/analysis resource are described in the deliverable D4.3. In addition, the information from the FAIRiCUBE Catalog on data resources is also exposed via STAC (SpatioTemporal Asset Catalog) API (Application Programming Interface). It is planned to deploy software harvesting the static STAC metadata and expose it via STAC API later in the project. To ensure that the diverse processing-analysis resources created within the FAIRiCUBE Hub are truly FAIR, these principles must be fully applied. Pertaining to these resources, the application of the FAIR principles is:

- **Findable:** the necessary metadata must be provided for these resources, to enable potential users to easily find them as well as assess their suitability for the task at hand.
- **Accessible:** the resources must be available to potential users under clear conditions. If licensing conditions apply, these must be provided in a transparent manner
- **Interoperable:** it must be possible to apply the available processing-analysis resources to different spatiotemporal settings and source data
- **Reusable:** it must be possible to execute the available processing-analysis resources in different settings, whereby tailoring of these resources to specific requirements should be possible.

Of primary concern to this document is the findability aspect, what metadata is required to enable potential users to identify suitable data resources. The accessibility, interoperability, and reusability aspects are of secondary concern, as these are enabled by the FAIRiCUBE Hub. However, the requirements stemming from FAIRiCUBE Hub functionality must also be covered by the metadata foreseen for data resources.



2 Metadata Requirements for Data Resources

The requirements for the Metadata, describing the datasets to be used within FAIRiCUBE, have been collected within the consortium. To collect all desired field names, discuss, harmonize, streamline and explain them was a very time-consuming task. A spreadsheet formed the basis for the collection of the desired field names needed. This spreadsheet was then translated to the Table 1 in this document.

A need for additional metadata has been identified in use case discussions that should be covered by FAIRiCUBE Catalog. The need is to convey information about how to portray a data resource. This includes information like legends mapping values from the dataset to colours to be displayed. As this is tightly coupled with the datasets, such information should be provided on the dataset metadata level. The initial idea to manage all the information using GitHub issues had to be dropped during the course of the metadata definition phase, due to limitations of the number of fields (max 55 fields supported) supported by GitHub.

Therefore, EOX developed a Web-GUI (Figure 2) as an Input Frontend allowing to collect and edit the metadata for the data resources. The input from this Web-GUI is collected, checked for consistency and errors and then directly stored as static STAC json items in GitHub. This ensures that all items stored in the GitHub repository act as the single "*Source of Truth*". The same interface is also available to edit already ingested metadata items.

Once the respective item is stored in GitHub a review process is initiated where the metadata responsible will manually check the correctness of the input and, if verified, will apply the Label '*approved*'. Once the item is labelled '*approved*' an automatic deployment procedure is available via GitHub pages and the content is directly harvested (STAC-fastapi/pgSTAC to provide a STAC API) and available via the Catalog Client, based on STAC Browser, currently deployed at <https://catalog.eoxhub.fairicube.eu/?language=en>. This catalog will be available in the future at deployed at <https://catalog.fairicube.eu>.

In case new data has to be added, a Data Request WebGUI has been created. The respective ingestion procedure is described in depth in D4.1 [16] and D5.2 [17]. Any new data request is addressed by the requester together with one of the ingestion handling partners. Any progress, problems, discussions, etc. shall be documented in a GitHub issue associated to the respective Pull Request, so that everybody interested can follow the progress and provide additional feedback or information as necessary. The following procedure for a data request has been set up and is shown in Figure 1.

