

REGACE – Crop Responsive Greenhouse Agrivoltaics System with CO2 Enrichment for Higher Yields

Project Deliverable Report

DELIVERABLE 9.4 – DATA MANAGEMENT PLAN

WORK PACKAGE NUMBER: WP9

WORK PACKAGE TITLE: PROJECT MANAGEMENT

TYPE: REPORT

AUTHORS: MARIAM AMER

REGACE Action Information	
Action full title	Crop Responsive Greenhouse Agrivoltaics System with CO2 Enrichment for Higher Yields
Action acronym	REGACE
Grant agreement number	101096056
Project coordinator	Prof. Ibrahim Yehia
Project start date and duration	1 February 2023, 36 months
Project website	https://regaceproject.com/

Document Information

Deliverable Information	
Work package number	9
Work package title	Project Management
Deliverable number	9.4
Deliverable title	Data Management Plan
Description	This deliverable outlines the framework for handling, storing, and sharing data generated throughout the project.
Lead beneficiary	Alzahrawy
Lead Author(s)	Mariam Amer
Contributor(s)	All partner organisations
Revision number	v2.0
Revision Date	2024/01/31
Status (Final (F), Draft (D), Revised Draft (RV))	F - Final
Dissemination level (Public (PU), Restricted to other program participants (PP), Restricted to group speci-fied by consortium (RE), Confidential for consortium members only (CO))	PU - Public



Document History			
Revision	Date	Modification	Author
v1.0	2023/05/29	Initial version submitted early in project	Michael Eilan
v1.1	2024/11/30	Updates and revisions after 1st reporting period	Mariam Amer
v2.0	2025/01/31	Second version for submission	Mariam Amer

Approvals				
	Name	Organisation	Date	Signature (initials)
Coordinator	Prof. Ibrahim Yehia	Alzahrawy Society	2025/01/31	IY
WP Leaders	Raymond Slaughter	Agora Partners	2025/01/31	RS

Table of Contents

EXECUTIVE SUMMARY	4
LIST OF ABBREVIATIONS	5
CHAPTER 1 – INTRODUCTION	6
CHAPTER 2 – DATA MANAGEMENT OUTLINE AND ETL PROCESS (EXTRACT, TRANSFORM, LOAD)	6
CHAPTER 3 – DATA GOVERNANCE AND RACI MATRIX (RESPONSIBLE, ACCOUNTABLE, CONSULTED, INFORMED)	10
CHAPTER 5 – DATA STORAGE AND ACCESS	11
CHAPTER 7 – OPEN SCIENCE AND FAIR PRINCIPLES	14
CHAPTER 8 – OTHER RESEARCH OUTPUTS	16
CHAPTER 9 – ALLOCATION OF RESOURCES	17
CHAPTER 10 – QUESTIONS TO INFORM CONTINUOUS IMPROVEMENT	17
CONCLUSION	17



Executive Summary

The REGACE Data Management Plan (DMP) outlines the framework for handling, storing, and sharing data generated throughout the project. REGACE, which focuses on developing an advanced agrivoltaic system with CO₂ enrichment, integrates agricultural innovation with renewable energy production. By systematically managing data, the project ensures that research findings are transparent, accessible, and aligned with industry best practices.

This document establishes the principles governing data collection, processing, and security, ensuring compliance with Open Science and FAIR (Findable, Accessible, Interoperable, and Reusable) guidelines. It details how numerical and qualitative data—ranging from sensor readings and crop yield measurements to farmer interviews and digital modelling outputs—will be handled across multiple locations in Israel, Italy, Germany, Greece, and Austria. A structured governance framework is in place, assigning clear responsibilities to project partners and maintaining accountability throughout the data lifecycle.

Storage solutions leverage cloud-based platforms with automated backups, ensuring data integrity and accessibility for research partners. Proprietary data, particularly concerning the electrical performance of the tracking system, is protected while balancing the project's commitment to knowledge-sharing and open science.

The long-term sustainability of the project's data is also a key focus, with a structured plan for post-project storage and controlled access to valuable research outputs. By implementing a rigorous yet practical data management strategy, REGACE supports scientific collaboration while protecting intellectual property, creating a foundation for future research in agrivoltaics, renewable energy, and climate-resilient farming systems.

