FI.ICT-2011.1.8 FINESCE

D7.12 version 1.0

FINESCE Data Repository in FIWARE Lab Data

Contractual Date of Delivery to the CEC:	*N/A Extra Deliverable*
Actual Date of Delivery to the CEC:	
Author(s):	Sonja Kolen, Jonas Otten
Participant(s):	*RWTH Aachen University*
Workpackage:	*WP7 – FINESCE Project Management*
Estimated person months:	**8**
Security:	**PU**
Nature:	**R** = Report
Version:	1.0
Total number of pages:	**10**

Abstract:

FINESCE provides live data from its trial sites over the FINESCE API. This document describes how this data has been published as open, historical data through the FIWARE Lab Data platform, ensuring that the Smart Energy data of FINESCE remains available after the live data streams stop.

Keyword list:

FIWARE, Open Data, Smart Energy

Disclaimer:

All information provided reflects the current status of the trial site testbeds at the time of writing and may be subject to change.

Executive Summary

FINESCE makes live trial site data publically available via the FINESCE API. The functioning of this API depends on having live FINESCE trial sites; once the trial sites are no longer on-line, the FINESCE API will no longer be able to provide the data from the non-working parts of the trial sites. In order to ensure that the trial site data remain available in such an eventuality, an effort has been made to collect all historical data from the FINESCE trial sites and to publish them in a platform which is independent of the FINESCE project, which is referred to in this report as the FINESCE Data Repository (FDR).

This report describes the approach to the FINESCE Data Repository in FIWARE Lab Data and the current status of work. The live data of all FINESCE trial sites have been downloaded via FINESCE API and stored on computers at the RWTH Aachen University. Sensitive data are anonymized or left out and the resulting datasets have been stored in FIWARE Lab Data.

Authors

Partner	Name	Phone / Fax / e-mail
RWTH Aachen University		
	Sonja Kolen	Phone: +49 241 80 49718 Fax: +49 241 80 49709 e-mail: skolen@eonerc.rwth-aachen.de
	Jonas Otten	Phone: +49 241 80 49742 Fax: +49 241 80 49709 e-mail: jotten@eonerc.rwth-aachen.de

Table of Contents

1. Introduction

FINESCE makes live trial site data publically available via the FINESCE API. The functioning of this API depends on having live FINESCE trial sites; once the trial sites are no longer on-line, the FINESCE API will no longer be able to provide the data from the non-working parts of the trial sites. In order to ensure that the trial site data remain available in such an eventuality, an effort has been made to collect all historical data from the FINESCE trial sites and to publish them in a platform which is independent of the FINESCE project, which is referred to in this report as the FINESCE Data Repository (FDR).

The objective of the FDR is to make the data collected by all FINESCE trials publicly available for evaluation and use beyond the duration of the FINESCE project. As a means of storage and publication, the FIWARE Lab Data platform is employed. External entities, e.g., researchers, SMEs, can use the interface of FIWARE Lab Data to search for and download FINESCE data, which are referred to as “Smart Energy Data” in FIWARE Lab. The FIWARE Lab Data platform is freely available to be used by the public, without the need to register or open an account for browsing or downloading data.

The overall process of transferring the data from the internal FINESCE trial site systems to FIWARE Lab Data, including the anonymization of data, is shown in Figure 1 and described in the following Section 2. Section 3 describes the results of the work.

The FINESCE datasets are published as open data under the Creative Commons CC-BY-SA 4.0 license. This license allows the datasets to be used if the data source is accredited and if the same licensing conditions (CC-BY-SA 4.0) are applied to its derivative use.

2. Approach

This section describes the way that all live data have been collected from the FINESCE trial sites, processed, and stored on the FIWARE Lab Data platform. Furthermore, it is outlined how users can access the FDR.

2.1 Step 1: Read FINESCE Live Data from all Trial Sites

As shown in Figure 1, the first step towards the FDR is the reading of all live data from the trial sites via the RESTful FINESCE API. The data are stored on computers at RWTH Aachen University for further processing. At this point, only authorized RWTH Aachen University people can access the data, it is not published yet. The download process has been automated through Python scripts as far as possible.