D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

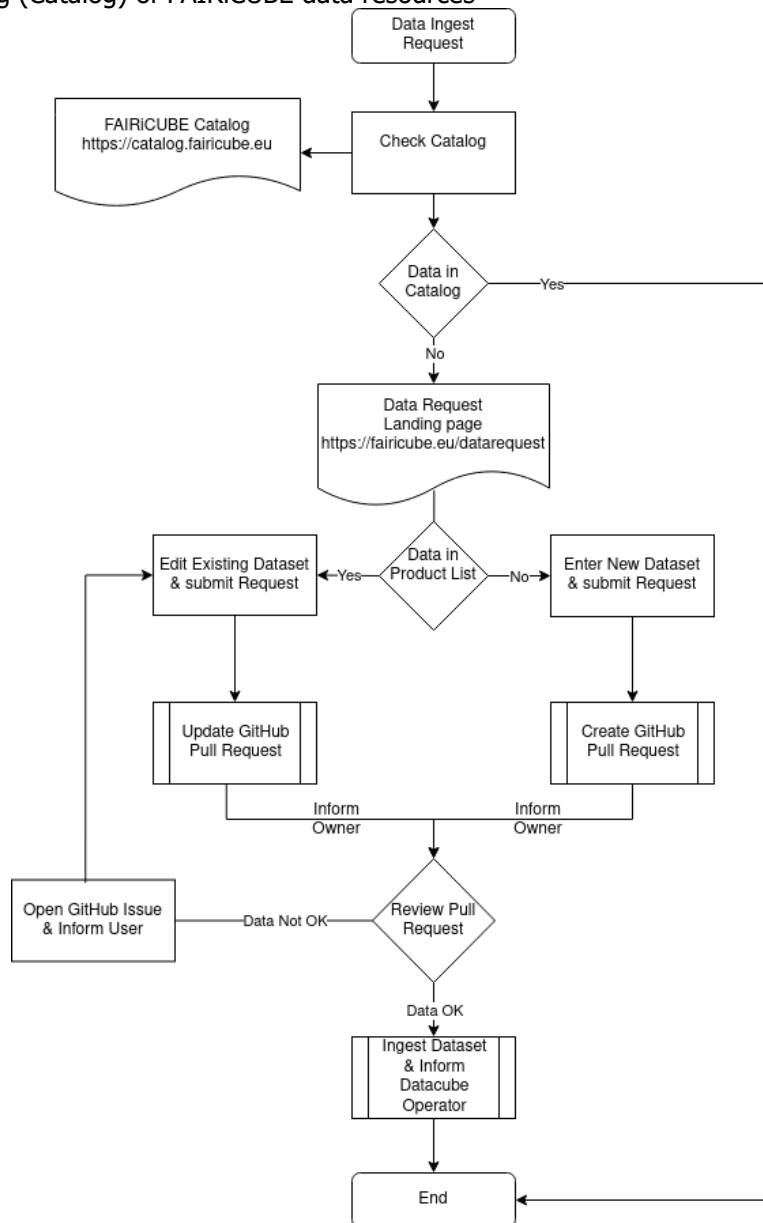


Figure 1: Data Ingestion Request Procedure

D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

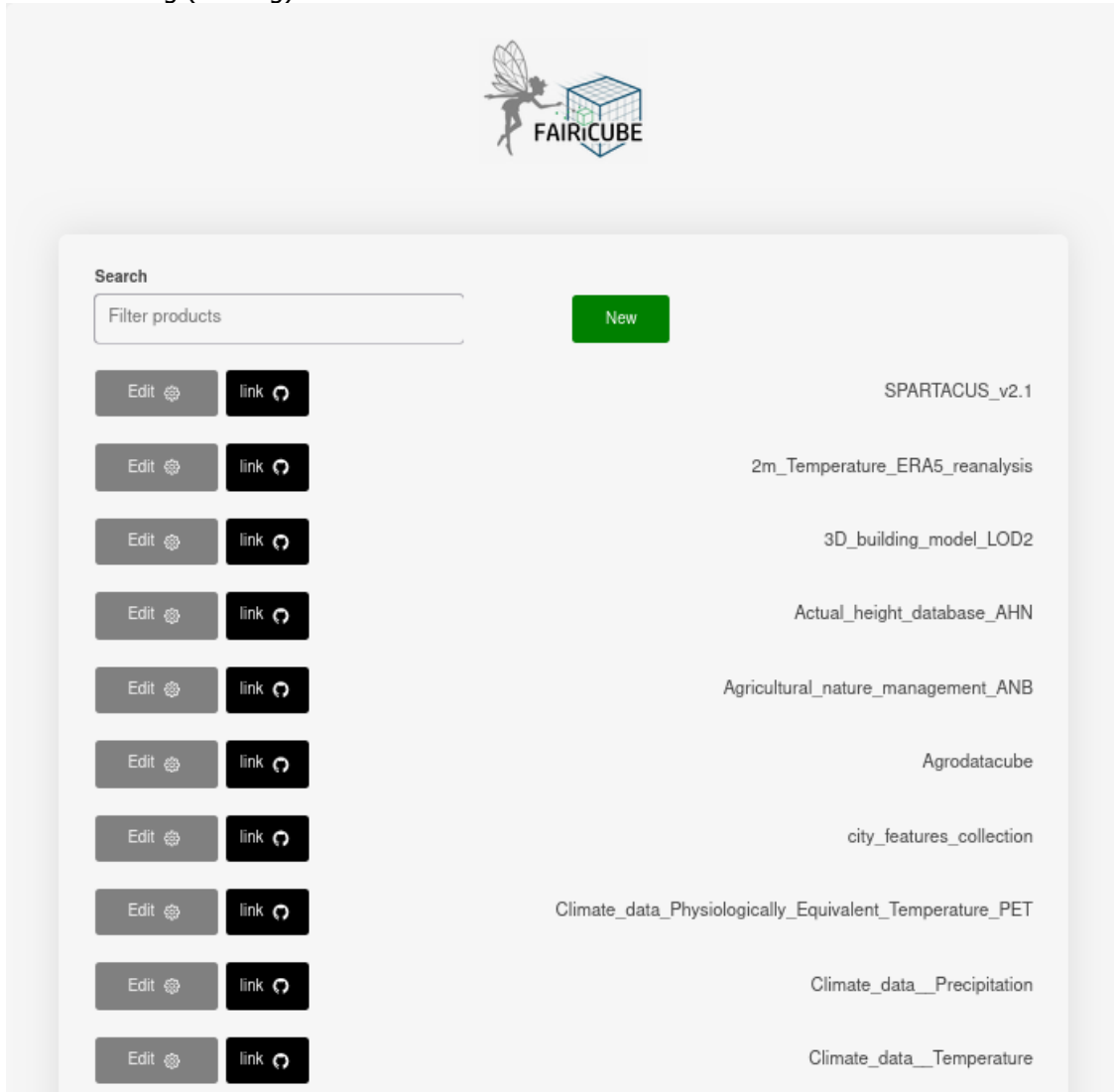


Figure 2: Data Ingestion Request Web GUI – Landing page

Once all metadata and data requirements are fulfilled, confirmed, and validated by the data requester, the ingestion handling partners will perform the merge and the request will be closed. Any discussion remains available. Figure 2 shows the landing page of the WebGUI which can also be used to create new or update existing datasets. In Figure 2 a sample list of available datasets is shown, each with an Edit Button associated. Additionally, a *Link Button* is provided which directly links to the pull request in the respective GitHub repository, which also provides access to the *.json files and allows reviewing any changes.

When a user enters a name in the *Search field* and presses the *Add button* on the Landing page, then the data entry form will be displayed (Figure 3, Figure 4), allowing the user to create a new record. If a user chooses an already existing dataset and uses the *Edit button* the same entry form will be shown with the values available already filled in. When the merge is done the newly submitted data is available in the STAC Browser deployed at <https://catalog.fairicube.eu>.

D4.2 Public Listing (Catalog) of FAIRiCUBE data resources



[back](#)

Title

The title of the issue request

ID

The ID of the requested stac item

Description

Brief, nontechnical explanation of the datacube.

Source

The link to the source of the dataset.

Source Type

The data source type

Organizations

[+ Add bands](#)

Figure 3: Data Ingestion Request Web GUI - Data entry Part-1

Resolution

Resolution (or 'irregular'). Should be 1 value as required by UC, not all resolutions of dataset

[Remove](#)
remove the extra dimension(Axis)

[+ Add another](#)
Add another dimension(Axis)

Bands

cell components <input type="text"/>	Unit of Measure <input type="text"/>	Data Type Select a data type ▾	Null values <input type="text"/>
Definition <input type="text"/>	Description <input type="text"/>	Category List <input type="text"/>	Comment <input type="text"/>

[Remove](#)
remove band

[+ Add bands](#)
Add another band

Re-projection axis

Horizontal CRS

Unit of Measure

Figure 4: Data Ingestion Request Web GUI - Data entry Part-2



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Once all metadata and data requirements are fulfilled, confirmed and validated by the data requester, the ingestion handling partners will perform the merge and the pull request will be closed. The respective branch in GitHub will also be closed and deleted. Any issues and discussions associated with the pull request are still available after the branch has been merged and deleted. When the merger is done the newly submitted data is available as a STAC item to the STAC Browser, currently deployed at <https://catalog.eoxhub.fairicube.eu/?language=en>. This catalog will be available in the future at deployed at <https://catalog.fairicube.eu>.



