

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

Work Package 2: Use
Deliverable 2.4: Sharing UC workshops presentations
and minutes

Deliverable Lead: 4sfera
Deliverable due date: 30/06/2023

Version: 1.0
2024-06-26

Document Control Page

Document Control Page	
Title	D2.4 Sharing UC workshops presentations and minute
Creator	Jaume Targa, María Colina and Cristina Carnerero
Description	Sharing UC workshops presentations and minute
Publisher	"FAIRICUBE – F.A.I.R. information cubes" Consortium
Contributors	4SF, NIL, WER, NHM, S4E
Date of delivery	30/06/2024
Type	Text
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	17/05/2024	Jaume Targa	Internal draft
0.2	24/06/2024	Jaume Targa	Further improvements
1.0	25/06/2024	Maria Ricci	Internal review and 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 European Union's Horizon research and innovation programme under grant agreement No. 101059238. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the European Commission.

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



Table of Contents

Document Control Page	2
Disclaimer	3
Table of Contents	4
List of Figures.....	7
List of Tables.....	8
1 Contextualisation of this deliverable	9
1.1 Overall objective of WP2	9
1.2 Description of WP2 work.....	9
1.3 Description of Task 2.5	9
2 Introduction	10
2.1 The aim of workshops.....	10
3 Workshops	13
3.1 WP2 UC synergy workshop.....	13
3.2 General summary of workshops.....	13
3.2.1 Key achievements on WP2 UC Synergy meeting.....	13
3.2.2 Key achievements on WP2-WP3 UC Synergy meeting.....	14
3.3 WP2 UC synergy monthly workshop (no.1).....	15
3.3.1 UC1 - Urban Climate	15
3.3.2 UC2 - Biodiversity-Agriculture Nexus	16
3.3.3 UC4 - Building Stock.....	16
3.3.4 UC3 - Drosophila Genetics	16
3.4 WP2 UC synergy monthly workshop (no.2).....	17
3.4.1 UC1 - Urban Climate	17
3.4.2 UC3 - Drosophila Genetics	17
3.4.3 UC4 - Building Stock.....	18
3.5 WP2 UC synergy monthly workshop (no.3).....	19
3.5.1 UC1 - Urban Climate	19



3.5.2	UC2 - Biodiversity-Agriculture Nexus	19
3.5.3	UC3 - Drosophila Genetics	20
3.5.4	UC4 - Building Stock.....	20
3.6	WP2 UC synergy monthly workshop (no.4)	21
3.6.1	UC1 - Urban Climate	21
3.6.2	UC2 - Biodiversity-Agriculture Nexus	21
3.6.3	UC3 - Drosophila Genetics	21
3.6.4	UC4 - Building Stock.....	22
3.7	WP2 UC synergy monthly workshop (no.5)	23
3.7.1	UC1 - Urban Climate	23
3.7.2	UC2 - Biodiversity-Agriculture Nexus	23
3.7.3	UC3 - Drosophila Genetics	24
3.7.4	UC4 - Building Stock.....	24
3.8	WP2 UC synergy monthly workshop (no.6)	25
3.9	WP2-WP3 UC synergy monthly workshop (no.1)	26
3.9.1	UC1 - Urban Climate	26
3.9.2	UC2 - Biodiversity-Agriculture Nexus	26
3.9.3	UC3 - Drosophila Genetics	26
3.9.4	UC4 - Building Stock.....	26
3.9.5	UC5 – Phytosociological method	26
3.10	WP2-WP3 UC synergy monthly workshop (no.2)	28
3.10.1	UC1 - Urban Climate	28
3.10.2	UC2 - Biodiversity-Agriculture Nexus	28
3.10.3	UC3 - Drosophila Genetics	28
3.10.4	UC4 - Building Stock.....	28
3.10.5	UC5 – Phytosociological method	28
3.11	WP2-WP3 UC synergy monthly workshop (no.3)	29
3.11.1	UC1 - Urban Climate	29
3.11.2	UC2 - Biodiversity-Agriculture Nexus	29
3.11.3	UC3 - Drosophila Genetics	29
3.11.4	UC4 - Building Stock.....	29
3.11.5	UC5 – Phytosociological method	29



- 3.12 WP2-WP3 UC synergy monthly workshop (no.4) 31
 - 3.12.1UC1 - Urban Climate31
 - 3.12.2UC2 - Biodiversity-Agriculture Nexus31
 - 3.12.3UC3 - Drosophila Genetics31
 - 3.12.4UC4 - Building Stock.....31
 - 3.12.5UC5 – Phytosociological method31
- 3.13 WP2-WP3 UC synergy monthly workshop (no.5) 32
 - 3.13.1UC1 - Urban Climate32
 - 3.13.2UC2 - Biodiversity-Agriculture Nexus32
 - 3.13.3UC3 - Drosophila Genetics33
 - 3.13.4UC4 - Building Stock.....33
 - 3.13.5UC5 – Phytosociological method33
- 4 Conclusion 35



List of Figures

<i>Figure 1 : Initial schedule for initial WP2 UC Synergy Workshops</i>	10
<i>Figure 2 : Summary of UC1 Technical setup.</i>	15
<i>Figure 3 : UC3 Objective summarised.</i>	17
<i>Figure 4 : UC2 workflow schema.</i>	19
<i>Figure 5 : UC4 stakeholder and outreach support.</i>	22
<i>Figure 6 : UC1 data analysis update.</i>	23
<i>Figure 7 : UC5 data analysis plan.</i>	27



List of Tables

Table 1 : UC Synergy monthly workshops. _____ 12



1 Contextualisation of this deliverable

1.1 Overall objective of WP2

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

1.2 Description of WP2 work

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

1.3 Description of Task 2.5

Sharing individual experiences with the FAIRiCUBE HUB framework is essential for the progress of our project and the success of each task. A series of workshops to enhance knowledge sharing and learning between UCs have been organised.

This delivery aims to summarise the series of workshop carried out up to M24.

2 Introduction

The primary goal of WP2 is to ensure efficient execution of the UCs. In alignment with the FAIRiCUBE objectives, a series of UC synergies' workshops are scheduled to knowledge sharing and facilitate learning transfer across all UCs involved.

The workshops' main aim is knowledge sharing and learning across various UCs. A series of workshops were initiated for experience sharing. These workshops were designed to provide a structured platform for the exchange of insights, progress updates, and collaborative learning among the participating UCs.