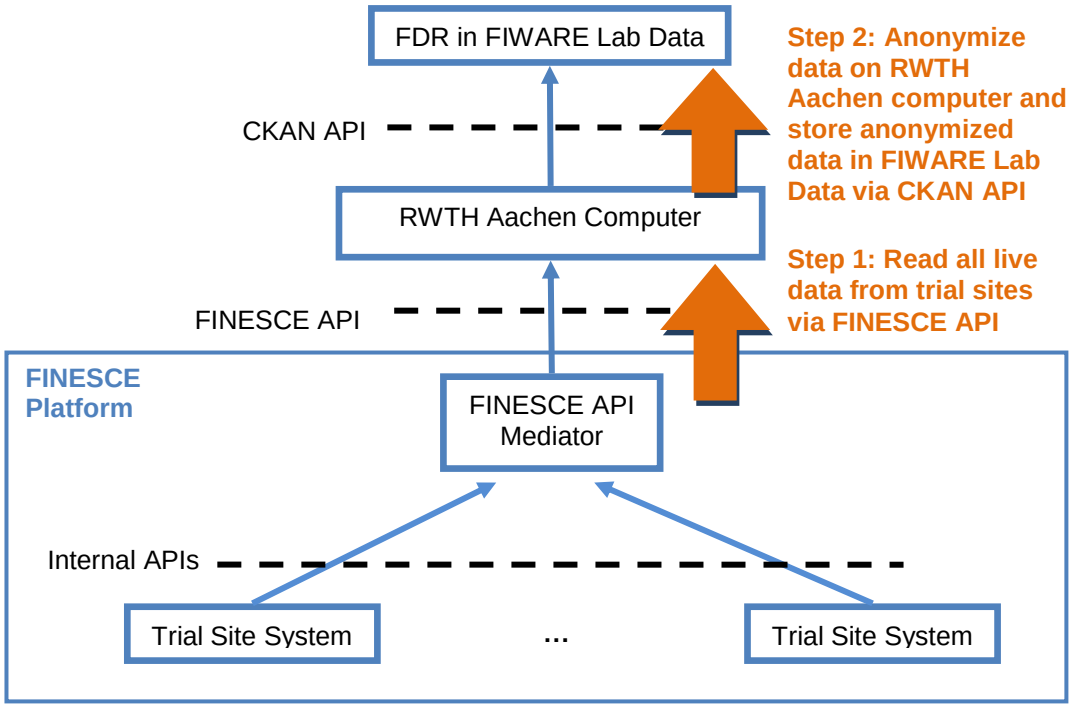

Figure 1: Approach to FINESCE Data Repository in FIWARE Lab Data

2.2 Step 2: Anonymize and Upload Data to FIWARE Lab Data

The live data contain sensitive information such as IDs, addresses, or names that need to be anonymized before the data can be published. The anonymization is done according to sensitive data lists which outline the sensitive attributes per trial and state the anonymization action that is required. Tables 1 to 5 provide the sensitive information lists for the trials Aachen, Horsens, Ireland, Malmö and Terni. No sensitive information could be identified for the Madrid trial. In the given tables only data that are supposed to be published are taken into account and all endpoints refer to FINESCE API version 0.1. A local backup of the anonymized data is stored at RWTH Aachen University.

The anonymized data are uploaded to FIWARE Lab Data via CKAN API. Furthermore, a brief description is provided in FIWARE Lab Data for each dataset and resource. External entities can use the web interface of the FIWARE Lab Data platform to browse and download the FINESCE datasets. Additionally, the data are accessible via CKAN API.

