

TwinERGY Integrated Platform – Beta Release

D5.4

February 2022



Deliverable

PROJECT ACRONYM	GRANT AGREEMENT #	PROJECT TITLE
TWINERGY	957736	Intelligent interconnection of prosumers
		in positive energy communities with twins
		of things for digital energy markets

DELIVERABLE REFERENCE NUMBER AND TITLE

D5.4 TwinERGY Integrated Platform – Beta Release

Revision: v1.0

AUTHORS

Tasos Tsitsanis	Marios Phinikettos		
Suite5	Suite5		



Funded by the Horizon 2020 programme of the European Union **Grant Agreement No 957736**

DISSEMINATION LEVEL

- P Public
 - C Confidential, only for members of the consortium and the Commission Services



Version History

REVISION	DATE	AUTHORS	ORGANISATION	DESCRIPTION
v0.1	08/03/2022	Tasos Tsitsanis Marios Phinikettos	Suite5	Ready for review version
v0.2	11/03/2022	Ana Isabel Martinez Garcia Moises Anton Garcia	ETRA	Minor changes and corrections
v0.3	11/03/2022	Konstantinos Kotsalos Apostolos Kapetanios	ED	Minor changes and corrections
v0.4	15/03/2022	Axel Balke Fynn Christian Bollhofer	TH-OWL	Minor changes and corrections
v0.5	16/03/2022	Tasos Tsitsanis Marios Phinikettos	Suite5	Consolidated version
v0.6	16/03/2022	Tasos Tsitsanis Marios Phinikettos	Suite5	Final version submitted to the Project Coordinator
v1.0	17/03/2022	Tasos Tsitsanis Marios Phinikettos	Suite5	Draft submitted to EC by the PC



Statement of Originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.



Legal Disclaimer

The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose. The above referenced authors will have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law. The sole responsibility for the content of this publication lies with the authors. It does not necessarily reflect the opinion of the European Union. Neither the CINEA nor the European Commission is responsible for any use that may be made of the information contained therein. © 2021 by TwinERGY Consortium

Executive Summary

The Beta Release of the TwinERGY Integrated Platform is presented in Deliverable 5.4, along with a description of the platform's services (Data Collection, Data Security, Data Storage, Platform Management). In this document, all services and their corresponding functionalities are detailed, using as a basis the TwinERGY conceptual architecture, as introduced in D5.3, in order to inform the TwinERGY Core Data Management Platform (CDMP) users about the range of functionalities that they are enabled with.

Deliverable D5.4 focuses on Deliverable D5.3's [1] research, which detailed the TwinERGY CDMP Services. This deliverable expands on platform functions per service, addressing the Data Collection Service's requirements for data ingestion, data mapping, data curation, and metadata editing. The Data Security Service also includes access policy controlling in order to prevent unauthorized data access. The Data Storage Service includes both data store and the CIM store, as well as data backup and recovery mechanisms. Finally, the Platform Management Service includes identity, search and data retrieval management, as well as notification handling functionalities. Appropriate screenshots of the TwinERGY Integrated Platform are provided for this purpose, in order to show the functionalities included in the Core Data Management Platform Beta Release. The most important functionalities are presented and described in detail, together with suggested guidelines, so that TwinERGY CDMP users have a comprehensive understanding of what is available to them.

Based on the results of this deliverable, the next release of the TwinERGY Integrated Platform (due in M24) will focus on: a) enriching the already introduced functionalities of the TwinERGY Core Data Management Platform, and b) evaluating the identified needs towards the potential introduction of new Platform functionalities, if they are deemed necessary by the pilot partners.



Twinergy has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 957736

Index

1. Introduction	
1.1 Purpose of this deliverable	8
1.2 Scope of this deliverable	8
1.3 Structure of the document	8
1.4 Abbreviation List	9
2. TwinERGY Core Data Management Platform	10
3. Data Collection Service	
3.1 Data Ingestion Job Creation	12
3.1.1 Data Ingester	13
3.1.2 Data Mapper	15
3.1.3 Data Curator	16
3.2 Data Storage Service	16
3.2.1 Data Store	16
3.3 Metadata Editor	17
4. Data Security Service	19
4.1 Access Policies Controller	19
6. Platform Management Service	20
6.1 Search & Retrieval Manager	20
6.2 Notifications Handler	21
7. Conclusions	23
REFERENCES	