3 FAIRiCUBE Metadata for Data Resources

In this section, the metadata identified for the description of data resources and their mapping to the STAC standard are described. In addition, new STAC fields are proposed which are desired by the FAIRiCUBE team members but are currently not covered by the STAC standard or by a STAC extension.

The FAIRiCUBE consortium decided to rely on the STAC (SpatioTemporal Asset Catalog) specification¹. The four key components of STAC include items, catalogs, collections, and the STAC API²:

- **STAC Item:** A STAC item is the foundational building block of STAC. It is GeoJSON supplemented with additional metadata that enables clients to traverse through catalogs.
- **STAC Catalog:** A Catalog is usually the starting point for navigating a STAC. A catalog.json file will contain contains links to some combination of other catalogs, collections, and/or items. This combination is quite variable and flexible depending on how the data is being organized.
- **STAC Collection:** A STAC Collection builds upon the STAC Catalog specification to include additional metadata about a set of items that exist as part of the collection.
- **STAC API:** STAC Catalogs can be static, by creating the json files and storing them. This makes static STAC Catalogs highly portable, reliable, providing a solid foundation for building dynamic versions through the use of APIs. A STAC API is a RESTful API specification for querying STAC Catalogs in a dynamic way. It is designed with a standard set of endpoints for searching Catalogs, collections, and items.

The entry point to the FAIRiCUBE Catalog is a STAC Catalog³. It lists all the available STAC collections, where each STAC Collection includes data as individual STAC Items linking to the describing json files. An example snippet from the entry STAC Catalog representing the FAIRiCUBE Catalog for the CORINE Land Cover looks like this:

```
{
  "rel": "item",
  "type": "application/json",
  "href": " https://fairicube.github.io/data-requests/CORINE_LAND_COVER/CORINE_LAND_COVER.json",
}
```

The full item description for the CORINE Land Cover utilizing the STAC Datacube extension⁴ looks like below. This item describes the CORINE Land Cover which is a typical land cover dataset consisting of 44 land cover and land use classes derived from a series of satellite missions since it was first established in 1990 currently up to 2018.

```
{
  "type": "Feature",
  "stac_version": "1.0.0",
  "id": "corine_land_cover",
  "properties": {
    "license": "other",
    "description": "It consists of an inventory of land cover in 44 cases.",
    "providers": [
      {
        "organization_name": "European Environment Agency",
        "organization": "EEA",

```

1 <https://stacspec.org>

2 <https://stacspec.org/en/tutorials/intro-to-stac/>

3 <https://catalog.fairicube.eu/stac/index.json>

4 <https://github.com/stac-extensions/datacube>



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

"comments": "The present 100m raster dataset is the 2018 CLC status layer modified for the purpose of consistent statistical analysis in the land cover change accounting system at EEA.\n\nCORINE Land Cover (CLC) data are produced from 1986 for European (EEA member or cooperating) countries. Altogether five mapping inventories were implemented in this period, producing five status layers (CLC1990, CLC2000, CLC2006, CLC2012, CLC2018) and four CLC-Change (CLCC) layers for the corresponding periods (1990-2000, 2000-2006, 2006-2012, 2012-2018). Pan-European CLC and CLCC data are available as vector and raster products.",

"doc_link": "https://land.copernicus.eu/pan-european/corine-land-cover",

"organization_email": "copernicus@eea.europa.eu",

"project_purpose": "CORINE Land Cover is a pan-European land cover inventory with 44 classes. Initiated in 1985 (the 1990 reference year) the inventory is available for the 1990, 2000, 2006, 2012 and 2018 reference years including change layers 1990-2000, 2006-2012 and 2012-2018. CORINE Land Cover is part of the European Union's Copernicus Land Monitoring Service."

```
    },
  ],
  "dataSource": "https://land.copernicus.eu/pan-european/corine-land-cover",
  "cube:dimensions": {
    "x": {
      "axis": "x",
      "extent": [
        -31.561261,
        44.820775
      ],
      "reference_system": "EPSG:4326",
      "type": "spatial",
      "unit": "m",
      "interpolation": "nearest",
      "step": 100
    },
    "y": {
      "axis": "y",
      "extent": [
        27.405827,
        71.409109
      ],
      "reference_system": "EPSG:4326",
      "type": "spatial",
      "unit": "m",
      "interpolation": "nearest",
      "step": 100
    },
    "time": {
      "extent": [
        "1990-01-01T00:00:00Z",
        "2018-01-01T00:00:00Z"
      ],
      "type": "temporal",
      "step": ""
    },
    "z": {
      "extent": [
        null,
        null
      ],
      "reference_system": null,
      "unit": "m",
      "interpolation": "nearest",
      "type": "spatial"
    }
  },
  "datetime": "2000-01-01T00:00:00Z",
  "raster:bands": [
    {
      "band_name": "CLC",
      "unit": "categorical, numeric code",
      "data_type": "int8",
      "nodata": 0,
      "definition": "CLC code - each code represents a colour",
      "description": "Each colour is a landform feature (urban area, forests, etc)",
      "category_list": "Multiple values",
      "comment": null,
      "interpolation": "Nearest"
    }
  ]
}
```

D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

```

    ],
    "title": "Corine Land Cover",
    "datasource_type": "grid",
    "keywords": [
      "copernicus services",
      "sentinel hub",
      "xcube",
      "raster",
      "systematic",
      "machine learning",
      "land cover",
      "CLMS",
      "corine",
      "derived data",
      "open data"
    ],
    "area_cover": "EEA-38+UK",
    "crs": "EPSG:3035",
    "start_datetime": "1990-01-01T00:00:00Z",
    "end_datetime": "2018-01-01T00:00:00Z",
    "personalData": null,
    "Provenance_name": "EO-based, produced by countries",
    "preprocessing": null,
    "modification": null,
    "provision": null,
    "apis": null,
    "models": null,
    "data_quality": null,
    "quality_control": null,
    "metada_standards": null,
    "distributions": null,
    "access_control": null,
    "resolution": null,
    "use_case_S4E": 1,
    "use_case_WER": 2,
    "use_case_NHM": 1,
    "use_case_NILU": null,
    "use_case_NHM_2": null,
    "ingestion_status": "Completed, year as integer, (1990,2000,2006,2012,2018)",
    "platform": "Both"
  },
  "geometry": {
    "type": "Polygon",
    "coordinates": [
      [
        [
          -31.561261000000002,
          27.405827000000002
        ],
        [
          -31.561261000000002,
          71.409109
        ],
        [
          44.820775,
          71.409109
        ],
        [
          44.820775,
          27.405827000000002
        ],
        [
          -31.561261000000002,
          27.405827000000002
        ]
      ]
    ]
  }
},
"links": [
  {
    "rel": "root",
    "href": "../catalog.json",
    "type": "application/json",
    "title": "data-access catalog"
  }
]

