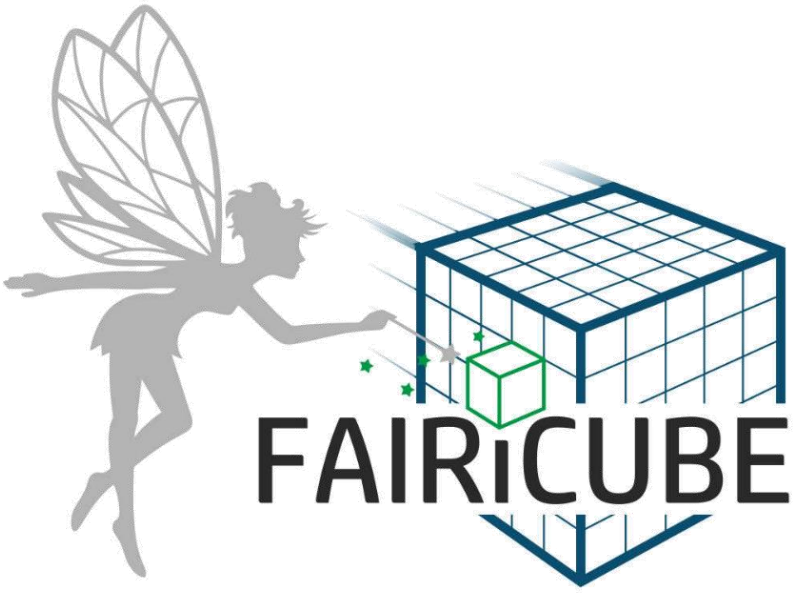

FAIRICUBE – F.A.I.R. INFORMATION CUBES

WP2: Use

M12: Final version of the Report on UC data source synergies released

Deliverable Lead: 4sfera
Deliverable due date: 30/06/2024

Version: 1.0
05/08/2024

Document Control Page

Document Control Page	
Title	Final version of the Report on UC data source synergies released
Creator	4SF
Description	Milestone M12 Final version of the Report on UC data source synergies released
Publisher	"FAIRICUBE – F.A.I.R. information cubes" Consortium
Contributors	
Date of delivery	30/06/2024
Type	Text
Language	EN-GB
Rights	Copyright "FAIRICUBE – F.A.I.R. information cubes"
Audience	☒ Public ☐ Confidential ☐ Classified
Status	☐ In Progress ☐ For Review ☒ For Approval ☐ Approved

Revision History			
Version	Date	Modified by	Comments
0.1	26/06/2024	Jaume Targa	First draft and complete edit
0.2	30/06/2024	Maria Ricci	Internal Review
1.0	30/06/2024	Stefan Jetschny	Final document, format checking

Disclaimer

This document is issued within the frame and for the purpose of the FAIRICUBE project. This project has received funding from the Horizon Europe 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.

Table of Contents

List of Tables

1 Introduction

To demonstrate, execute and improve the use of the FAIRiCUBE Hub, 5 Use Cases (UCs) were designed. To ensure efficiency on data ingestion and harmonisation of data, we need to identify and leverage potential synergies between UCs. Where possible, data sources, processing and ingested data need to be re-used. To this end, 4SF has coordinated the elaboration of an initial inventory of data sources with all 5 UC. The aim of the inventory was to gather an initial list of all the different data sources that all UCs envisage to be using and identify common data source needs between UC. The identification of data source synergies is a clear example of knowledge sharing across UCs and is essential for efficiently executing the UCs, not only to minimise data access but to cut down on processing/ingestion time.

The initial inventory has been upgraded to a data request process through a web GUI. This process is the same for both the rasdaman and EOX technical platform pillars. When a Use Case needs an additional dataset, a new record is created on the FAIRiCUBE data requests GitHub repository via a new issue. This streamlined process is facilitated by a web GUI available at https://data-request.fairicube.eu, where users can provide necessary information through a form. This form generates a machine-readable request that starts the ingestion process.