List of Abbreviations

Full Name	Abbreviation
Grant Agreement	GA
Consortium Agreement	CA
Project Officer	PO
Data Management Plan	DMP
Project Coordination Team	PCT
Description of Action	DoA
Work Package	WP
Consortium Members and Affiliates	
Alzahrawy Society	AZS
Fattoria Solidale Del Circeo Cooperativa Sociale Di Produzione E Lavoro	FSC
Università degli Studi di Roma Tor Vergata	UNITOV
Panepistimio Thessalias	UTH
Humboldt-Universitaet zu Berlin	HU
Tel Aviv University	TAU
Universitaet fuer Bodenkultur Wien	BOKU
Bio-Gaertnerei Watzkendorf Gmbh	BW
Interteam Ltd	IT
Agora P.S.v.D. Ltd	AP
Trisolar - Innowadi Group Ltd	TS
Timelex	TL



Chapter 1 – Introduction

Effective data management is critical to ensuring the integrity, accessibility, and usability of the data collected in our project. This plan outlines the processes and policies for managing data generated from our research, which focuses on investigating the use of CO₂ enrichment to increase energy yields in a novel responsive tracking system developed by TriSolar, one of the consortium's members. Another major target is to investigate ways in which to incentivize farmers to adopt this new technology.

In the course of the project, the responsive tracking system will be installed in six locations in Israel, Italy, Germany, Greece, and Austria. Numerate data on the performance of the system, with and without CO₂ enrichment will be collected in terms of electrical yield, crop yield, and microclimate inside the greenhouses. Non-numerate data from the greenhouses will include images of various kinds, 3D plans, and others. In addition to this progress on the project to introduce the technology to farmers will be collected in the forms of text and video files.

This document details the lifecycle of the data, from collection and storage to sharing and long-term preservation. The plan aligns with the principles of FAIR (Findable, Accessible, Interoperable, and Reusable) data management to maximize the value and impact of our research. Furthermore, we commit to open-access publishing, ensuring that data is accessible to the broader scientific and professional community, fostering transparency, collaboration, and innovation.

The primary objectives of this DMP are to:

1. Describe the methods for data collection, storage, and backup.
2. Ensure proper documentation and organization for ease of understanding and use.
3. Provide a framework for use, long-term preservation and future reuse of datasets.

By adhering to these guidelines, we aim to maintain high standards of data quality, protect data integrity, and contribute to the global exchange of knowledge in the fields of agriculture, renewable energy, and environmental science.

Chapter 2 – Data Management Outline and ETL Process (Extract, Transform, Load)

Chapter 2.1 – Previous Data

Data from previous experiments by AZS and TS using the same technology without CO₂ enrichment is relevant to this project but will not be used because new panels were designed for

this project and different panel specifications render the results incomparable. However, lessons learned from previous research have been applied to designing the experiments in REGACE.

Chapter 2.2 – Types and Formats

Data by WP:

Work Package (WP)	WP lead	Data description
WP1 Potentials of the synergies of innovative and intelligent agrivoltaic systems	UTH	Mixed data in CSV and written reports in Microsoft Word
WP2 Implementation	TS	Plans in DWG or SLDDRAW, component list in CSV installation guides in Microsoft Word
WP 3 Testing	BOKU	Data in CSV and XLSX, images in JPEG, reports in Microsoft Word
WP 4 Modelling & Digital Twins	UR	Data in CSV, reports in Microsoft Word
WP 5 Validation	AZS	Data in CSV and XSLX, reports in Microsoft Word
WP 6 Sustainability (Environmental, economic, and social)	UR	Data in CSV, reports in Microsoft Word, Video in MP4
WP 7 Innovation Management & Business Planning	TS	Reports in Microsoft Word
WP 8 Dissemination	IT	Reports in Microsoft Word. Images in JPEG Video in MP4
WP 9 Project Management	AZS	Reports in Microsoft Word

Chapter 2.3 – Measurements

Data Management Overview

This section outlines the data types, formats, and measurement criteria for different aspects of the project, ensuring clarity and consistency in data handling and reporting.