2.1 The aim of workshops

A series of workshops were designed to gather and share experience from individual UCs. The workshop series were scheduled to run every month since February 2023, with each session focusing on specific aspects of the UCs' progress and plans. The following outlines the schedule and key objectives of each workshop:

Who	Topic	Time	
1. February	UC1	Update of UC objectives	2 min
		Description of data analysis planned	10 min
	UC2	Update of UC objectives	2 min
		Description of data analysis planned	10 min
	UC3	Update of UC objectives	2 min
		Description of data analysis planned	10 min
	UC4	Update of UC objectives	2 min
		Description of data analysis planned	10 min
	All	Discussion	30 min

Who	Topic	Time	
3. April	UC1	Update of data sources	2 min
		Update of data analysis and expected results	10 min
	UC2	Update of data sources	2 min
		Update of data analysis and expected results	10 min
	UC3	Update of data sources	2 min
		Update of data analysis and expected results	10 min
	UC4	Update of data sources	2 min
		Update of data analysis and expected results	10 min
	All	Discussion	30 min

Who	Topic	Time	
2. March	UC1	Update of data analysis plan	2 min
		Description of expected results	10 min
	UC2	Update of data analysis plan	2 min
		Description of expected results	10 min
	UC3	Update of data analysis plan	2 min
		Description of expected results	10 min
	UC4	Update of data analysis plan	2 min
		Description of expected results	10 min
	All	Discussion	30 min

Who	Topic	Time	
4. May	UC1	Update of data sources/ingestion	2 min
		Update of data analysis and expected results	10 min
	UC2	Update of data sources/ingestion	2 min
		Update of data analysis and expected results	10 min
	UC3	Update of data sources/ingestion	2 min
		Update of data analysis and expected results	10 min
	UC4	Update of data sources/ingestion	2 min
		Update of data analysis and expected results	10 min
	All	Discussion	30 min

Figure 1 : Initial schedule for initial WP2 UC Synergy Workshops

The first workshop (February 2023) set the stage for the series by focusing on updating the objectives of each UC. Participants provided detailed descriptions of the data analysis planned for their respective UCs. This initial meeting aimed to ensure that all UCs were aligned on the goals and methodologies, setting a clear direction for subsequent workshops.

The second workshop (March 2023) focused on updates to the data analysis plans. Participants shared their progress and provided descriptions of the expected results. This session was crucial for refining analysis approaches and setting realistic expectations for outcomes.

In April, in the third workshop participants concentrated on updating data sources and further elaboration on the data analysis plans and expected results. This workshop aimed to ensure that all UCs had access to the necessary data and that their analysis strategies were on track.



The fourth workshop in May 2023 continued the focus on data, with updates on data sources and ingestion processes. UCs also provided further updates on their data analysis efforts and the anticipated results. This session emphasized the importance of data integrity and accuracy in the overall analysis process.

As the series progressed, the fifth workshop in June marked a significant point with comprehensive updates on the data analysis plans and the current state of analysis and expected results. Additionally, this workshop introduced stakeholder and outreach support, highlighting the importance of engaging with relevant stakeholders to ensure broader support and impact.

The workshop in July served as a culmination of the series, with participants providing final updates on their data analysis efforts. This workshop was designed to consolidate the insights gained over the previous months, celebrate the achievements, and outline next steps for ongoing collaboration and improvement.

In November 2023, it was decided together with WP3 lead, to join WP2 and WP3 monthly meetings. In continuation of the effort to enhance collaboration and knowledge sharing among UCs, the series of workshops has been upgraded to WP2 and WP3 synergy meetings from December 2023. These workshops have provided a platform for detailed discussions on progress, challenges, and future plans. The following workshops were designed to provide a regular update of outcomes of the UCs, covering all aspects. *Table 1* provides an overview of the planned monthly workshops. The following paragraphs provide a description of the conducted and planned activities.

Table 1 : UC Synergy monthly workshops.

Date	Workshop (link to minutes)	Description
7/2/2023	WP2 UC synergy monthly workshop (no.1)	Set the structure of next meeting and discuss the aim of the workshop series, which is to better understand data analysis for each UC, and knowledge sharing, with specific aim to identify synergies.
7/3/2023	WP2 UC synergy monthly workshop (no.2)	The workshop aimed to provide an update of the data analysis plan and description of expected results.
1/4/2023	WP2 UC synergy workshop (no.3)	Aim to provide an update of data sources, and update of data analysis and expected results.
6/6/2023	WP2 UC synergy workshop (no.4)	Aim to provide an update of data sources, progress on the UC, update of data analysis and expected results
4/7/2023	WP2 UC synergy workshop (no.5)	Updates on the progress and challenges of each UC, including data ingestion, taxonomy, and 3D modelling of buildings. Collaboration and data sharing between UCs were emphasised for more accurate and comprehensive analysis.
03/10/2023	WP2 UC synergy workshop (no.6)	Focus on potential cross-UC synergies and discussed the validation of data ingestion, with a revised agenda to accommodate more time for key topics.
05/12/2023	WP2 & WP3 UC synergy monthly workshop (no.1)	Focus on discussing the current progress and future requirements of UCs, including data ingestion, ML applications, and validation protocols. UCs shared updates, identified challenges, and proposed solutions to enhance data analysis and processing capabilities.
06/02/2024	WP2 & WP3 UC synergy monthly workshop (no.2)	UCs provided updates on their progress, including data analysis, visualization tools, and future plans. Discussion on challenges such as data availability and the need for collaboration on energy estimation and building materials.
12/03/2024	WP2 & WP3 UC synergy monthly workshop (no.3)	The workshop focused on updating, refining and aligning the presentations for each UC, emphasizing the UC and their workflows.
07/05/2024	WP2 & WP3 UC synergy monthly workshop (no.4)	Updates from UCs. Emphasizing the importance of synergies and collaboration to achieve project goals. Key discussions included data ingestion, platform usage, and future planning, with a notable presentation on the Vienna Fly project.
04/06/2024	WP2 & WP3 UC synergy monthly workshop (no.5)	UCs provide full update including progress and challenges. Emphasis on the importance of data validation and stakeholder engagement. Discussions on potential synergies between UCs, the need for data ingestion, and upcoming training sessions.

3 Workshops

3.1 WP2 UC synergy workshop

A series of workshops, starting in February 2023 have been organised to enhance knowledge sharing and learning between UCs. This set of monthly workshops complement UC specific weekly/bi-weekly meetings done individually by each UC. Up to October 2023, 6 workshops have been organised under WP2. From December 2023 to June 2024, 5 workshops have been organised together with WP3.