The web GUI serves as a frontend to a GitHub repository. When the form is submitted, it creates a GitHub Pull Request as a new branch, with the user automatically following the branch for updates based on their GitHub notification settings. Progress, issues, and discussions are documented in a GitHub issue linked to the Pull Request, allowing everyone involved to monitor and provide feedback.

2 Formal deliveries contributing to M12

As a formal measure to validate the achievement of the M4 milestone, the deliverable shown in (Table 1) was formulated and documented.

Description	Lead Beneficiary	Type	Dissemination Level	Due Date
D2.1: Report on UC data source synergies	4SF	R	PU	30/06/2024

Table 1: Formal deliverables contributing to M12

The **deliverable D2.1** and the respective report provided a summary data source need of all UC and the synergies between them. The initial FAIRiCUBE data sources inventory is available in FAIRiCUBE website. The FAIRiCUBE catalog includes data requested and ingested in the different platforms. The different data sets can be accessed here:

- https://catalog.eoxhub.fairicube.eu/ (all available collections)
- https://fairicube.rasdaman.com/rasdaman/ows#/services

3 Summary of synergies

The initial aim of compiling an inventory of data sources from all UCs was to identify initial data source synergies across these UCs. All 5 UCs initially identified a total of 54 data sources. Once the initial draft of data sources inventory was compiled, the team worked with all UCs to identify those data sources which could be of use, despite not being initially listed as a required data source. The identification of further data source synergies is a clear example of knowledge sharing across UCs.

The exercise detected 25 data sources which were identified as potentially useful sources with high or low priority further use by at least 2 UCs. A summary of these sources is presented in Table 1. The synergetic data sources are related to geographical parameters, land cover, data on buildings, biological parameters and climate variables. It is worth noting that all of the 25 synergetic sources have been identified as high priority by at least one UC. Additionally, 3 data sources (Open Street map, Temperature data and Digital Elevation Model) have been identified as further use by all 5 UCs.

Data sources are classified as high priority (1), low priority (2) and a blank meaning that the UC does not plan to use the source. This has assisted in identifying the common priorities across UCs.

Table 1: Summary of synergetic data sources found. Columns UC1, UC2, UC3, UC4 and UC5 represent the priority of use of each data source, with 1 meaning high priority, 2 meaning low priority and a blank meaning that the UC does not plan to use the source.

Name	Source	Origin	Resp. maint.	Coverage	Spatial res.	Time cov.	UC1	UC2	UC3	UC4	UC5
Open Street Map	www.openstreetmap.org		OSM	World	vector		2	2	1	1	2
INSPIRE Buildings	https://inspire-geoportal.ec.europa.eu/overview.html?view=themeOverview&theme=bu		EEA	EEA-38	vector		2			1	
Corine Land Cover	https://land.copernicus.eu/pan-european/corine-land-cover	EO-based, produced by countries	EEA	EEA-38+UK	100m	1990 2000 2006 2012 2018	1	2	1		2
Imperviousness	https://land.copernicus.eu/pan-european/high-resolution-layers/imperviousness/status-maps	EO-based	EEA	EEA-38+UK	10m 20m	2006 2009 2012 2015 2018	1	2	2		2
Forest type	https://land.copernicus.eu/pan-european/high-resolution-layers/forests/forest-type-1/status-maps	EO-based	EEA	EEA-38+UK	10m 20m	2012 2015 2018	1	2	2		2
Forest - tree cover	https://land.copernicus.eu/pan-european/high-resolution-layers/forests/tree-cover-density/status-maps	EO-based	EEA	EEA-38+UK	10m 20m	2012 2015 2018	1	2	2		2
Forest - Dominant Leaf Type	https://land.copernicus.eu/pan-european/high-resolution-layers/forests/dominant-leaf-type/status-maps	EO-based	EEA	EEA-38+UK	10m 20m	2012 2015 2018	1	2	2		2
Grassland	https://land.copernicus.eu/pan-european/high-resolution-layers/grassland	EO-based	EEA	EEA-38+UK	10m 20m	2015 2018	1	2	2		2
Water & Wetness	https://land.copernicus.eu/pan-european/high-resolution-layers/water-wetness	EO-based	EEA	EEA-38+UK	10m 20m	2015 2018	1	2	2		2
Small Woody Features	https://land.copernicus.eu/pan-european/high-resolution-layers/small-woody-features	EO-based	EEA	EEA-38+UK	10m 20m	2015 (2018)	1	2	2		2
European Settlement Map	https://land.copernicus.eu/pan-european/GHSL/european-settlement-map	EO-based	JRC	EEA-38+UK	2,5m 10m 100m	2012 2015	1		1		
CLMS Urban Atlas	https://land.copernicus.eu/local/urban-atlas	EO-based	EEA	EEA-38+UK	vector	2006 2012 2018	1	2		2	2