```

D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

```

    },
    {
      "rel": "parent",
      "href": "../catalog.json",
      "type": "application/json",
      "title": "data-access catalog"
    },
    {
      "href":
"https://catalog:JdpsUHpPoqXtbM3@fairicube.rasdaman.com/rasdaman/ows?&SERVICE=WCS&VERSION=2.1.0&REQUEST=DescribeCoverage&COVERAGEID=corine_land_cover&outputType=GeneralGridCoverage",
      "rel": "about",
      "type": "text/xml",
      "title": "Link to the rasdaman coverage description in XML"
    },
    {
      "href": "https://catalog:JdpsUHpPoqXtbM3@fairicube.rasdaman.com/rasdaman-
dashboard/?layers=corine_land_cover",
      "rel": "service",
      "type": "text/html",
      "title": "Link to the rasdaman web application to Access, process gridded data"
    }
  ],
  "assets": {
    "thumbnail_rasdaman": {
      "href":
"https://catalog:JdpsUHpPoqXtbM3@fairicube.rasdaman.com/rasdaman/ows?service=WMS&version=1.3.0
&request=GetMap&layers=corine_land_cover&bbox=27.405827000000002,-
31.5612610000000002,71.409109,44.820775&time=\"1990-01-
01T00:00:00Z\"&width=800&height=600&crs=EPSG:4326&format=image/png&transparent=true&styles=",
      "roles": [
        "thumbnail"
      ]
    }
  },
  "bbox": [
    -31.5612610000000002,
    27.4058270000000002,
    44.820775,
    71.409109
  ],
  "stac_extensions": [
    "https://stac-extensions.github.io/raster/v1.1.0/schema.json",
    "https://stac-extensions.github.io/datacube/v2.0.0/schema.json"
  ]
}

```

The STAC Datacube extension is currently already in *candidate* stage which represents the highest maturity stage before a stable released version. The aim of the STAC Datacube Extension is to provide a way to specify datacube related metadata, especially their dimensions and potentially more in the future as needed in FAIRiCUBE. Non-cubed data can also be described by STAC metadata and included in the FAIRiCUBE Catalog.



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

3.1 Metadata mapping

The table below provides a mapping from the metadata fields collected by the FAIRiCUBE consortium to the STAC naming schema. FAIRiCUBE tries to use existing STAC names whenever possible. However, a set of items used in FAIRiCUBE are currently neither available in the STAC basic standard nor in any of the extensions. The FAIRiCUBE project will therefore propose the expansion of the basic standard or of extensions, where appropriate, to the respective responsible STAC teams.

Notes - Explanations to Table 2::

<i>R / O</i>	<i>Required/ Optional</i>
<i>stac:common</i>	<i>refers to: https://github.com/radiantearth/stac-spec/blob/master/item-spec/common-metadata.md#provider-object</i>
<i>fic</i>	<i>the tag for a proposed FAIRiCUBE-extension</i>
<i>currently implemented STAC name</i>	<i>the current catalog uses these given STAC names</i>
<i>targeted STAC ame</i>	<i>in the next version of the catalog these STAC names shall be used</i>
<i>new: to be proposed to STAC</i>	<i>these STAC names or STAC extension shall be proposed to the STAC committee- Either as new extensions of preferable as additions/changes to existing specifications (extensions or "basic" specs)</i>

Table 1: Mapping from metadata requirements to STAC

Input Term	R / O	multipl icity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
Data item								
Datacube Name	O	0...1	string	A human readable title describing the Item	Corine Land Cover	stac:common:title		
Datacube Identifier (no blanks)	R	1	string	Provider identifier. The ID should be unique within the Collection that contains	Corine_Land_Cover Format: [_a-zA-Z][-\._0-9a-zA-Z]*	stac:item:id		



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				the Item.				
Description	R	1	string	Detailed multi-line description to fully explain the Item	The CORINE Land Cover (CLC) inventory was initiated in 1985 (reference year 1990). Updates have been produced in 2000, 2006, 2012, and 2018. It consists of an inventory of land cover in 44 classes.	stac:common:description		
Data source								
Source	O	0..1	string	Origin of data (institution, etc.) including download URL / email as appropriate	https://land.copernicus.eu/pan-european/corine-land-cover	stac:source_data	stac:processing:lineage	
Thumbnails	O	0..1	string	an URL pointing to the thumbnail		stac:item:assets:thumbnail		
Source Type (grid, vector)	O	0..1	string	Source Type (grid, vector)	e.g. grid	stac:datasource_type		X stac:fid
Organisation	R	0..1	string	REQUIRED if name is missing. Organization/affiliation of the contact.	e.g. EEA	stac:common:providers:organization	stac:contacts:organisation	
Name	R	0..1	string	REQUIRED if organization is missing. The name of the		stac:common:providers:organization_name	stac:contacts:name	



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				responsible person.				
Email	O	0...1	string	Email address at which contact can be made.		stac:common:providers:organization_email	stac:contacts:emails	
ORCID ID	O	0...1	string	ORCID provides a persistent digital identifier (an ORCID iD) that you own and control, and that distinguishes you from every other researcher.		stac:common:providers:ORCID_ID		X field already proposed (https://github.com/stac-extensions/scientific/pull/11) Propose to move it to contacts
Project Purpose	O	0...1	string	Short description of the project		stac:common:providers:project_purpose		X stac:contacts:project_purpose
Documentation link	O	0...*	string	List related publications / public deliverables, state DOI and/or permanent location	https://land.copernicus.eu/pan-european/corine-land-cover	stac:common:providers:doc_link	stac:sci:publications	
Comments	O	0...1	string	Multi-line description to add further	The present 100m raster dataset is the 2018 CLC status layer modified for the	stac:common:providers:comments		X stac:fic