3.2 General summary of workshops

The WP2 UC Synergy Monthly Workshops, held from February to October 2023 brought together teams working on different UCs to discuss their objectives, share progress, and address challenges. Key achievements of the workshops included:

3.2.1 Key achievements on WP2 UC Synergy meeting

3.2.1.1 Establishment of Collaborative Framework

The workshops set up a structure for ongoing discussions and collaboration across UCs, enabling participants to understand each other's objectives and methodologies.

3.2.1.2 Progress in Data Infrastructure

Significant strides were made in setting up the necessary infrastructure for data ingestion and analysis, including the implementation of cloud-based solutions and data cubes.

3.2.1.3 Identification of Synergies

Teams identified multiple synergies, particularly between UCs working on related topics, such as urban climate and building stock, which led to shared expertise and resources.

3.2.1.4 Technical Challenges and Solutions

The workshops provided a platform to discuss and address technical challenges, such as data ingestion procedures, model implementation, and database access. Solutions and best practices were shared among participants.

3.2.1.5 Data Quality and Validation

Emphasis was placed on ensuring data quality and integrity, with discussions on developing guidelines for data validation and addressing discrepancies in data sources.

3.2.1.6 Stakeholder Engagement

The workshops facilitated discussions on engaging stakeholders, particularly in presenting project benefits and addressing data confidentiality issues. Overall, the WP2 UC Synergy Monthly Workshops achieved substantial progress in enhancing collaborative efforts, improving data infrastructure, and addressing technical and procedural challenges, thereby advancing the objectives of the WP2 project.



3.2.2 Key achievements on WP2-WP3 UC Synergy meeting

3.2.2.1 Enhanced Collaboration and Synergy

Continued and strengthened collaboration across UCs, focusing on integrating Earth Observation data and refining analysis methodologies. Identified and leveraged synergies, particularly between UC3 (Drosophila Genetics) and UC5 (Phytosociological Method), enhancing data tools and methodologies.

3.2.2.2 Technical Progress and Challenges Addressed

Progressed in ML integration and visualization tools to support complex data interpretation (e.g., data cubes and dashboards in UC1). Each UC refined their workflows, addressing challenges in data preprocessing and enhancing methodologies like species distribution modelling (UC2) and energy performance estimation (UC4). Emphasized rigorous data validation protocols, including the establishment of standards for genetic and environmental data (UC3) and ensuring accuracy in species distribution mapping (UC5). Addressed technical hurdles such as data access issues and system performance optimizations (UC4, UC5), ensuring robustness in data handling and analytical processes.

3.2.2.3 Stakeholder Engagement and Application Development

Engaged city administrations (e.g., Luxembourg, Vienna) to collaborate on urban climate and building stock analyses. Developed tools such as the "worm picker" for genetic data analysis (UC3), enhancing data accessibility and analytical capabilities.

3.2.2.4 Progress in Machine Learning and Visualization

Planned and executed machine learning pipelines for advanced data analysis (UC5), leveraging platforms like EOX for enhanced computational capabilities.

Implemented data cubes and visualization tools (UC1) to facilitate the interpretation and presentation of complex datasets.

3.3 WP2 UC synergy monthly workshop (no.1)

The first synergy monthly workshop was held on 7th February 2023. The workshop aimed to set the structure and discuss the aim of the workshop series, which was to better understand Data Analysis for each UC, and knowledge sharing, with specific aim to identify synergies. Each UC was asked to provide an overview of UC objectives and provide a description of data analysis planned.

General discussions were held on notable challenges in comparing the two platforms, EOX and Rasdaman. Conceptual challenges were discussed, highlighting difficulties in understanding concepts such as registration, ingestion, and update within the EOX system. For future meetings, a template was proposed. Reporting failures was emphasized, with the suggestion of a "failure of the week" approach to facilitate learning from each other.

3.3.1 UC1 - Urban Climate

At the time, there were ongoing discussions with EOX on how to generically upload data. Over the following 2-3 weeks, high-level questions were to be formulated to develop meaningful relationships. The analysis was planned to occur on two levels: the first level involved using European datasets (e.g., Copernicus Land Monitoring Service, Climate Change, Atmosphere) to provide an overview for cities at the European level. The second level focused on four identified cities, using local datasets from these cities and combining them with EU data. The primary users of this analysis were intended to be the European Environment Agency (EEA), with ongoing discussions on how to approach individual cities.

UC1 team summarised the technical setup, which included setting up tools for processing 10m and 100m Cloud Optimized GeoTIFF files, connecting to an AWS S3 bucket, uploading the first data to the S3 bucket, and creating a collection to transfer registry data from S3 to EuroDataCUBE.

STATUS FEB 2023 - Urban adaptation to climate change

- ✔
Set up of tools for processing the 10m& 100m Cloud Optimized GeoTIFF files
 - <https://github.com/FAIRiCUBE/uc1-urban-climate/blob/master/notebooks/pre-processing/>
- ✔
Connect to AWS s3 bucket

name	Last modified
s3	53 years ago
uc1-urban-climate	7 days ago
- ✔
Upload first data to the s3 bucket

Simple Process

The following process will help you to enrich the functionality:

 - Store your own data in S3 to use on your own S3 bucket
 - Enrich the dataset with external data from external sources
 - Use the registry API to create index
 - Use the services with your data
- ✂
Create collection -registry data form s3 to EuroDataCUBE

Next week technical meeting with EOX to clarify open questions about the use of own data on EuroDataCube

Figure 2 : Summary of UC1 Technical setup.



3.3.2 UC2 - Biodiversity-Agriculture Nexus

UC2 focused on biodiversity mapping and establishing causal relationships. On the technological side, the team aims to implement deep learning with Rasdaman. There were challenges with converting the local Rasdaman model (C++ library) to PyTorch, as the results were not as expected due to differing data patterns and gaps. Discussion issues highlighted that the model was hardcoded and needed to be more generic, requiring further debugging and development. The team was in touch with the Rasdaman team and FAIRiCUBE PMs to advance these efforts.

3.3.3 UC4 - Building Stock

UC4 presented their primary objective which remained to build an urban stock model progressing from Level of Detail 1 (LoD1) to LoD3. Data analysis steps included estimating building height using 3D city models at LoD1 and gathering more data to achieve LoD2 with roof shapes. The analysis techniques were summarised in the presentation. There were plans for collaborations with cities like Oslo, Vienna, Luxembourg, and Barcelona. An application example discussed by the team involved assessing building properties to classify them and estimate their greenhouse emissions and end-of-life material management.

3.3.4 UC3 - Drosophila Genetics

UC3 objectives included understanding evolution by linking genetic data with environmental data and facilitating the link between genetic data and environmental factors using Drosophila as a model. The current status involved finalising the data pipeline and applying various statistical models.