List of Figures

Figure 1 TwinERGY Core Data Management Platform Conceptual Architecture	10
Figure 2 TwinERGY CDMP Landing Page	11
Figure 3 User registration	11
Figure 4 Data Ingestion Job Creation	13
Figure 5 Display of the data ingestion jobs page	13
Figure 6 File uploading in the TwinERGY CDMP	13
Figure 7 Overview of the sampling of the files uploaded	14
Figure 8 API Data Ingestion	14
Figure 9 Mapping data to the TwinERGY CIM	15
Figure 10 Curation Routines definition	16
Figure 11 Data Store Configuration	17
Figure 12 Data Ingestion Jobs overview and completion status	17
Figure 13 Dataset's Metadata Editing	18
Figure 14 Provision of information on the granularity and coverage of the dataset	18
Figure 15 Access Level Definition	19
Figure 16 Access Policies Definition	19
Figure 17 Retrieval of datasets	21
Figure 18 Notifications on ongoing/finished data ingestion jobs	22



1. Introduction

1.1 Purpose of this deliverable

Deliverable 5.4 presents the TwinERGY Integrated Platform's beta release, including the current implementation state of the platform's services, namely the Data Collection, Data Security, Data Storage, and Platform Management services. This deliverable is an updated version of D5.3, "TwinERGY Integrated Platform - Alpha, Mock-ups Release," which covers all services and their related functionalities, as well as the extent of the TwinERGY Core Data Management Platform users' access to those functionalities.

The process for establishing and performing the core platform functionalities is outlined via intuitive and extensive screenshots.

1.2 Scope of this deliverable

The TwinERGY Core Data Management Platform's implementation status is detailed in Deliverable D5.4, "TwinERGY Integrated Platform - Beta Release", which outlines the latest functionality available to its intended users. The most essential user routines for performing fundamental platform functions, as well as the many functionalities implemented, are available in this beta release. To that aim, an end-to-end usage overview is provided, completed with specific instructions and appropriate screenshots.

1.3 Structure of the document

D5.4 is structured as follows:

Section 2 presents an overview of the TwinERGY Core Data Management Platform.

The Data Collection, Data Security, Data Storage, and Platform Management Services are described in detail in Sections 3 through 6. In order to advise and inform TwinERGY CDMP users on the available functionalities of the TwinERGY Integrated Platform, the state of implementation of their respective functionalities is also disclosed, along with appropriate screenshots.

Section 7 concludes the deliverable by presenting a brief overview of what has been described within the document.

1.4 Abbreviation List

Acronym	Full Name
API	Application Programming Interface
CA	Consortium Agreement
CDMP	Core Data Management Platform
CIM	Common Information Model
CSV	Comma Separated Values
D	Deliverable
DER	Distributed Energy Resources
EC	European Commission
EMS	Energy Management System
ICT	Information and Communication Technology
GA	Grant Agreement
H2020	Horizon 2020 The EU Framework Programme for Research and Innovation
JSON	JavaScript Object Notation
PubSub	Publish-Subscribe
XML	Extensible Markup Language
WP	Work package



2. TwinERGY Core Data Management Platform

The project's backbone is an "open", modular and interoperable Big Data Management Platform which enables open standards-based data collection and management communication across the TwinERGY project's energy value chain. The TwinERGY Core Data Management Platform follows existing open energy standards. The CDMP also incorporates a homogenized Common Information Model (TwinERGY CIM) that ensures semantic interoperability of the digitalized energy assets deployed in the pilot sites, as well as seamless integration, communication, and operation on top of any Energy Management System and Smart Home systems and devices. The TwinERGY Data Management Platform is equipped with appropriate data security, privacy, authentication, and authorization processes to ensure end-user data protection and nonrepudiation of DER assets.



Figure 1 TwinERGY Core Data Management Platform Conceptual Architecture

The TwinERGY Core Big Data platform is made up of the following essential services, as shown in Figure 1 above:

- Data Collection Service
- Data Security Service
- Data Storage Service
- Platform Management Service

The TwinERGY Core Data Management Platform, which can be found at <u>https://twinergy.s5labs.eu/</u>, is included in this deliverable.



As shown in Figure 2, the landing page welcomes users to the TwinERGY CDMP and invites them to register or log in using their existing accounts.



Figure 2 TwinERGY CDMP Landing Page

Users of the TwinERGY CDMP can commence the registration process by providing their first and last names, a username, an e-mail address, and a password that must meet specific requirements and be submitted twice, as shown in Figure 3.