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				provider information such as processing details for processors and producers, hosting details for hosts or basic contact information	purpose of consistent statistical analysis in the land cover change accounting system at EEA. CORINE Land Cover (CLC) data are produced from 1986 for European (EEA member or cooperating) countries.			
Datacube description								
Total area covered	O	0...*	string	Comma separated list of the area(s) covered	World or NO,ES,NL,FI,EE	stac:area_cover	stac:osc:fields_region	
CRS	O	0...1	string number	The spatial reference system for the data, specified as numerical EPSG code. Defaults to EPSG code 4326		stac:proj:epsg		
Horizontal CRS	R	1	string number object	The spatial reference system for the data, specified as numerical EPSG code, WKT2 (ISO 19162) string	Epsg:4326	stac:cube:reference_system		



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				or PROJJSON object. Defaults to EPSG code 4326				
Bbox X – Extent	R	1	number	X – Extent (lower and upper bounds) of the dimension as two-element array. Open intervals with null are not allowed.	-180,18	stac:cube:extent (xmin,xmax)		
Bbox Y – Extent	R	1	number	Y – Extent (lower and upper bounds) of the dimension as two-element array. Open intervals with null are not allowed.	-90 90	stac:cube:extent (ymin,ymax)		
unit of measure	R	1	string	The unit of measurement for the data, preferably compliant to UDUNITS-2 units (singular)	e.g. deg	stac:cube: unit		
Interpolation/Aggregation	O	0...1	string	The method which has	e.g. NA/nearest	stac:cube:interpolation		X stac:cube:interpolation



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				been applied				
Resolution	R	1	string null	If the dimension consists of interval values, the space between the values. Use null for irregularly spaced steps.	e.g. 4 or irregular Should be 1 value as required by UC, not all resolutions of dataset	stac:cube:step		
Vertical CRS	if V-CRS R	0...1	string number object	The spatial reference system for the data, specified as numerical EPSG code, WKT2 (ISO 19162) string or PROJJSON object. Defaults to EPSG code 4326.	EPSG:9389 (EVRF2019 height)	stac:cube:reference_system		
Lower/Upper Bound	if V-CRS R	if V-CRS 2	number	If the dimension consists of ordinal values, the extent (lower and upper bounds) of the values as two-element array.	e.g. 0 - 10	stac:cube:extent (zmin,zmax)		



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multipl icity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				Use null for open intervals.				
unit of measure	if V-CRS R	if V-CRS 1	string	The unit of measurement for the data, preferably compliant to UDUNITS-2 units (singular).	e.g. m	stac:cube:unit		
Interpolation/Aggregation	O	0..1	string	The method which has been applied	Linear/Bilinear	stac:cube:interpolation		X stac:cube:interpolation
Resolution	if V-CRS R	if V-CRS 1	string null	If the dimension consists of interval values, the space between the values. Use null for irregularly spaced steps.	e.g. null or 4 or "irregular" Should be 1 value as required by UC, not all resolutions of dataset	stac:cube:step		
Time – Lower Bound	R	1+1	string	Extent (lower bounds) of the dimension. The dates and/or times must be strings compliant to ISO 8601. null is allowed for open date	2015 - 2021	stac:start_datetime + stac:cube:extent:t (tmin, tmax)		



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				ranges.				
Time – Upper Bound	R	1	string	Extent (upper bounds) of the dimension. The dates and/or times must be strings compliant to ISO 8601. null is allowed for open date ranges.		stac:end_datetime		
unit of measure	R	1	string	The space between the temporal instances as ISO 8601 duration, e.g. P1D. Use null for irregularly spaced steps.	e.g. y	stac:cube:unit		
Interpolation/Aggregation	O	0...1	string	The method which has been applied	e.g. NA/nearest	stac:cube:interpolation		X stac:cube:interpolation
Resolution	R	1	string null	If the dimension consists of interval values, the space between the values. Use null for irregularly	e.g. 3 or "irregular" Should be the values as required by UC, not all resolutions of dataset	stac:cube:step		



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multipl icity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				spaced steps.				
Other - CRS or enumeration of dimension categories	O	0...1	string number object	The spatial reference system for the data, specified as numerical EPSG code, WKT2 (ISO 19162) string or PROJJSON object. Defaults to EPSG code 4326.	Species list: aus bus, aus cus, aus dus...	stac:cube:reference_system		
lower/upper bound (if ordinal)	O	0...1	number	Extent (lower and upper bounds) of the dimension as two-element array. Open intervals with null are not allowed.	n.a.	stac:cube:extent (min,max)		
unit of measure	O	0...1	string	The unit of measurement for the data, preferably compliant to UDUNITS-2 units (singular).	n.a.	stac:cube: unit		
Interpolation/Aggregation	O	0...1	string		n.a.	stac:cube:interpolation		X stac:cube:interpolation



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
Resolution (or "irregular")	O	0...1	string null	If the dimension consists of interval values, the space between the values. Use null for irregularly spaced steps.	categorical	stac:cube:step		
Range Type - Cell components	R	1	string	The descriptive name. Can be distinct from title or id, but is available for historic reasons.	windspeed_u, windspeed_v Note: these strings will be used as the name in the rangeType	stac:raster:bands:band_name		X stac:raster:bands:band_name
Data type	R	1	string	The data type of the pixels in the band.	ie.g. nt8	Stac:raster:bands:data_type		
unit of measure	R	1	string	The unit of measurement for the data, preferably compliant to UDUNITS-2 units (singular).	e.g. m/s	stac:raster:bands:unit		
Null values	R	1	number string	Pixel values used to identify pixels that are no data in the band either by	["no data", 240];["outside area", 255] Null values are to be provided as; separated pairs: [label, null-value]	stac:raster:bands:nodata		



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				the pixel value as a number or nan, inf or -inf (all strings).				
Definition	O	0...1	string	Band definition	https://qudt.org/vocab/quantitykind/Velocity	stac:raster:bands:definition		X stac:raster:bands:definition
Description	O	0...1	string	Band description		stac:raster:bands:description		X stac:raster:bands:description
Category List	O	0...1	string		Set of comma separated triples: [value, name, color_hint]	stac:raster:bands:category_list		X stac:raster:bands:category_list
Comment	O	0...1	string	Multi-line description to add further provider information such as processing details for processors and producers, hosting details for hosts or basic contact information. CommonMark 0.29 syntax		stac:raster:bands:comment		X stac:raster:bands:comment