3.4 WP2 UC synergy monthly workshop (no.2)

The second synergy monthly workshop was held on 7th March 2023. The workshop aimed to provide an update of the data analysis plan and description of expected results. UC2 could not participate.

3.4.1 UC1 - Urban Climate

UC1 data analysis plan was described. Main data sources are European/global and open access. Local data could be sensitive or confidential and not open access, requiring a plan to manage these aspects. The ongoing work involved ingestion into S3, enabling them to import data with EOX. The team anticipated a connection with UC4 regarding building heights, as they shared the same ML expert.

UC1 is identifying 4/5 test case cities: Barcelona (BCN), Vienna, Oslo, and Luxembourg. A professional-looking package was needed to present the project, covering all the UCs approaching that city. Visual content was preferred over text-only descriptions. They planned to create 1 or 2 pages describing each UC, highlighting benefits for cities, and composing a common mail.

3.4.2 UC3 - Drosophila Genetics

UC3 were in the process of consolidating existing global data (raw genomic data management), doubling their data input. Each data point had coordinates and a time of sampling. They were developing a "worm picker," which extracted environmental data stored in Fairicube based on the coordinate position. There was progress in interactions with Rasdaman, with access now available. Local implementation was also progressing.

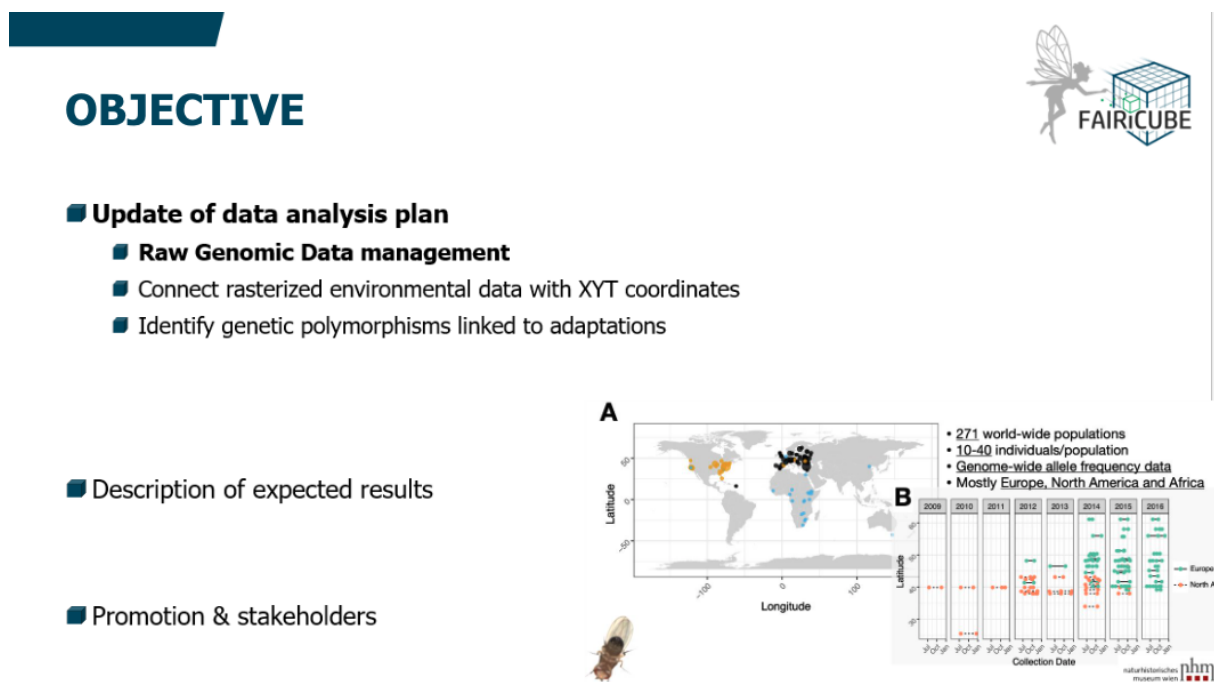


Figure 3 : UC3 Objective summarised.



3.4.3 UC4 - Building Stock

UC4 explain how the team were expanding the applications. Initially, they were too quick to state they wanted to quantify emissions and decided to focus on building stock instead. No further updates or information on stakeholders were provided.

3.5 WP2 UC synergy monthly workshop (no.3)

The third synergy monthly workshop was held on 1st April 2023. The workshop aimed to provide an update of data sources, and update of data analysis and expected results.

3.5.1 UC1 - Urban Climate

UC1 explained that since last meeting, the team spent most of their time setting up the infrastructure. The team identified potential synergies with UC3 and UC4 and decided to discuss these further with UC3 before involving the rest of the group.

UC1 had set up the environment for data cubes. Although there were no European-level stakeholders directly relevant to the execution of their work, UC1 emphasized the need to clarify capabilities within FAIRiCUBE to determine limitations. The team suggested including interlinkages between UCs if stakeholders involved more than one UC. Additionally, they believed having a local partner for city contacts would be advantageous due to language and cultural considerations. They concluded that videos were unnecessary for stakeholder engagement.

3.5.2 UC2 - Biodiversity-Agriculture Nexus

UC2 focused on biodiversity mapping and establishing causal relationships. On the technological side, the team implemented biodiversity data cubes and saw potential synergies with UC3 on occurrence cubes, leveraging their existing experience in this area.