FIRST NAME	LAST NAME
First Name	Last Name
USERNAME	EMAIL
Username	Email
PASSWORD	REPEAT PASSWORD
Password	Repeat Password

Figure 3 User registration

After completing the registration process, users are routed to the login page, where they can submit their credentials and log in to the TwinERGY Core Data Management Platform. Both actions (registration and login) are part of the Platform Management Service and serve the purpose of the Identity Manager functionality.

3. Data Collection Service

The Data Collection service is responsible for the data ingestion process in the TwinERGY Core Data Management Platform; thus, it is a vital aspect of the overall platform configuration. Data ingestion from various sources, semantic mapping of the data ingested in the platform, data curation to assure the correctness and integrity of the data to be processed later, and metadata editing to enrich the dataset with the necessary additional information are all part of the process.

The Data Collection Service presents a set of functionalities for configuring the data ingestion process. Any TwinERGY CDMP user can upload data to the TwinERGY Integrated Platform infrastructure via stable and secure mechanisms, such as direct batch data uploading or real-time data uploading through APIs, making use of the Data Ingester functionality. Additionally, any outcome of a data ingestion job is saved as a dataset in the TwinERGY Core Data Management Platform. This process is enabled by the Metadata Editor functionality.

The platform's user interface is duly built, to provide platform users with appropriate advice so that they can map all of the fields of ingested data using the TwinERGY Common Information Model (CIM), as part of the Data Mapper functionality.

The user can also specify the constraints and limitations that data ingested into the platform may have, as well as the specific actions that must be performed if any of these constraints are breached. The Data Curator functionality guarantees that the ingested data is of high quality and value. A curation routine is created by combining a validation option with a corrective action for a specific field, and then all of the curation routines are applied to the converted data (deriving from the mapping process to the Common Information Model), resulting in curated data.

3.1 Data Ingestion Job Creation

The Data Collection page is displayed when the users first access the platform. After clicking the Create button on the top navigation bar, the data ingestion job is created. In addition, the platform users must submit data ingestion job parameters and select suitable processing rules.

The process begins with the creation of a new data ingestion job and the specification of appropriate data processing routines. The user is asked to enter some basic information about the data ingestion job, such as the title and a brief description of the dataset that will be ingested, as well as the data processing routines that will be applied. The data ingester, data mapper, and data store data processing routines are pre-selected; however, because the data curator routine is optional, the platform users can choose whether or not to curate the dataset, as depicted in Figure 4.





TWIN	Data Collection	Datasets	Search			🗘 👩 TwinERGY User. ~
Create Data Ing	gestion Job					X Cancel 🗸 Create
				Job Details Basic Information about the job	NAME PV Generation Essentiations Information about the generation of the photo-ontac	
				Data Processing Routines Entry reporting your data to the partners, you can use determine you to where preven them	ALX INVESTIGE Collect data through files, reference JAN or other services ALX INVESTIGE The Collect data through files, such as collect of insectiones recents The Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones recents Collect data through files, such as collect of insectiones Collect data through files, such as collect of insectiones	

Figure 4 Data Ingestion Job Creation

After the required information is provided, the job is created and listed in the data ingestion jobs list, as shown in Figure 5.



Figure 5 Display of the data ingestion jobs page

3.1.1 Data Ingester

When the TwinERGY Core Data Management Platform user selects the Data Ingester on a new data ingestion job, the user can determine how the data will be ingested into the platform, either via direct file uploading or data ingestion via API. In the file uploading option, platform users can choose between different file types. As shown in Figure 6, this process is completed by submitting a sample and the actual file to be uploaded.

Data Ingestion Method How do you plan to store your data to the platform?	FILE UPLOAD Direct file upload (CSV, JSON)		
Data Ingestion Routines Select the format of the file you will use	O CSV ⊖ JSON		
Sample Upload Upload a sample of your data to be used in next steps	BROWSE DPV_Generation_Sample.csv 289.0 B		
Upload File(s) Upload your file(s) to be processed (if in csv, json)	BROWSE DPV_Generation.csv 255.3 KB		

Figure 6 File uploading in the TwinERGY CDMP



When all of the required information is provided, platform users are provided with a summary of the sample file they selected, as depicted in Figure 7. Then they are able to Save and Finalize the configuration.