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				MAY be used for rich text representation.				
DataCube reprojection								
Horizontal Axis – Horizontal CRS	O	0...1	string number object	The spatial reference system for the data, specified as numerical EPSG code, WKT2 (ISO 19162) string or PROJJSON object. Defaults to EPSG code 4326	e.g. EPSG:3035	stac:re_projection_crs		X stac:cube:reproj_axis:reference_system
unit of measure	O	0...1	string	The unit of measurement for the data, preferably compliant to UDUNITS-2 units (singular).	e.g. m	stac:unit		X stac:cube:reproj_axis:unit
Resolution	O	0...1	string null	The space between the values. Use null for irregularly spaced steps.	e.g. 10	stac:resolution		X stac:cube:reproj_axis:resolution
Legal								
License	R	1	string	Item's		stac:common:license		



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				license(s), either a SPDX License identifier, various if multiple licenses apply or proprietary for all other cases. Should be defined at the Collection level if possible.				
Personal Data	O	0...1	string			stac:personalData		X stac:fic
Provenance								
Origin	O	0...1	string		EO-based, produced by countries	stac:provenance_name		X stac:fic
Preprocessing	O	0...1	string	Description of the preprocessing		stac:preprocessing	stac:links(title:preprocessing-execution)	
Source Data	O	0...1	string	link to data		stac:source_data		X stac:links(href:preprocessing_data_source a non-STAC url)
Models	O	0...1	string	A url link to the processing Model		stac:models	stac:links(href:preprocessing-execution)	
Documents &	R	1	string			stac:documentation	stac:links(href:cit	



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
Publications							e-as)	
Data Quality								
Data Quality	O	0...1	string			stac:data_quality		X stac:fic
Quality control	O	0...1	string			stac:quality_control		X stac:fic
Accessibility								
Metadata standards	O	0...*	string	Standards applied		stac:metadata_standards		X stac:fic
APIs	O	0...*	string	access APIs		stac:apis	srtac:links(rel:(external or service)) relationship needed	
Distributions	O	0...*	string			stac:distributions		X stac:fic
Access Control	O	0...*	string			stac:access_control	stac:auth (more fields required)	
Keywords								
Keywords	R	1	string	Multi-line description to add further provider information such as		stac:keywords (from collections)		



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
				processing details for processors and producers, hosting details for hosts or basic contact information. Common Mark 0.29 syntax MAY be used for rich text representation.				
Administrative								
Dates - Creation	O	0...1	string	creation date and time of the corresponding data in UTC.		stac:datetime		
modification	O	0...1	string	description of modifications		stac:modification	stac:processing:lineage	
Provision	O	0...1	string	date of provision		stac:provision		X stac:fic
Ingestion relevant / Use cases								
Climate change (S4E)	O	0...1	number	Use Case		stac:use_case_S4E		X stac:fic:use_case
Biodiversity & agri (WER)	O	0...1	number	Use Case		stac:use_case_WER		X stac:fic:use_case
Biodiversity occurrence	O	0...1	number	Use Case		stac:use_case_NHM		X stac:fic:use_case

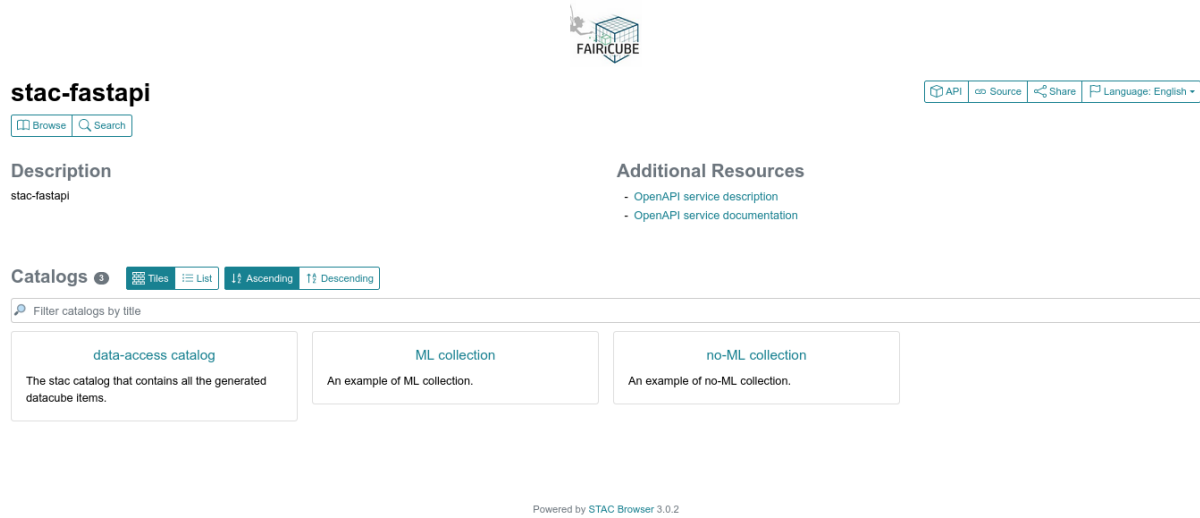


D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

Input Term	R / O	multiplicity	Element type	Notes	Examples	currently implemented STAC name	targeted STAC name	new: to be proposed to STAC
cubes (NHM)								
Neighbourhood building stock (NILU)	O	0..1	number	Use Case		stac:use_case_NILU		X stac:fic:use_case
Platform/Stack	O	0..*	string	deployment	e.g. rasdaman, xcube	stac:platform		X stac:fic:plattorm
Ingestion Status (rasdaman)	O	0..*	string	Version of the context this fields is used in (e.g. Asset, Item or Collection).		stac:ingestion_status		stac:fic:ingest_status