1. Electricity



Format: CSV files; images of panels in conventional formats; thermal camera images.

Data Types:

- Voltage, current, current-voltage curves.
- Power generated and extrapolated values (e.g., efficiencies, fill factors).
- PV module incident irradiance and temperatures.
- Visual records of panels and their temperature distributions (infrared camera).

2. Horticultural Measurements

Format: CSV files; images of fruit.

Data Types:

- **Quantity Parameters:**
 1. Weight of fruit.
 2. Number of fruits and flowers per plant at harvest.
- **Quality Parameters:**
 1. Fruit length and diameter.
 2. Fruit shape index.
 3. Fruit firmness.*
 4. Fruit colour analysis.*
 5. Titratable acidity (TA).*
 6. Chlorophyll and total carotenoids.
 7. Total phenolics.
 8. Lycopene content (e.g., for tomatoes).
 9. Ascorbic acid.
 10. Antioxidative capacity.
 11. Sugar and acid content (HPLC).
- **Lettuce, Cucumber, and Basil Measurements:**
 1. **At Harvest:**
 - Plant height and diameter.
 - Separation into leaves, stems, and roots (if possible).
 - Fresh mass, dry matter, and water content.
 - Leaf colour analysis.
 - Chlorophyll and total carotenoids.
 - Total phenolics, ascorbic acid, antioxidative capacity.
 - Sugar and acid content (HPLC).
 - Leaf nitrate content.
 2. **Growth Analysis Parameters** (weekly for selected plants):
 - Plant height, stem diameter, canopy diameter.
 - Leaf number, leaf area, specific leaf area (SLA). SLA = amount of leaf area per unit of leaf dry mass [m² kg⁻¹]**

- The fraction of total plant biomass that is allocated to leaves, leaf mass fraction (LMF) = leaf dry weight per plant dry weight [kgkg⁻¹]**.
- Leaf area ratio (LAR) = SLA x LMF [m² kg⁻¹]**.
- Net assimilation rate (NAR) = net increment in plant dry mass per unit of mean leaf area during the experiment [kg m⁻² days⁻¹].
- Absolute growth rate (AGR) = (DM₂-DM₁)/(t₂-t₁), where DM₂ and DM₁ are the mass of the plant at time t₂ and t₁, respectively [kg days⁻¹].
- Relative growth rate (RGR) = the dry mass difference at the beginning (DM₁) and end (DM₂) assuming an exponential growth of the plants:
RGR = (Ln(DM₂) – Ln(DM₁))/(t₂-t₁) ~? LAR x NAR [kg m⁻² days⁻¹].

* Subjective measurement by grower.

** Values for calculation measured only at initial and/or final harvest.

3. Gas Exchange Measurements

Format: Units in Vienna for A/Ci and light curve measurements using CIRAS-2.

Measured Parameters:

- Substomatal CO₂ concentration (ppm).
- Transpiration rate (mmol m⁻² s⁻¹).
- Stomatal conductance (mmol m⁻² s⁻¹).
- Net photosynthetic rate (μmol m⁻² s⁻¹).
- Vapor pressure deficit (mbar).

4. Microclimate Measurements

Format: CSV files.

Data Types:

- Light (watts per square meter).
- Relative humidity (%).
- Temperature (°C).
- CO₂ levels (ppm).

5. Greenhouse Plans

Format: DWG or SLDDRW files.

Data Types:

- 3D models for greenhouse assembly planning.

6. Work with Farmers

Format: MP4 videos; Microsoft Word documents.

Data Types:

- Videos of meetings with farmers.
- Written reports.

Chapter 2.3 – Data Generation

The overarching aim of REGACE is to examine the hypothesis that CO₂ enrichment can enhance electrical yields in greenhouse agrivoltaics. However, the data generated in the project will serve other purposes:

- Scientific papers by members of the consortium.
- The work on introducing new technologies to farmers will be especially useful at a time of rapid climate change when farmers will have to adapt new approaches very quickly.