Table 1: Sensitive Information List for Trial Aachen

Trial Site	Endpoint	Class	Attribute	Action
Aachen	/Aachen/factory/equipment/ machines	Machine	id	use synonym
Aachen	/Aachen/factory/equipment/ machines	Machine	name	leave out, synonym for id is enough
Aachen	/Aachen/vpp/components	Component	access	leave out, synonym for id is enough
Aachen	/Aachen/vpp/components	Component	url	leave out
Aachen	/Aachen/vpp/{comp_type}/ {comp_id}/data	Component Data	url	leave out

Table 2: Sensitive Information List for Trial Horsens

Trial Site	Endpoint	Class	Attribute	Action
Horsens	/Horsens/buildings	Building	address	leave out
Horsens	/Horsens/vehicles	Vehicle	id	use synonym
Horsens	/Horsens/vehicles	Vehicle	related_entities	use synonyms

Table 3: Sensitive Information List for Trial Ireland

Trial Site	Endpoint	Class	Attribute	Action
Ireland	/Ireland/charging_states	ChargingState	ev_id	use synonym
Ireland	/Ireland/vehicles/{id}/ connections	Connection	ev_id	use synonym
Ireland	/Ireland/vehicle_types	VehicleType	Id	leave out
Ireland	/Ireland/vehicle_types	VehicleType	brand	leave out
Ireland	/Ireland/vehicle_types	VehicleType	manufacturer	leave out
Ireland	/Ireland/connections	Connection	ev_id	use synonym
Ireland	/Ireland/connections	Connection	ev_supply_ equipment_id	use synonym
Ireland	/Ireland/supply_equipment	VehicleSupply Equipment	Id	use synonym

Table 4: Sensitive Information List for Trial Malmö

Trial Site	Endpoint	Class	Attribute	Action
Malmö	/Malmo/weather/{from}/{to}	all weather measurement types		accredit www.yr.no as origin
Malmö	/Malmo/prices/{from}/{to}	Price	id	leave out, not relevant

Table 5: Sensitive Information List for Trial Terni

Trial Site	Endpoint	Class	Attribute	Action
Terni	/Terni/meters	Meter	address	leave out
Terni	/Terni/meters	Meter	id	leave out, not relevant
Terni	/Terni/meters	Meter	customer_id	use synonym
Terni	/Terni/meters	Meter	latitude	leave out
Terni	/Terni/meters	Meter	longitude	leave out
Terni	/Terni/meters	Meter	related_entities	use synonyms

2.3 FINESCE Data Repository DSE

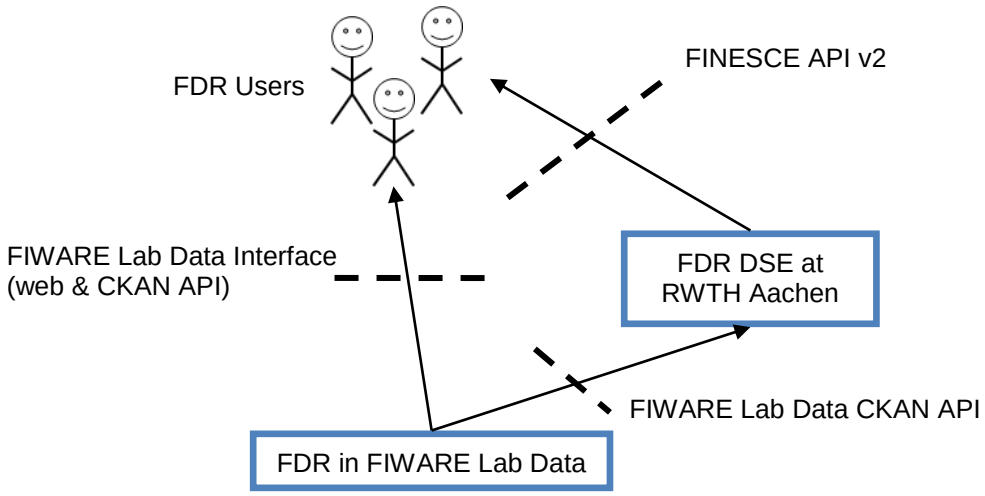

Figure 2: FINESCE Data Repository DSE

In order to offer another access possibility for FDR, a FINESCE Data Repository DSE (Domain Specific Enabler) is established at RWTH Aachen University. This can be seen as a modified version of the FINESCE API Mediator DSE. It offers a FINESCE API v2 (Figure 2) to the FDR users. The advantage of this additional access option is a (possibly high) downward compatibility for applications that use the FINESCE API during the FINESCE project. Furthermore, the data access options that the FIWARE Lab Data platform offers could be extended leading to enhanced API features for FDR users.