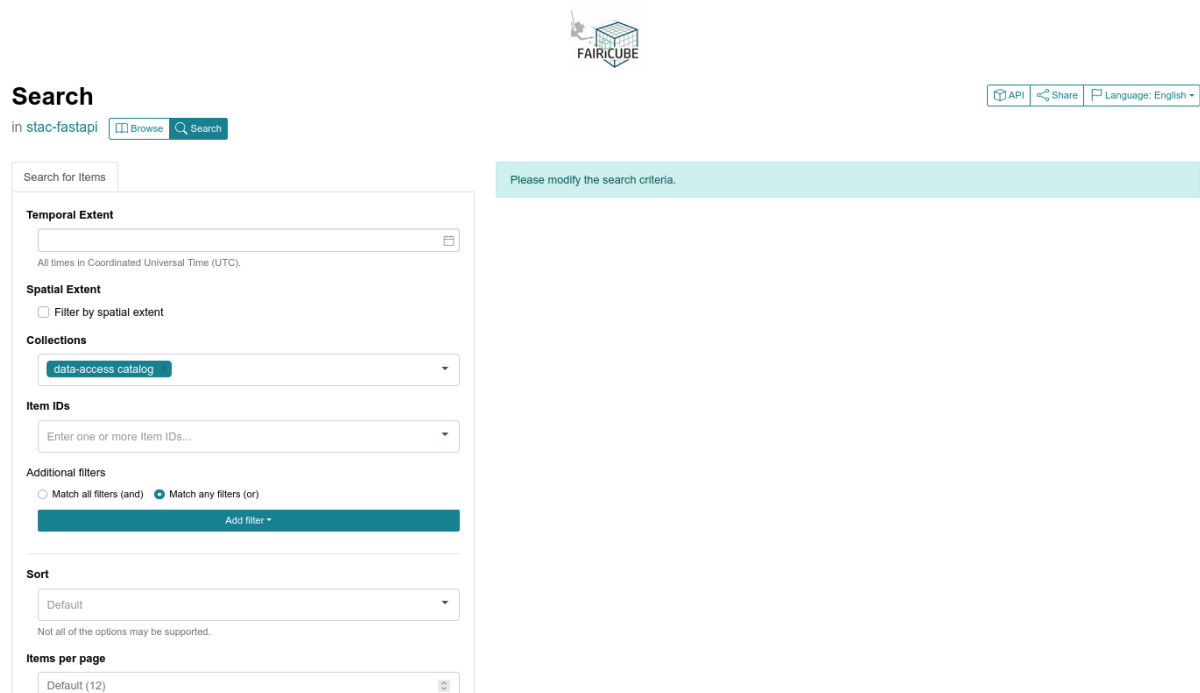
4 FAIRiCUBE Catalog for Data Resources

Below are some screenshots showing the dynamic STAC Browser based on STAC-fastapi.



The screenshot shows the 'stac-fastapi' interface. At the top, there is a FAIRiCUBE logo and navigation links for API, Source, Share, and Language (English). Below the title, there are 'Browse' and 'Search' buttons. The main content is divided into two columns: 'Description' (stac-fastapi) and 'Additional Resources' (OpenAPI service description and documentation). A 'Catalogs' section features a filter input and three catalog cards: 'data-access catalog' (containing all generated datacube items), 'ML collection' (an example of ML collection), and 'no-ML collection' (an example of no-ML collection). The interface is powered by STAC Browser 3.0.2.

Figure 5: Dynamic Catalog based on STAC -fastapi



The screenshot shows the search interface for 'stac-fastapi'. It includes a FAIRiCUBE logo and navigation links for API, Share, and Language (English). The search area has a 'Search for Items' input and a 'Please modify the search criteria.' message. The search filters are organized into sections: 'Temporal Extent' (with a date range input and UTC note), 'Spatial Extent' (with a 'Filter by spatial extent' checkbox), 'Collections' (with a dropdown menu showing 'data-access catalog'), 'Item IDs' (with an input field for item IDs), 'Additional filters' (with radio buttons for 'Match all filters (and)' and 'Match any filters (or)', and an 'Add filter' button), 'Sort' (with a dropdown menu set to 'Default' and a note that not all options may be supported), and 'Items per page' (with a dropdown menu set to 'Default (12)').

Table 6: Dynamic Catalog - Search Interface



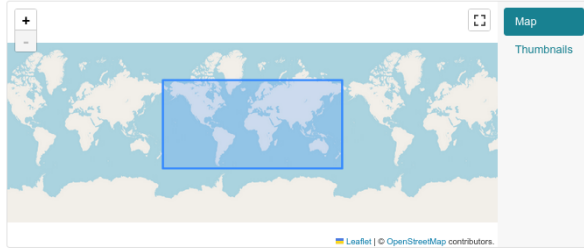
D4.2 Public Listing (Catalog) of FAIRiCUBE data resources



Sentinel-2 L2A 120m Mosaic

in [stac-fastapi](#) [Up](#) [Collection](#) [Browse](#) [Search](#)

[API](#) [Source](#) [Share](#) [Language: English](#)



Asset

> Thumbnail [THUMBNAIL](#) [PNG](#)

Additional Resources

About this resource

- [Website describing the collection](#)
- [Details about running Evalsripts](#)

OGC WMTS web map

- [205fb2e0-0deb-464d-9103-82d33ba73b5d \(services.sentinel-hub.com\)](#)

Processing instructions/code

- [Evalsript to generate True Color imagery](#)
- [Evalsript to generate False Color imagery](#)

Description

Sentinel-2 L2A 120m mosaic is a derived product, which contains best pixel values for 10-daily periods, modelled by removing the cloudy pixels and then performing interpolation among remaining values. As clouds can be missed in the removal process and as there are some parts of the world, which have lengthy cloudy periods, clouds might be remaining in some parts. The actual modelling script is available [here](#).

Collection

[data-access catalog](#)

The stac catalog that contains all the generated datacube items.

Provider

> Sentinel Hub

[PROCESSOR](#)

General

License	License
Keywords	<ul style="list-style-type: none"> - sentinel hub - xcube - raster - systematic - satellite imagery - multi spectral imagery - machine learning - agriculture - open data - sentinel

Table 7: Dynamic Catalog – Data item description



5 Data Resources

At the initial setup of FAIRiCUBE Hub the following existing data collections were made available via data federation to the FAIRiCUBE users:

- European Data Cube
- Sentinel Hub
- EarthServer
- CoperniCUBE

In addition, FAIRiCUBE users can request additional datasets to be integrated. For the initial collection of additional data resources, a Spreadsheet (Inventory-Sheet) had been setup to collect users data needs. The following Table 2: shows the relevant data requests claimed the different Use Cases. A list of data resources currently available in FAIRiCUBE can always be viewed using the catalog at <https://catalog.fairicube.eu>.