Name	Source	Origin	Resp. maint.	Coverage	Spatial res.	Time cov.	UC1	UC2	UC3	UC4	UC5
CLMS Urban Atlas Street Tree Layer	https://land.copernicus.eu/local/urban-atlas	EO-based	EEA	EEA-38+UK	vector	2012 2018	1	2			2
Population by Urban Atlas polygon	https://land.copernicus.eu/local/urban-atlas/population-estimates-by-urban-atlas-polygon	EO-based	EEA	EEA-38+UK			1			2	
Natura 2000 landcover/land use	https://land.copernicus.eu/local/natura	EO-based	EEA	EEA-38+UK	vector	2006 2012 2018	1	2			2
NUTS regions	https://ec.europa.eu/eurostat/web/nuts/background	national reporting	Eurostat	EEA-38+UK, -BA, XK	vector		1	2	2		2
Urban Audit city delineations (FUA, city, commuting zone)	https://ec.europa.eu/eurostat/web/cities/background		Eurostat	EU-27+EFTA	vector		1			2	
Temperature	https://cds.climate.copernicus.eu/cdsapp#!/dataset/derived-near-surface-meteorological-variables	modelled	C3S	global	grid		1	2	1	2	2
Physiologically Equivalent Temperature	https://climate.copernicus.eu/thermal-assessment-tool		C3S	global	grid		1			2	
Climate extremes indices and heat stress indicators	https://cds.climate.copernicus.eu/cdsapp#!/dataset/sis-extreme-indices-cmip6?tab=overview		C3S		grid		1	2	1		2
Copernicus DEM	https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1	EO-based	Copernicus	global	grid 10m 30m 90m		2	2	1	2	2
Global Biodiversity Information Facility	https://www.gbif.org/	in-situ observations	GBIF Secretariat	World	various	2014-2022		1	1		1
HR VPP (NDVI, PPI, FAPAR, LAI)	https://land.copernicus.eu/pan-european/biophysical-parameters/high-resolution-vegetation-phenology-and-productivity	Sentinel 1, 2	CLMS	Europe	10 m	2014-2022	2	2			2
Soil grids	https://soilgrids.org/		ISRIC	World	250m	2014-2022	2	1			1
3D building model (LOD2)	https://geoe3platform.eu/geoe3	GeoE3	GeoE3	NO, ES, NL, FI, EE	vector		2			1	

4 Data source management and ingestion

The FAIRiCUBE catalogue (https://catalog.eoxhub.fairicube.eu/?.language=en) contains around 100 diverse data sources with various temporal and spatial resolutions. These sources provide data on meteorology, oceans, land use, geography, demographics, vegetation, and air quality. Some datasets have been identified by Use Cases (UCs) as useful for future analysis. You can explore the catalogue at FAIRiCUBE Catalogue. To add new data, a data request WebGUI has been created. More information can be found in deliveries D4.1 and D5.2.

As UCs implement their machine learning tasks under WP3, the data source list will grow. UCs have unique needs that may require additional data over time. For example, two UCs focusing on urban areas will analyse specific cities like Oslo, Vienna, and Barcelona. Once these locations are set, local data will be added to the repository. New data requests should follow the procedure outlined at Catalog Editor. The catalogue is expected to expand and update as it develops.