Chapter 3 – Data Governance and RACI Matrix (Responsible, Accountable, Consulted, Informed)

The RACI matrix for the DMP clearly delineates roles and responsibilities to ensure effective collaboration and accountability. **AZS** is designated as the **Responsible** party for data management and storage, overseeing data quality assurance, backups, sharing, and archiving to maintain integrity and accessibility. **AZS** is the **Accountable** entity, ensuring all tasks are completed to the required standards and that the overall data management aligns with organizational goals. Data collection is performed by the **WP Leaders and their teams**, who are also **Responsible** for gathering accurate and complete data from relevant sources. **WP Leaders** are **Consulted** in monthly meetings, providing expertise and feedback on data handling to ensure alignment with project objectives. Finally, the **Coordination and Management Team** is **Informed** during weekly meetings to maintain oversight, facilitate communication, and ensure the plan is implemented smoothly. This structured approach ensures that responsibilities are clearly assigned, and all stakeholders are appropriately involved.

Chapter 5 – Data Storage and Access

Chapter 5.1 – Data Types

All locations with greenhouses have installed standardized microclimate sensors that continuously feed data into the tracking system website created and hosted by AZS. This website is accessible to all WP leaders and provides graphs displaying the trends in various measurements per time. It also describes map sensors describing the location and nature of the installed sensors for each location.

Other types of data include inverter data, installed in AZS, FSC, UTH, and will be installed in HU soon; local sensor data, comprising of sensors that are not automatically standardized with other locations (in AZS, UTH, HU, BOKU, BW, and shortly in FSC); crop data, as measured throughout the planting and harvesting seasons; modelling data as per WP4, described in deliverable 4.1; and single-PV module monitoring system, which provides information on the temperature, light, and electric output on a single-PV module scale, rather than as a system, each location has their own.

The table below summarises the types of data collected and their formats:

Data Source	Upload Method	Frequency	Responsibility	Type of Data	File Format
Tracking system website	Automatic	Daily	System automated process	Microclimate sensor data (REGACE): light (solar radiation), temperature, humidity, CO ₂ , PV modules (light above & below, temperature, inclinometer)	.csv, https://www.pitec.hgh.com/ (manually exported as .xlsx)
Inverter data	Manual	As needed	AZS	Voltage, current, current-voltage curves, power generated, extrapolations (efficiencies, fill factors), PV module incident irradiance, PV module temperatures	https://www.semportal.com (exported as .xls)
Local sensor data (independent of tracking system website)	Manual	Weekly	Local group	Microclimate sensor data (local): light, temperature, humidity, CO ₂	.xlsx
Crop data	Manual	As needed	Local group	Crop growth metrics, planting records	.xlsx
Modelling data	Manual	As needed	UNITOV	Simulation and modelling outputs	.csv .xlsx
Single-PV module monitoring system	Manual	Weekly	Locations with single-PV systems	PV electric data: voltage, current, current-voltage curves, power generated, extrapolations (efficiencies, fill factors), PV module incident	.csv

				irradiance, PV module temperatures	
--	--	--	--	------------------------------------	--

A comprehensive list of sensors is available to all partners, as are maps describing the locations of each sensor, these are updated as necessary.

Chapter 5.2 – Repository

The data is hosted in a Google Drive space leased by REGACE partner TriSolar Innowadi Group Ltd. It is agreed between the coordinator, the Alzahrawy Society, and TriSolar to host the REGACE data free of charge for the duration of the project plus one year after the end of the project i.e., January 31, 2027. All WP leaders have access and editing permissions via link.

Each WP leader is responsible for analysing the relevant data in accordance with their specific work package and uploading it to their folder within Google Drive.

The Google Drive folder also contains a dedicated section for storing reports generated by WP leaders. Access to this centralized Google Drive is granted to all WP leaders, who have viewing and editing privileges. To ensure consistency, all WP leaders have been trained in the appropriate use of the folder and will continue to receive updated instructions by email as data management policies evolve. Files should not be deleted unless absolutely necessary and with agreement from all parties involved.

To ensure data security, the data management team at AZS backs up the Google Drive folder to a separate backup Google Drive folder, which is accessible only to the data management team. Additionally, tracking system data from the website is automatically backed up daily to AZS servers. This multi-tiered approach ensures the integrity, availability, and security of critical data throughout the project lifecycle.