			255 2 KB
			200.0 KD
DATETIME	PV PRODUCTION, KW	RELATEDGENERATIONTECHNOLOGY	
12-03-19 00:00	0	3	
12-03-19 01:00	7987.40113	3	
12-03-19 02:00	3113.56546	3	
12-03-19 03:00	3235.074906	3	
12-03-19 04:00	4793.323944	3	
12-03-19 05:00	3563.919926	3	
12-03-19 06:00	5626.647834	3	
12-03-19 07:00	9901.117318	3	

Figure 7 Overview of the sampling of the files uploaded

Platform users can also use the TwinERGY platform's API retrieval feature if they retain data in their organization systems and expose it through APIs. They are presented with a configuration screen, as illustrated in Figure 8, to begin the process of ingesting their data. After selecting the type of authentication, the platform user is asked to provide the whole API URL, as well as the method to be used (e.g., GET), as well as define any potential request parameters.

Data Ingestion Method How do you plan to store your data to the platform?	DATA PROVIDER'S AVAILABLE API Collect data from the APIs provided by applications and systems of the data provider or from open APIs					
Authentication Details Details about the authentication policies of the API	O None 🔘 Bearer	Custom				
Method, URL & Body O The method, URL and query body of the request.	GET https://a	pi.openweathermap.org/data/2.5/weather?q=athens≈	ppid=290eef18c9bea	a92b70b0e04f4c602609		C
Request Parameters						
Any url, query or body parameters that will be used on the API	PARAMETER	VALUE	TYPE	SENSITIVE		
Calls	1 g	athens	Query		C	۵
	② appid	290eef18c9bea92b70b0e04f4c	Query		C	۵

Figure	8 API	Data	Ingestion
--------	-------	------	-----------

Users of the platform are also requested to choose how frequently they like to retrieve data via the API and to set appropriate scheduling intervals. In order to quickly evaluate



the data that will be saved on the platform, the user is given a summary of the structure that is recorded, as well as selected concepts.

3.1.2 Data Mapper

Following the completion of the Data Ingester configuration, the process of mapping the ingested data to the TwinERGY CIM specified concepts in D5.1 [2] begins, by clicking the Configure Data Mapper button. Initially, the platform displays the model's name, the standards considered in the development of the TwinERGY Common Information Model, which can be selected if the user knows the corresponding standard to which the ingested data adheres, and a list of the TwinERGY CIM concepts defined to select the mapping process's entry point.

Then, users of the TwinERGY CDMP are redirected to the mapping playground, where they can map the ingested data with the TwinERGY CIM's concepts and fields, as shown in Figure 9.

DATA MODEL	PLAYGROUND		Clear Selection All Unidentified Invalid	DATA MAPPER DETAILS
O Search	SOURCE DATA	CONFIDENCE LEVEL	COMMON DATA MODEL	TITLE
	DateTime Date		observedDateTime Datetime ×	observedDateTime
~ EnergyGenerationMeasurements			GenerationMeasurements > observedDateTime	DATE FORMAT
ne activeEnergyExport				DD-MM-YYYY hh:mm:ss
M activeEnergyExportBiomass	PV production, kW Number	• ./	activeEnergyExport Double ×	
activeEnergyExportCoal		A 1	GenerationMeasurements > activeEnergyExport	
Real activeEnergyExportFossilFuel				Will be ignored if already available in your dat
activeEnergyExportGas	relatedGenerationTechnology		Id String ×	
NO activeEnergyExportGeotherma			> relatedPhotovoltaic > id	
NR activeEnergyExportHydroPowe				
🔞 activeEnergyExportNuclear				
N2 activeEnergyExportPetroleum	SAMPLE VALUES			
ne activeEnergyExportPhotovoltai	DateTime			
Real activeEnergyExportRenewable	"12-03-19 00:00" "12-03-19 01:00" "12-03-19 02:00" "12-03-19 03:00" "1	2-03-19 04:00" "12-03-19 05	:00" "12-03-19 06:00" "12-03-19 07:00"	
activeEnergyExportSolar				
activeEnergyExportSolarTherm				
activeEnergyEvportMind				

Figure 9 Mapping data to the TwinERGY CIM

On the left side of the mapping playground page, the users of the TwinERGY CDMP can search through the TwinERGY CIM and find the list of fields for the concept specified in the previous step. In addition, in the centre of the page, the user can view the data that has been ingested into the platform, and on the right side of the page, the user can define the mapping details for each field selected. To make the mapping process easier, the user is given appropriate assistance and mapping recommendations, as well as relevant confidence levels, search and drag-and-drop functionality, field descriptions, and a data sample display. A validation option is also available, which indicates to the user whether the mapping performed is correct. If not, the user can always make the required changes in order to complete the process.