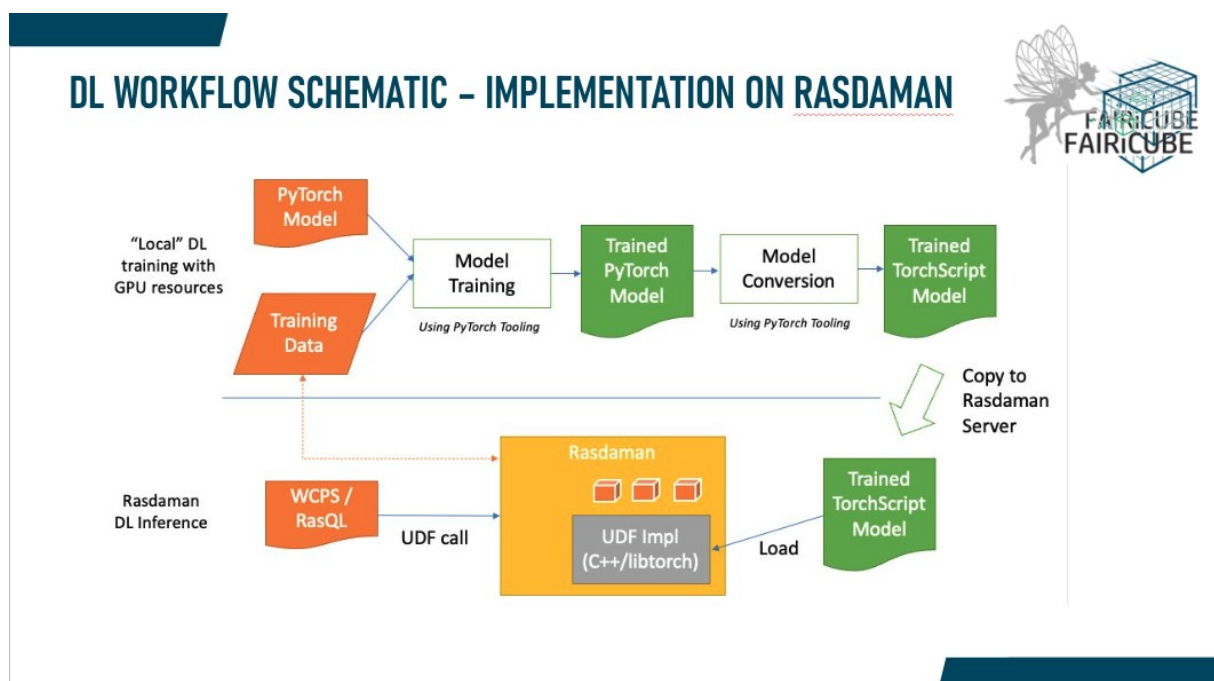


Figure 4 : UC2 workflow schema.



3.5.3 UC3 - Drosophila Genetics

UC3 were reported an increase in their dataset over the past month and had started working on quality control (QC). They were determining the value of the data using heuristics and aimed to provide best practice recommendations for future users. The team also utilized data proximity to impute missing genomic positions in other samples, considering distance and time between samples.

Experts from NILU offered assistance with machine learning to address this issue. UC3 was investigating how climate change influenced genetic variation and biodiversity across European populations. Efforts were ongoing to "cubify" their dataset.

3.5.4 UC4 - Building Stock

UC4 focused on new data in addition to their existing datasets like OSM, Urban Atlas, and elevation data. NILU experts incorporated building material intensity, carbon emission factors, and data from the Tabula project. Data analysis was progressing well. However, they needed to identify sources and methods for obtaining an inventory of building materials to validate their method. UC4 also required support to determine whom to contact to acquire this data.

3.6 WP2 UC synergy monthly workshop (no.4)

The fourth synergy monthly workshop was held on 6th June 2023. The workshop aimed to provide an update of data sources, progress on the UC, update of data analysis and expected results.

3.6.1 UC1 - Urban Climate

UC1 expressed that progress was slower than anticipated, with efforts focused on setting up the infrastructure and addressing various procedural challenges. The team aimed to utilise tabular data and was awaiting a response from EOX to proceed further.

The ingestion procedure was operational, but additional data was required for their machine learning analysis, particularly cluster analysis, which was causing delays. The main data sources were European and global, primarily open access and available in EDC/SentinelHub. Ingestion into S3 was ongoing but blocked due to database issues, and discussions with EOX were in progress.

Synergies were identified with UC4 regarding building heights and with UC3, which could provide *Drosophila* data for their analysis. Data analysis efforts were focused on ingesting more datasets, including land use/cover of city surroundings, climate, and socio-economic data, and selecting appropriate climate models in terms of resolution and time-period. At the local level, the lack of gridded data posed a challenge, with only city centre spatial information available. The climate data being imported was point data, extracted as individual points.

3.6.2 UC2 - Biodiversity-Agriculture Nexus

No updates were given by UC2 in that meeting.

3.6.3 UC3 - *Drosophila* Genetics

UC3 highlighted an upcoming synergy with UC1, focusing on datasets collected near urban areas to compare genetic variation and adaptation signals between rural and urban areas. Additional datasets to be used included land usage, floods, floor sealing, and temperature differences.

The development of the "wormpicker" tool was underway to automate the download of layer data for point coordinates. Challenges persisted with Rasdaman, including server hosting issues, changes in datasets over time, and expiration of security certificates. UC3 team suggested notifications when datasets were permanently ingested and unchanged, while EPSILON noted that the ingestion process was slow, potentially requiring prioritisation of layers.

Updates to the data analysis plan involved identifying genetic polymorphisms linked to adaptations. Expected results included the development of the wormpicker as a standalone script and an imputation pipeline comparing the performance of heuristics/ML. Stakeholders had provided positive feedback on the presented results at the DrosEU conference in Wageningen and the NHM internal conference.

3.6.4 UC4 - Building Stock

UC4 reported the completion of building height estimation and the compilation of results into a report and deliverable. Preparations to move to the cloud were underway, including setting up an Amazon account. The team planned to repeat the building height analysis for another city, likely Oslo or Barcelona, to synchronise it with UC1.

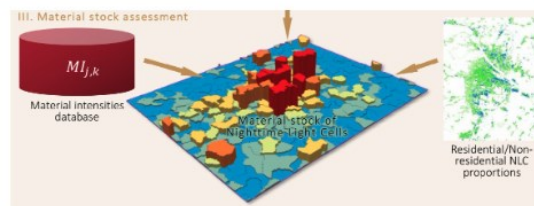
Data analysis aimed to describe the building stock using various parameters such as building age, material intensity, volume, and building type classification using images. However, the lack of a domain expert was slowing the development of an efficient data analysis plan.

Expected results could include many exportable outputs, but defining these outputs was dependent on having a domain expert. It was suggested the creation of a centralized notebook for everyone to contribute, listing datasets explored and their usefulness.

STAKEHOLDERS & OUTREACH SUPPORT



- We need to figure out where and how to get the inventory of building materials (to compare with our mapping)



- May need support on who to contact, who may have such data, etc.
 - Structural building materials (steel, bitumen, roof gravel, ceramic, wool, gypsum, etc.)
 - Components materials (Steel, copper, cast iron, iron, aluminum, plastics, etc.)

Figure 5 : UC4 stakeholder and outreach support.

To sum up, UCs demonstrated progress in data collection, ingestion, and analysis. Challenges included obtaining complete and accurate data, establishing guidelines for data ingestion, and resolving discrepancies in data sources. Synergies between the UCs were actively explored to share knowledge and improve analysis methods, fostering a collaborative approach to achieving their objectives.