Chapter 5.3 – Data Size

WP1 Coordination and Management Reports, manuals, workshop and consortium presentations, recordings, pictures: up to 7GB in 3 years.

WP2 Implementation Documentation of installation in schematic and picture form: up to 2MB.

WP3 Testing Measurements are to be uploaded to Google Drive on a daily or weekly basis

(depending on the type of data) during ongoing experiments or seasons, at the conclusion of which a summarising document can replace or be added to the existing files: up to 1GB.

WP4 Modelling The data acquired by the different instruments and the data processed in the digital twins will represent the main output of this WP in terms of data. The expected volume is approximately 100GB. However, during computational analysis, the amount of temporary data can reach 1.5TB in the local computers.

WP5 Validation Manual and evaluation reports in PDF form, analysis of modelling versus real data comparison: expected up to 1GB.

WP6 Sustainability (Environmental, Economic, Social)

Environmental: up to 3MB.

Social: Video – 200 hours of MP4 video at 1.4 GB per hour = 280GB.

WP7 Innovation management & Business Planning Business case: 2MB.

WP8 Dissemination & Visibility Graphics, outreach and communication plans, networking and conference materials: at least 2GB.

WP9 Project Management: at least 10MB.

Total data needs: approximately 400GB.

Our Google Drive currently allows for up to 2TB of data, with 5.36GB taken up thus far, and more space expected to be taken up as experiments conclude and their results and analyses uploaded. Another source of variation for these estimates is the number of hours of video, both from recordings of meetings, conferences, workshops, and from WP6. Our backup Google Drive allows for up to 200GB of storage, the usage of both will be monitored closely by the data management team and will be adjusted should the need arise.

Chapter 6 – Data Privacy and Security

The data will be stored on a Google Drive server and the security will be as good that that afforded by Google to commercial customers.

Furthermore, no personal data is stored by REGACE.

Chapter 7 – Open Science and FAIR Principles

The consortium is committed to implementing the mandatory open science practices as outlined

in Horizon Europe. It will ensure that all peer-reviewed publications resulting from the research are immediately available in open access and deposited in trusted repositories such as OpenAIRE. The project will ensure the reproducibility of its research outputs, including documentation of methodologies and tools to facilitate reuse and dissemination.

All open data and publications produced in REGACE (open REGACE results) will be identifiable and locatable by means of a persistent Uniform Resource Locator (URL). If possible, open REGACE results will be assigned a Digital Object Identifier (DOI) in order to make content easily and uniquely citable. Thereby, REGACE relies on external services, since DOIs can only be assigned by DOI registrants through a DOI registration agency so DOI identification will probably only be available for papers published based on REGACE results.

Chapter 7.1 – Metadata

Metadata will be attached to the header of each document uploaded to the REGACE repository (see Chapter 5.2- Repository). Our priority is the members of the REGACE consortium who will need the information for scientific publications, and then other researchers who will be allowed access to some of the data upon request, with permissions determined on a case-by-case basis.

Chapter 7.2 – Metadata Fields

Field Name	Notes
File Identifier	A one-use file number ascribed by the Google script described above
File name*	See below for file naming conventions
Work Package	WP number
Institution	Following abbreviations: UTH, BOKU, HU, TAU, FSC, UNITOV1, UNITOV2, BW, TS, IT, TL— UNITOV 1 refers to solar energy group, UNITOV 2 refers to sociology group
Upload person	name
File type*	See list below for options. Use of the provided acronyms
Greenhouse number	All experimental partners with more than one greenhouse need to number them consistently and include the data in a text file to be uploaded
File date range	From dd/mm/yy to dd/mm/yy
Keywords	Up to 3
Description	Short text

File format	CSV, MS WORD, DWG MP4, etc
--------------------	----------------------------

Description of file types:

File type	Function	Acronym
Electricity files from greenhouse	All files with data on voltage, amperage, watts, and resistance	ELEC
Microclimate files from greenhouse	All files in irradiance, humidity, CO2 levels and temperature	MICRO
Reports	All written reports in MS Word	REP
Images	All images in various formats	IMG
Video files	All video files	VID

File Naming conventions:

{upload institution}-{file type}-{upload date DD-MM-YYYY}

The REGACE metadata will be updated to include keywords and be organized in a manner conducive to efficient search and harvesting.