Once the users are satisfied with the applicable mapping, they can choose to Save and Finish the process and continue with the Data Curator routine (if selected).



3.1.3 Data Curator

The users are redirected to the Data Curator configuration page, after completing the Data Mapper Configuration, where they can apply curation routines to the ingested data. As shown in Figure 10, users are given the option of defining one or more curation routines, as well as associated outliers' criteria, for each of the data fields in order to eliminate inaccurate values before they are eventually stored in the platform.

Data Curator curation routine execution order information Mandatory curation routines will be executed first Curation routines with outlier rule of DROP will be executed All other curation routines will be executed in the order pro	second						×
FIELDS	Clear Selection	All	Datetime	Double	String	CURATION ROUTINES	
observedDateTime Datetime						CURATION ROUTINE	
EnergyGenerationMeasurements > observedDateTime						Mandatory Curation Routine	
						Field values must not be null	
activeEnergyExport Double EnergyGenerationMeasurements > activeEnergyExport			1 Curat	tion Routine	Defined	Attention The field values you define may have been set to null during a transformation in the data mapper step	
						OUTLIERS RULE	
SAMPLE VALUES						Select Outliers Rule	
EnergyGenerationMeasurements > observedDateTime						Select Outliers Rule	
"12-03-18 00:00" "12-03-19 01:00" "12-03-19 02:00" "12-03-19 03 03-19 07:00"	00" *12-03-19 04:00" *	12-03-19	9 05:00" "12	-03-19 06:	00" "12-	Diop Replace with default value Replace with rections value Replace with most Request value Replace with max value Replace with max value	

Figure 10 Curation Routines definition

Afterwards, the TwinERGY CDMP users can review the curation routines defined for each and every data field once the Data Curator configuration is finished, and then they can complete the curation process.

3.2 Data Storage Service

The Data Storage service of the TwinERGY Core Data Management Platform provides robust measures and multiple indexing approaches to suit the demand for trustworthy data storage and indexing. The Data Storage Service is responsible for the effective and reliable storing of a wide range of data, as well as the metadata that goes with it. Depending on the type of data stored in the TwinERGY CDMP and how it is accessible when it comes to data storage, different storage and indexing capabilities are intended to fulfil different needs. At the same time, suitable metadata is collected in order for all TwinERGY modules to access it. This service also stores log-related data for the TwinERGY Core Data Management platform's operation and usage, such as users and any administrative data required to keep the TwinERGY CDMP up and running.

3.2.1 Data Store

Data Store is the final stage of a successful data ingestion job. At this point, TwinERGY CDMP users are required to provide a title for the dataset and a brief description of the data that will be saved in the platform, as shown in Figure 11.



Destination How do you want your data to be handled?	• NEW DATASET Create a new dataset and load the processed data
Dataset Information Enter a title and a short description for your dataset. You	NAME
will be able to change these once the dataset is created	Enter dataset name DESCRIPTION
	Enter a short description for your dataset

Figure 11 Data Store Configuration

Accordingly, the users of the platform are notified that the data ingestion job is complete after providing all necessary information about the dataset to be stored in the TwinERGY Core Data Management Platform, and they are redirected to the data ingestion job page, as shown in Figure 12, where they can find all the data ingestion jobs initiated as well as their respective completion status.

PV Generation # TwinERGY User Updated on Mar 2, 2022	COMPLETED	
Weather API Athens 1 TwinERCY User Dupdated on Mar 2, 2022	CONFIGURATION: DATA MAPPER	

Figure 12 Data Ingestion Jobs overview and completion status

3.3 Metadata Editor

In the TwinERGY Core Data Management Platform, any outcome of a data ingestion job is preserved as a dataset, therefore platform users must specify a complete profile of the dataset. Platform users are prompted to give a title for the dataset and a description of what it contains after heading to the Datasets tab on the top navigation bar, as depicted in Figure 13. When completing the dataset profile, users are also asked to define any applicable tags.