3.7 WP2 UC synergy monthly workshop (no.5)

The fifth synergy monthly workshop was held on 4th July 2023. The workshop aimed to discuss the progress and updates of the use cases. Information was shared on data ingestion, validation, and analysis plans. The UC also discussed the need for guidelines on data validation and taxonomic classification.

3.7.1 UC1 - Urban Climate

UC1 exchanged their significant progress, with ingestion procedures functioning well and access to a database secured, allowing for the ingestion of climate data. However, there were challenges due to delays in the database availability. Despite this, obtaining the necessary database marked a considerable achievement. Synergies were identified with Use Case 4 regarding building heights and with Use Case 3 concerning transfer analysis, highlighting opportunities for collaborative advancements.

DATA ANALYSIS

■ Summary of UC data analysis plan

- Main data sources are European and global databases (e.g., CLMS, C3S, Eurostat, JRC), mostly free and open access, partly already available in [EDC/SentinelHub](#)
 - Once the local level is addressed, city and regional data will be added, it is expected that they are often confidential/sensitive and not open
 - Ingestion into S3 ongoing (Urban Atlas 2012, climate data, Eurostat data → database → discussion with EOX pending)
- UC1 is expected to be linked to UC4 on the building heights
 - a link to UC3 is currently set up, exchange concerning urban data that could be used in the Drosophila analysis

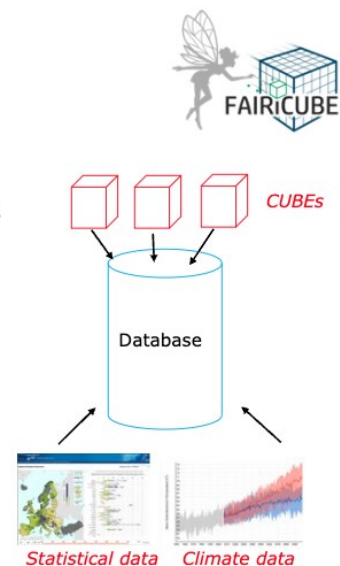


Figure 6 : UC1 data analysis update.

3.7.2 UC2 - Biodiversity-Agriculture Nexus

UC2 explained the ongoing efforts in data exploration and ingestion procedures demonstrated progress. Challenges included waiting for additional data from the NDF database, though a request had already been sent to address this issue. UC2's next steps involved calculating the biodiversity index, determining analysis boundaries, and developing gap-filling methods. These steps aimed to enhance the robustness and completeness of the biodiversity and agriculture data nexus.



3.7.3 UC3 - Drosophila Genetics

UC3 explain how the team saw the creation of 3D building models using various data sources, alongside the acquisition of data on the construction year of buildings. Challenges included the need to consolidate taxonomy data from diverse sources and establish comprehensive guidelines for data ingestion. The next steps involved exploring image processing techniques to create more detailed building models and continuing with data collection efforts. These actions aimed to refine the understanding of building energy performance.

3.7.4 UC4 - Building Stock

UC4 focused on progress made with the acquisition of 3D building models and construction year data. However, challenges persisted in reconciling data discrepancies and establishing connections with other cities. The team planned to use image processing techniques to improve building models and explore methods for estimating energy performance. These steps aimed to enhance the accuracy and utility of building material and energy performance data.

To sum up, the use cases were making substantial progress in data collection, ingestion, and analysis. Key challenges included obtaining complete and accurate data, establishing guidelines for data ingestion, and addressing discrepancies in data sources. Synergies between the use cases were actively explored to share knowledge and improve analysis methods, fostering a collaborative approach to achieving their objectives.



3.8 WP2 UC synergy monthly workshop (no.6)

The sixth (and last) WP2 synergy monthly workshop was held on 3rd October 2023. The workshop focused on potential cross-UC synergies focusing on gap filling. No specific UC presentations were given. The meeting started with a discussion on potential cross-use case synergies, focusing on gap filling. The participants agreed to start looking at data ingestion validation and how to embed it in the ingestion process. They discussed the importance of quantifying the ingestion process and ensuring data integrity and completeness. The participants also mentioned the need to validate social, economic data and incorporate it into the validation process.

The meeting concluded with a discussion on the need for clear and consistent UC descriptions and project updates. It was suggested to allocate time for each use case to present their achievements and updates, followed by a discussion on best practices and future plans. Overall, the meeting was productive in discussing important topics related to data validation and project updates.

3.9 WP2-WP3 UC synergy monthly workshop (no.1)

The first WP2-WP3 synergy monthly workshop was held on 5th December 2023. The workshop aimed to discuss the progress and challenges of UC, focusing on data analysis plans, validation protocols, and the integration of Earth observation data. Use case 5, initially part of UC3 is included.

3.9.1 UC1 - Urban Climate

UC1 provided an update. The team discussed the challenge of filling gaps in socioeconomic datasets and decided against pursuing a gap-filling algorithm due to extensive data gaps. They planned to initiate clustering analysis with the identified datasets.

3.9.2 UC2 - Biodiversity-Agriculture Nexus

UC2's team described how they have continued working on species distribution modeling using the maximum entropy model, refining inputs and environmental variables. They updated the workflow diagram and deliverable D3.2 with progress indications and addressed taxonomy and coordinate issues in GBIF data. Deliverable updates were in progress, with reviews of other deliverables as needed.

3.9.3 UC3 - Drosophila Genetics

UC3 explain success in accessing data via Rasdaman and a user-defined function, with plans to run a complete analysis early in 2024 using temperature data from Rasdaman. Collaboration with Use Case 5 was ongoing to test the "wormpicker" tool, and there were plans to migrate the application to a notebook for hackathon sessions.

3.9.4 UC4 - Building Stock

UC4's team progressed from building height estimation to estimating the energy needed for heating buildings based on classification. They were working on associating energy sources with greenhouse gas emissions. Data loss due to system issues highlighted the importance of using GitHub for syncing and backup.

3.9.5 UC5 – Phytosociological method

UC5's team explained how they aim to validate classical phytosociological methods and propose new approaches to predict vegetation communities. They planned to link vegetation occurrence data with biotic and abiotic Earth observation data, starting with EUNIS habitat data and using temperature data from Rasdaman. The goal was to create data cubes based on habitat types and occurrences and publish species distribution patterns.



DATA ANALYSIS PLAN



- EUNIS Habitats
- Taxa present in the habitat types
- Correct taxonomy (species/synonyms/subsp./varieties) + coordinates from GBIF
 - GBIF cubes of the taxa selected
- Occurrence cubes in Rasdaman with the selected layers
- Using Wormpicker (UC3) tool to get the layers information
- ML:
 - distribution of communities
 - predict locations with favorable conditions for unknown communities similar to those from the occurrence cubes.