3. Results

The CKAN-based FINESCE Lab Data platform offers data producers the opportunity to publish their data as a "CKAN organization". Such an organization comprises different members (users) who can upload and maintain datasets on behalf of an organization. The FINESCE data is published by the organization "RWTH Aachen University" in FIWARE Lab Data. All FINESCE datasets are equipped with respective tags to enhance findability. "FINESCE" and "Smart Energy" are tags common to all datasets. Searching this tag on the FIWARE Lab Data platform gives you all available FINESCE datasets.

3.1 Step 1 and Step 2

All available live data from all trial sites have been collected and stored at RWTH Aachen University (step 1). An anonymized version of the data has been uploaded with respective descriptions to the FIWARE Lab Data Platform in several datasets (step 2).

We published the FINESCE data in 30 datasets in FIWARE Lab Data. Their names follow, if applicable, the FINESCE API categories. In order to enhance usability and comprehensibility for FDR users who are not familiar with FINESCE details and FINESCE API, we made some of the dataset names more descriptive and detailed as the FINESCE API categories specify. Table 6 provides a list of the available datasets.

Table 6: FINESCE Datasets in FIWARE Lab Data

No.	Dataset name
1	Smart Energy Data: Aachen/ Cologne Virtual Power Plant
2	Smart Energy Data: Aachen/ Cologne Smart Factory
3	Smart Energy Data: Horsens Building Information House01
4	Smart Energy Data: Horsens Building Information House02
5	Smart Energy Data: Horsens Building Information House03
6	Smart Energy Data: Horsens Building Information House04
7	Smart Energy Data: Horsens Building Information House05
8	Smart Energy Data: Horsens Building Information House06
9	Smart Energy Data: Horsens Building Information House07
10	Smart Energy Data: Horsens Building Information House08
11	Smart Energy Data: Horsens Building Information House09
12	Smart Energy Data: Horsens Building Information House10
13	Smart Energy Data: Horsens Building Information House11
14	Smart Energy Data: Horsens Building Information House12
15	Smart Energy Data: Horsens Building Information House13
16	Smart Energy Data: Horsens Building Information House14
17	Smart Energy Data: Horsens Building Information House15
18	Smart Energy Data: Horsens Building Information House16
19	Smart Energy Data: Horsens Building Information House17
20	Smart Energy Data: Horsens Building Information House18
21	Smart Energy Data: Horsens Building Information House19
22	Smart Energy Data: Horsens Building Information House20
23	Smart Energy Data: Ireland Energy in Regions
24	Smart Energy Data: Ireland Electric Vehicles
25	Smart Energy Data: Madrid Weather Forecast
26	Smart Energy Data: Madrid Smart Office Building
27	Smart Energy Data: Malmo Energy Prices/ Weather Forecast
28	Smart Energy Data: Terni Energy Consumption Profiles
29	Smart Energy Data: Terni Power Demand/ Supply Profiles
30	Smart Energy Data: Terni Weather Forecast/ Social Events

3.2 FINESCE Data Repository DSE

During our work on FDR, we discarded the idea of a FDR DSE at RWTH Aachen University for the following reasons:

- No existing application requiring a backward compatible FINESCE API could be identified
- The CKAN DataStore extension incorporated in FIWARE Lab Data already offers an extended CKAN API with exhaustive search options for users that wish to browse/ download FINESCE data from FDR via an API.

For these reasons, the benefit of having a FDR DSE at RWTH Aachen University is not worth the effort required for its implementation and maintenance. Hence, no FDR DSE has been made.

In order to improve the service for FDR users nonetheless, we have uploaded a merged CSV data file for every dataset. This file contains the data of all resources that belong to it. Hereby, datasets are better searchable via the DataStore extension API in FIWARE Lab Data.

4. List of Abbreviations

API	Application Programming Interface
CKAN	Comprehensive Knowledge Archive Network
CSV	Comma-separated Values
DSE	Domain Specific Enabler
FDR	FINESCE Data Repository
REST	Representational State Transfer