TWIN FERGY Data Collection Datasets Search	1	🗘 🚺 TwinERGY User 🗸
Edit Dataset Details INCOMPLETE		X Cancel Save
General Information General Information about the profile of the specific dataset	TITLE The rance of the detraset by which it can be easily identified. IV Generation DESCRIPTION Ab ord service that acts as an account of the detraset's contents. Information about the generation of the Photovoltaic TAGS At ord service and/or arbitrary tendual tage associated with the detraset by its data provider. TAGS TABIONEE LINKIOME The language of the dataset. English	

Figure 13 Dataset's Metadata Editing

Information about the dataset's coverage and granularity, as well as specifics about the temporal coverage, spatial coverage, temporal resolution, and spatial resolution units of the data are also required, in order to delineate the type of dataset, the format of the dataset to which the data are available, the dataset's language, and the dataset's temporal coverage, spatial coverage, temporal resolution, and spatial resolution units, as depicted in Figure 14.

Extent Details Information regarding the coverage and granularity of the dataset from a temporal and spatial perspective	TEMPORAL COVERAGE UNIT The time period during which the data were collected or the time period the data are referring to.
	Select temporal coverage unit
	SPATIAL COVERAGE UNIT The location/area the data refer to or were collected from; either defined directly (e.g. geographical area) or indirectly (i.e. place of interest or an activity that is the subject of the collection).
	Select spatial coverage unit -
	TEMPORAL RESOLUTION UNIT The frequency of acquiring new data from the same data source (e.g. as part of a dynamic process from a system/location/sensor). Select temporal resolution unit
	SPATIAL RESOLUTION UNIT The granularity applied within the data allowing to distinguish different spaces using the data; either in terms of actual space/ground area (e.g. room, zone, building, country etc.) or as defined by a sensor/sensor network.
	Select spatial resolution unit

Figure 14 Provision of information on the granularity and coverage of the dataset

As part of the metadata editing process, specific access policies per dataset can be applied, which is covered as a functionality in the next chapter, under Access Policies Controller section.



4. Data Security Service

The TwinERGY CDMP users' data security and privacy concerns about data that will be ingested and handled in the TwinERGY Core Data Management Platform are addressed by the Data Security Service. By allowing users to define straightforward and adaptable access routines that restrict access requests to their data on the platform, the Data Security Service helps users gain confidence in the TwinERGY CDMP. Access policy routines for each dataset can be designed using the Access Policies Controller functionality to maximize effectiveness. Users can establish advanced access policy routines for their datasets using a graphical user interface. The routines are preserved, and the user can make changes to them right away using the provided interface.

4.1 Access Policies Controller

The visibility levels and access policy routines for the corresponding datasets can be configured by TwinERGY CDMP users. By choosing visibility levels, users can choose whether the dataset is for: i) Private access (access to the dataset is allowed only provided access policy routines are followed) or ii) General access (Free access to the dataset, no requirement for access policy routines to be met), as depicted in Figure 15.

|--|

Figure 15 Access Level Definition

The necessary access regulations, as well as the relevant approach (Allow-all) and the inclusion of certain exceptions, must be configured if the access level is set to Private, as shown in Figure 16. To add exceptions, platform users must select a user parameter, a condition (such as equal, not equal, etc.,), and the value of the parameter.

Access Policy Define who has and doesn't have access to this dataset	 Allow everyone to view this dataset
	+ Add an exception
	Field
	+ AND (add another condition)
	You need to define at least one condition for policy to be added
	Cancel Create policy exception

Figure 16 Access Policies Definition



5. Platform Management Service

The Platform Management Service in TwinERGY's platform is critical because it establishes the means and processes for users to register securely and reliably on the platform. Users are provided access to the data, they are eligible to use through proper authentication and safety methods. The Identity Manager functionality represents identity provider responsibilities such as creating and managing identification information for those who are authorized to access the TwinERGY CDMP, as well as providing authentication and authorization services to restrict access to approved users, as described in Section 2. TwinERGY Core Data Management Platform) of this document. With the introduction of a user-friendly data search that enables for different search options, platform users are able to search for data using the Search & Retrieval Manager functionality. Additionally, authorized apps are able to use the TwinERGY Open APIs to configure data retrieval from a single dataset when configuring APIs for data retrieval. Filters, which are represented by API request parameters, and the selection of specific data fields are used to fine-tune the delivered results.