Table 2: Data requests by Use Cases as per Inventory Sheet (as of December 2023)

Datacube Name	Climate change (S4E)	Bio-diversity & agri (WER)	Drosophila Genetics (NHM)	Neighbourhood building stock (NILU)	Biodiversity occurrence cubes (NHM)
Corine Land Cover				X (SentinelHub)	X
High Resolution Layer - Imperviousness				X	
High Resolution Layer - Forest type					X
High Resolution Layer - Forest -tree cover				X	X
High Resolution Layer - Water & Wetness					X
High Resolution Layer - Small Woody Features		X			X
CLMS Urban Atlas				X	
CLMS Urban Atlas Street Tree Layer				X	
Population by Urban Atlas polygon				X	
Riparian Zones Land Cover and Land Use					X
Natura 2000 land-cover / land use classification				X	X
NUTS regions				X	X
Urban Audit city delineations (FUA, city, commuting)				X	

D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

zone)					
Local Administrative Units (LAUs)				X	
Climate data - Temperature		X		X	X
Climate data - Precipitation		X		X	X
Climate data - Wind				X	
Climate data - Universal Thermal Comfort Index				X	
Climate extremes indices and heat stress indicators				X	X
Copernicus DEM				X (SentinelHub)	
Urban Audit statistical data				X	
National Database of Fauna and Flora (NDFP)		X			
Global Biodiversity Information Facility (GBIF)		X			X
Essential Biodiversity variables (EBV)					X
Agrodatacube		X			
Sentinel 2		X			
Actual height database (AHN)		X			
Soil grids		X			X

5.1 Resources made available in FAIRiCUBE

The following list presents a snapshot in time of the available datasets supplied to FAIRiCUBE. The actual listing of datasets available can always be investigated via the catalog, found at (<https://catalog.eoxhub.faircube.eu/>):

- 2m_Temperature_ERA5_reanalysis
- 3D_building_model_LOD2
- Actual_height_database_AHN
- Agricultural_nature_management_ANB
- Agrodatacube
- CLMS_Urban_Atlas
- CLMS_Urban_Atlas_Street_Tree_Layer
- Climate_data_Physiologically_Equivalent_Temperature_PET
- Climate_data__Precipitation
- Climate_data__Temperature
- Climate_data__Universal_Thermal_Comfort_Index



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

- Climate_data__Wind
- Climate_extremes_indices_and_heat_stress_indicators
- Copernicus_DEM
- Crop_Parcel_Registration_BRP
- DTM_and_DSM_data
- Distribution_data_of_Drosophila
- EPISCOPE_energy_calculation
- EPISCOPE_material_intensity
- ERA5_Land_monthly
- ERA5_global_climate_data
- ERA5_monthly_averaged_data
- Essential_Biodiversity_variables_EBV
- Genomic_data_of_Drosophila
- Global_Biodiversity_Information_Facility_GBIF
- Global_Pesticide_Grids_PESTCHEMGRIDS,_v1.01
- HR_VPP_NDVI,_PPI,_FAPAR,_LAI
- High_Resolution_Layer__Small_Woody_Features
- High_Resolution_Layer__Small_Woody_Features
- INSPIRE_Buildings_geoportal
- KPIs_on_agricultural_land_management
- LGN
- Local_Administrative_Units_LAUs
- Local_research_outcomes
- NUTS_regions
- National_Database_of_Fauna_and_Flora_NDFF
- Open_Street_Map_OSM
- Population_by_Urban_Atlas_polygon
- RS_derived_signals,_markers_and_indices
- Riparian_Zones_Land_Cover_and_Land_Use
- SPARTACUS_v2.1
- Sentinel_1
- Sentinel_2
- Soil_grids
- Statistical_data_on_species
- The_Earth_Observations_Toolkit_for_Sustainable_Cities_and_Human_Settlements
- Urban_Audit_city_delineations_FUA,_city,_commuting_zone
- Urban_Audit_statistical_data
- Vegetation_Indexes_TBD
- Weather
- city_features_collection
- corine_land_cover
- dominant_leaf_type
- dominant_leaf_type_virtual_coverage_index_2018
- eu_demography
- european_settlement_map
- forest_type_virtual_coverage_index_2012_2015
- forest_type_virtual_coverage_index_2018
- grassland_status_virtual_coverage_index_2015
- imperviousness_virtual_coverage_index_2006_2015
- imperviousness_virtual_coverage_index_2018
- tree_cover_density_virtual_coverage_index_2012_2015
- tree_cover_density_virtual_coverage_index_2018
- water_and_wetness_virtual_coverage_index_2015



D4.2 Public Listing (Catalog) of FAIRiCUBE data resources

6 References

- [1] STAC website: <https://stacspec.org> (accessed 30-05-2024)
- [2] STAC Datacube Extension: <https://github.com/stac-extensions/datacube> (accessed 30-05-2024)
- [3] STAC Common Metadata: <https://github.com/radiantearth/stac-spec/blob/master/item-spec/common-metadata.md> (accessed 30-05-2024)
- [4] STAC Extensions: <https://github.com/stac-extensions/> (accessed 30-05-2024)
- [5] STAC Best Practices: <https://github.com/radiantearth/stac-spec/blob/master/best-practices.md> (accessed 30-05-2024)
- [6] STAC Item Specification: <https://github.com/radiantearth/stac-spec/blob/master/item-spec/item-spec.md> (accessed 30-05-2024)
- [7] STAC Collection Specification: <https://github.com/radiantearth/stac-spec/tree/master/collection-spec> (accessed 30-05-2024)
- [8] Processing: <https://github.com/stac-extensions/processing> (accessed 30-05-2024)
- [9] Datacube Extension: <https://github.com/stac-extensions/datacube> (accessed 30-05-2024)
- [10] Raster Extension: <https://github.com/stac-extensions/raster> (accessed 30-05-2024)
- [11] Open Science Catalog Extension: <https://github.com/stac-extensions/osc> (accessed 30-05-2024)
- [12] Contacts Extension: <https://github.com/stac-extensions/contacts> (accessed 30-05-2024)
- [13] Scientific Citation Extension: <https://github.com/stac-extensions/scientific> (accessed 30-05-2024)
- [14] Version extension: <https://github.com/stac-extensions/version> (accessed 30-05-2024)
- [15] D4.1 FAIRiCUBE Hub Architecture
- [16] D5.2 Description of the datacube ingestion pipelines