Figure 7 : UC5 data analysis plan.

To sum up, all UCs discussed the progress and challenges of UC, focusing on data analysis plans, validation protocols, and the integration of Earth observation data. There was an emphasis on the need for collaboration on the potential for cross-use case synergies, particularly between use cases 3 and 5. Information was shared on data ingestion, validation, and analysis plans. The UC also discussed the need for guidelines on data validation and taxonomic classification.

3.10 WP2-WP3 UC synergy monthly workshop (no.2)

The second WP2-WP3 synergy monthly workshop was held on 6th February 2024. The workshop objective was for use cases to provide updates on their progress and look at potential collaboration and data sharing to enhance their work.

3.10.1 UC1 - Urban Climate

UC1 reported on the challenges with filling gaps in socioeconomic datasets and the decision not to pursue a gap-filling algorithm due to extensive data gaps. They are planning to start clustering analysis with the identified datasets. Additionally, they discussed organising use case seminars to gather policy and decision-maker needs.

3.10.2 UC2 - Biodiversity-Agriculture Nexus

UC2's team continues to refine species distribution modelling inputs for the MaxEnt model and update relevant deliverables. They are focused on cleaning datasets and testing different visualisation tools.

UC2 also plan to hold a use case seminar in March.

3.10.3 UC3 - Drosophila Genetics

UC3's team has managed to access data via Rasdaman and plans to run the entire analysis pipeline using the new surface air temperature dataset. The team continue working with Drosophila genomic data. The team established a pipeline to intersect genomic data with environmental data, such as temperature, and have been comparing different statistical approaches to assess correlations. They are also planning to include urban samples from Vienna and sequence the DNA of these populations.

3.10.4 UC4 - Building Stock

UC4's team has progressed in estimating building heights and energy consumption for heating based on building classifications. They discussed the challenges of obtaining data on energy consumption and building materials. They also plan for statistical analysis and machine learning approaches.

3.10.5 UC5 – Phytosociological method

UC5's team and Use Case 3 are collaborating, with Use Case 5 planning to use the "worm picker" tool from Use Case 3 for their analysis. To sum up, use cases provided updates on their progress and potential collaboration and data sharing. The next meeting date will be scheduled, focusing on continuing discussions on use case progress and deliverable contributions.

3.11 WP2-WP3 UC synergy monthly workshop (no.3)

The third WP2-WP3 synergy monthly workshop was held on 12th March 2024. The workshop objective was for use cases to provide updates on workflow, highlighting latest achievements.

3.11.1 UC1 - Urban Climate

UC1 presented the summary and objectives of Use Case 1, focusing on the integration of machine learning for clustering analysis and data cubes. The team presented detailed workflow starting from data ingestion to visualisation: Special attention was given to the challenges faced with gap filling and local climate data. The creation of dashboards and clustering results and discussed challenges such as database requirements and local climate data acquisition was highlighted.

3.11.2 UC2 - Biodiversity-Agriculture Nexus

UC2's team focused on updating the UC workflow. The workflow involves data processing and modeling, although there are some data limitations. Key achievements include making conceptual decisions and data preprocessing. Data acquisition and processing remain challenges, but some solutions have been identified.

3.11.3 UC3 - Drosophila Genetics

UC3's team presented the updated workflow to highlight the worm picker tool for data retrieval. Achievements include data integration and the development of the analysis pipeline. Challenges involve data gaps and inconsistencies in retrieval. Temporary workarounds and alternative methods for data access have been employed as solutions.

3.11.4 UC4 - Building Stock

UC4's team discussed investigating energy performance and material use in buildings. The workflow requires updating with new icons and streamlined processes. Achievements include data ingestion, building stock modelling, and energy performance estimation. Challenges include data access, VM performance, and data representation. Solutions will be incorporated into the presentation with a focus on practical outcomes.

3.11.5 UC5 – Phytosociological method

UC5's team presented the revised workflow to include new approaches for data retrieval and analysis. Achievements include acquiring occurrence data and preparing for machine learning analysis. Challenges involve large dataset downloads and VM issues. Next steps include finalising machine learning processes and data analysis.



To sum up, meeting focused on presenting updates of UC workflows. UC reviewed and refined the presentations for various use cases, emphasising the need for clarity, consistency, and conciseness in slides and explanations. Participants discussed the integration of new icons, the alignment of workflows.

3.12 WP2-WP3 UC synergy monthly workshop (no.4)

The fourth WP2-WP3 synergy monthly workshop was held on 7th May 2024. The workshop objective was for use cases to provide updates on workflow, highlighting latest achievements.

3.12.1 UC1 - Urban Climate

UC1 presented the summary and objectives of Use Case 1, focusing on the integration of machine learning for clustering analysis and data cubes. The team presented detailed workflow starting from data ingestion to visualisation: Special attention was given to the challenges faced with gap filling and local climate data. The creation of dashboards and clustering results and discussed challenges such as database requirements and local climate data acquisition was highlighted.

3.12.2 UC2 - Biodiversity-Agriculture Nexus

UC2's team focused on updating the UC workflow. The workflow involves data processing and modelling, although there are some data limitations. Key achievements include making conceptual decisions and data preprocessing. Data acquisition and processing remain challenges, but some solutions have been identified.

3.12.3 UC3 - Drosophila Genetics

UC3's team presented the updated workflow to highlight the worm picker tool for data retrieval. Achievements include data integration and the development of the analysis pipeline. Challenges involve data gaps and inconsistencies in retrieval. Temporary workarounds and alternative methods for data access have been employed as solutions.

3.12.4 UC4 - Building Stock

UC4's team discussed investigating energy performance and material use in buildings. The workflow requires updating with new icons and streamlined processes. Achievements include data ingestion, building stock modelling, and energy performance estimation. Challenges include data access, VM performance, and data representation. Solutions will be incorporated into the presentation with a focus on practical outcomes.

3.12.5 UC5 – Phytosociological method

UC5's team presented the revised workflow to include new approaches for data retrieval and analysis. Achievements include acquiring occurrence data and preparing for machine learning analysis. Challenges involve large dataset downloads and VM issues. Next steps include finalising machine learning processes and data analysis.

To sum up, meeting focused on presenting updates of UC workflows. UC reviewed and refined the presentations for various use cases, emphasising the need for clarity, consistency, and conciseness in slides and explanations. Participants discussed the integration of new icons, the alignment of workflows.