The metadata will be made available in a Creative Commons public domain repository as per the Grant Agreement. However, the metadata will not necessarily link to the datasets themselves as specified in Chapter 7.3 – Data Accessibility.

The level of security around the data is as high as Google affords its commercial customers.

Chapter 7.3 – Data Accessibility

Not all the data generated by REGACE will be freely accessible because the electrical performance of the responsive tracking system is proprietary information of TriSolar- protected under the GA and the Background of the CA.

Electrical data will be published during the course of the project, but this will be done following a consultation between the researchers and the stakeholders.

After a certain embargo on release of scientific results -- to be agreed upon by the researchers in REGACE -- the data sets will be released, on request, to qualified scientific researchers.

External researchers may request access to the data by contacting the corresponding author in relevant publications or by reaching out to AZS.

Chapter 7.4 – Increase Data Reuse

In order to facilitate analyses in an error-free manner, readme files will be added to all the directories with numerate data. Data not protected under the IPR issues outlined above will be freely available after the project’s scientists have published the relevant articles. The readme files will enable reuse of the data and will also document the provenance of data sources. The first tier of quality assurance will be provided by a relatively simple script that seeks major deviations in reading of numerate data. The second tier will be in monthly manual review of uploaded files.

Chapter 8 – Other Research Outputs

The main non-numerate research output from REGACE is around 200 hours of recorded video. These records are from a highly structured process including open interviews, focus groups, and world café sessions that will be held in the project’s greenhouses. All this work is intended to a) hear greenhouse farmers’ attitudes towards agrivoltaics and b) hear how they think it should be deployed. The subsequent sentiment analysis will be of great value beyond the confines of the REGACE project. At a time of accelerated climate change, farmers will have to rapidly adapt new technologies to ensure food security. The structured approach that will be developed in REGACE could be a resource of huge importance in the next few years.

Chapter 9 – Allocation of resources

The storage space for REGACE data is being provided free of charge by REGACE partner TriSolar. The coordinator, the AZS, is responsible for managing the data. The data will be stored in the TriSolar space for one year until after the end of the project. After that, partners or other bodies who might want to host it will be invited to do so.

Chapter 10 – Questions to inform continuous improvement

These will be examined on an iterative basis, as more results are recorded and analysed.

- Are we looking at the right metrics for decision-making?
- Is the analysis measuring meaningful, actionable outcomes?
- What is missing from the data? Ensure analysis is not skewed by data gaps.
- How do we minimize bias in analysis?
- Impact vs feasibility: What is causing the impact?

Conclusion

The REGACE project represents a significant advancement in the integration of agrivoltaic technology, CO₂ enrichment, and data-driven agricultural optimization. Through a structured data management approach, the project ensures that research data is collected, stored, and analyzed in a way that maintains integrity, accessibility, and long-term usability. By integrating CO₂ enrichment with agrivoltaic systems, this initiative will contribute to advancing agricultural productivity while optimizing energy generation. The data collected throughout the project—from microclimate sensors, crop performance metrics, and engagement with farmers—is being systematically stored, governed, and analyzed following Open Science and FAIR principles, ensuring its long-term usability and impact.

This DMP provides a clear framework for the collection, organization, and preservation of project data, supporting the broader scientific and agricultural communities. The implementation of standardized storage solutions, automated backups, and structured governance mechanisms allows for effective collaboration between partners while safeguarding proprietary information where necessary. Over the course of the project, careful monitoring and continuous improvement of data processes will allow us to refine methodologies, strengthen the reliability of results, and ensure that all outputs align with the project’s objectives. Furthermore, ensuring secure and well-structured access to data remains a key priority as the project moves forward.

The outcomes of this project will extend beyond its immediate duration, offering valuable insights for researchers, policymakers, and industry stakeholders. The methodologies developed here provide a model for future agrivoltaic initiatives, supporting the global transition to more sustainable and efficient agricultural systems. Through rigorous data stewardship, collaborative engagement, and technological innovation, REGACE is poised to drive meaningful advancements in the fields of agriculture, renewable energy, and climate adaptation.