Finally, the Notifications Handler functionality provides platform users with real-time updates on the progress of a data ingestion job's execution (successful or unsuccessful). They can also look into various notifications and respond to them, as well as delete them if necessary.

5.1 Search & Retrieval Manager

The functionality is critical for TwinERGY's Core Data Management Platform, as it allows users to search for and discover relevant data, determine and define which of the available data is crucial, and lastly, see a clear and comprehensive overview of the results. After accessing and clicking on the appropriate tab in the platform's top navigation bar, users of the TwinERGY CDMP can see a list of the datasets they possess. They are given the necessary filtering capabilities to filter the datasets according to their data type. It's worth noting that the platform users can choose from datasets that have been published by them or by other users, as illustrated in Figure 17.



All		۹)
8	RESULTS 2	~
	PV Generation Store of the Photovoltaic GenerationMeasurements	RESULTS
FILTERS		
	Due to access level limitations, the result shown is insisting 1 frem	

Figure 17 Retrieval of datasets

The users of the platform can also create queries utilizing a flexible free-text search and data filtering, providing them more options in how they find and search for data within the TwinERGY Core Data Management Platform. The platform's users can look for data that is relevant to their needs, explore the results, and go further into the data to find eligible candidates for retrieval. The options for retrieval must be defined through APIs, with the choice of the concepts that a user wishes to acquire from each dataset. Furthermore, concepts that are utilized as query parameters to filter query results must be defined. The retrieval configuration ends with instructions on how to utilize the TwinERGY APIs to acquire the retrieval results. To that purpose, authentication instructions are included alongside the endpoints (i.e., for GET and POST methods).

5.2 Notifications Handler

Users of the TwinERGY Core Data Management Platform get relevant notifications about the progress and current status of their data ingestion jobs, as depicted in Figure 18.



	4	TwinERGY User 🗸
Notifications	Mark all as read	+ Create
PV Generation suc execution!	ccessfully finished its 2 hours ago	
See	e All	

Figure 18 Notifications on ongoing/finished data ingestion jobs



6. Conclusions

The Beta Release of the TwinERGY Integrated Platform, as well as the services that constitute the platform (Data Collection, Data Security, Data Storage and Platform Management), are described in Deliverable 5.4. The current state of the platform's services, as well as the range of the functionalities provided to the TwinERGY CDMP users, are also assessed. In particular, this deliverable elaborated on:

- The platform's data ingestion process (through direct file uploading and APIs), the mapping with the TwinERGY CIM, and data processing routines definition during this process.
- The Data Security Service, which is responsible for addressing the platform users' data security and privacy concerns regarding the data that are ingested in the TwinERGY CDMP. Utilizing the functionalities offered by this Service, platform users can establish access policies for the datasets they ingest into the platform, in order to prevent unauthorized access to them.
- The Data Storage Service, which is responsible for addressing the demand of the platform users for reliable data storage and indexing.
- The platform management functionalities introduced within the TwinERGY Core Data Management Platform. Users of the TwinERGY CDMP can edit and modify their personal profile information, as well as change their existing password. They can also look for data that they are allowed to use according to the access policies that have been set up. The users are given the ability to filter the available results, select and use the ones that best suit their needs. Finally, they can be notified about the success or failure of a data ingestion job.

D5.5 "Data Collection, Security, Storage & Management Services Bundles – Release 1.00" [3] will provide a more mature and enriched version of the services in comparison to Deliverable D5.2 "Data Collection, Security, Storage & Management Services Bundles" [4], while D5.6 "TwinERGY Integrated Platform– Release 1.00" [5], will present an enriched version of the functionalities included in the current deliverable, including Data Ingester enhancements to cover data streaming input and the update of data by appending new rows.



REFERENCES

- 1. TwinERGY Consortium. (2020). TwinERGY D5.3 "TwinERGY Integrated Data Management Platform Alpha, Mock-ups Release"
- 2. TwinERGY Consortium. (2020). TwinERGY D5.1 "TwinERGY Common Information Model"
- 3. TwinERGY Consortium. (2020). TwinERGY D5.5 "Data Collection, Security, Storage & Management Services Bundles Release 1.00"
- 4. TwinERGY Consortium. (2020). TwinERGY D5.2 "Data Collection, Security, Storage & Management Services Bundles Beta Release"
- 5. TwinERGY Consortium. (2020). TwinERGY D5.6 "TwinERGY Integrated Data Management Platform– Release 1.00"