3.13 WP2-WP3 UC synergy monthly workshop (no.5)

The fifth WP2-WP3 synergy monthly workshop was held on 4th June 2024. The workshop objective was for use cases to provide detailed updates on the work, and progress done up to date. Despite not presented below, the presentation started with UC5 to UC1.

3.13.1 UC1 - Urban Climate

UC1 presented an update on their progress. The team is formalising collaboration with Luxembourg City to analyse the impact of land use, green spaces, and other factors on urban climate, particularly temperature. Luxembourg City will contribute data from seven weather stations to this effort. Concurrently, the team continues its collaboration with UC3 in Vienna, aiming to build a data cube for urban analysis and possibly replicate the analysis done in Luxembourg. Emphasis has been placed on transforming local data into a gridded format suitable for the data cube, starting with data resolutions of 100 meters and 10 meters, and initially working with coarser data such as climate data at one-kilometre resolution.

Stakeholder engagement remains a key focus, with ongoing discussions with Luxembourg City and plans to reach out to the Vienna City administration. Potential workshops and bilateral meetings are being considered to engage stakeholders and present findings. Although data analysis has not yet commenced, the team is concentrating on data preprocessing, planning to use machine learning approaches and ensuring quality checks. They are also exploring synergies with other UCs, particularly in Vienna, and considering workshops to demonstrate data cube analysis to a broader audience. The team emphasised the importance of collaboration, stakeholder engagement, and thorough data preprocessing as essential steps before beginning detailed analysis.

3.13.2 UC2 - Biodiversity-Agriculture Nexus

UC2 explored the significant progress made in data ingestion and preprocessing despite facing some technical issues in the registration process. The team has completed metadata collection and editing, extended the study area, and aligned all datasets.

UC2's workflow and objectives have remained consistent, focusing on specific task breakdowns and reviewing relevant indices. The team has identified synergies with other UCs, particularly in data processing workflows, underscore the collaborative potential across the project. The use of common resources like GBIF and climate data, despite geographical differences, has been noted, with no new data needs identified recently.

The team is using both Rasdaman and EOX platforms, prioritising EOX for ML due to its superior hardware resources and is considering ingesting data on both platforms if feasible. While validation of ingested data is yet to be conducted, plans are in place to initiate this crucial step in the coming weeks.

3.13.3 UC3 - Drosophila Genetics

UC3's team provided an update on the UC, highlighting the steady progress in data analysis primarily performed on the EOX's platform. The team has improved access to Sentinel hub data, though they face challenges dealing with NA values in specific time ranges and the need for approximating missing data. They also noted a need for more expertise in Earth observation data and establishing standards for genetic data. To address these issues, the workflow plan has been updated to integrate genomic data from Drosophila and environmental data from Vienna. UC3 is using the EOX hub for data analysis and processing, which has been crucial for their progress.

Collaboration and data sharing were significant themes in the update. The team identified a need for more shared data in common S3 buckets, as many use cases are accessing similar data repeatedly. Ongoing collaboration with UC1 on the Vienna data catalogue and progress in the citizen science aspect, with 95 participants collecting fly samples across Vienna.

3.13.4 UC4 - Building Stock

UC4's team presented updates, focusing on validating energy performance calculations against self-declared assessments from property owners. A key aspect of their work involves validation and model comparison, where they have compared their energy performance model with self-declared energy performance data from property owners in Norway. The findings revealed that their model generally overestimates energy performance, and they identified that the model's sensitivity to heated floor area might be causing these discrepancies. To address this, the team is currently conducting a sensitivity analysis to understand the impact of heated floor area on energy performance estimates, documenting their hypothesis testing, and preparing for further validation.

Looking ahead, the team plans to replicate the analysis for the city of Vienna, similar to what was done for Oslo. This will require accessing self-declared energy performance data for Vienna to ensure accurate validation. Throughout their work, the team continues to use the EOX platform for data processing and analysis, ensuring consistency and reliability in their methodologies.

3.13.5 UC5 – Phytosociological method

UC5 provided detail of the progress and ongoing efforts. The workflow remains aligned with the original plan, utilising data from the Global Biodiversity Information Facility (GBIF) to create GBIF cubes. Environmental data is then integrated into these cubes to generate occurrence cubes, facilitating the visualization and filtering of species distribution and environmental factors. To enhance data accuracy, a filter is applied to remove incorrect occurrences from the GBIF data. UC5's team also mentioned plans to download ERA5 monthly average reanalysis data, which includes temperature and precipitation information, and store it in the S3 common bucket for accessibility by other use cases. Additionally, there is potential to use the Eunice raster map of Europe, pending permission for sharing with other use cases.



Data processing and validation efforts are ongoing, with a particular focus on establishing a ML pipeline. The team is currently utilising the EOX platform for their work. Overall, the team emphasised the steady progress in data processing, the importance of further validation, and the planning required for ML integration.

To sum up, the meeting discussed the progress and challenges of various UCs, including data pre-processing, validation, and stakeholder engagement, with a focus on synergies between different UCs (including piloting cities). Specific updates included the use of EOX services, the need for additional data ingestion, and potential collaborations with local authorities. Plans for a training session in October and the importance of documenting computational resources were also highlighted.



4 Conclusion

The series of workshops organised between February 2023 and June 2024 have assisted the project in improving knowledge across Use Cases. The workshops successfully facilitated collaboration among various UCs involved in the FAIRiCUBE. They provided a platform for teams to share objectives, progress, and methodologies.

Knowledge sharing involved the usage of data infrastructure, implementation of cloud-based solutions, data cubes, and advanced visualization tools. This infrastructure supported complex data analysis and interpretation across UCs, enhancing overall project capabilities.

The workshops assisted in the identification of synergies between UCs working on related topics, such as urban climate and building stock analysis. These synergies enabled shared expertise and resources, optimising project outcomes and efficiency in data handling and analysis.

The workshops served as a platform to discuss and address technical challenges specific to each UC, including data ingestion, model implementation, and data validation. Solutions and best practices were shared among participants, ensuring robustness in data handling and analytical processes.

Engagement with stakeholders, including city administrations, was included in the discussion. This engagement facilitated collaboration on specific project aspects like urban climate analysis and building stock management.

Advances in machine learning integration and visualisation tools (e.g., data cubes, dashboards) were showcased. These should support data analysis techniques and facilitate the interpretation and presentation of complex datasets, contributing to informed decision-making within the project.

Overall, the series of workshops from February 2023 to June 2024 assisted collaborative efforts, improved technical capabilities, and enhanced the overall effectiveness of the project across all UCs. These conclusions highlight the workshops' role in achieving project objectives and while creating a synergistic environment for knowledge sharing.