Ref. Ares(2023)5937369 - 31/08/2023

H2020 – LC – SC3 – EE – 2019 – GA 894240

Operating System for Smart Services in Buildings



D3.6 Reference Data Sets

WP3 Common Ontology and Semantics

	Name	Date
Prepared by	Christian Thomsen (AAU), Pierre Ferrez (HES-SO), Henrik Lund Stærmose (NEOGRID), Morten Veis Donnerup (NEOGIRID)	13.08.2023
Peer reviewed by	Dominique Gabioud (HES-SO)	18.08.2023
Reviewed and approved by	Dominique Gabioud (HES-SO)	31.08.2023



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 894240.

Distribution list

External		Internal	
European Commission	X	Consortium partners	x

Change log

Version	Date	Remark / changes
1.1	13.08.2023	First version by AAU
2.1	20.08.3023	Second version by AAU
3.1	31.08.2023	Final version by HES-SO

To be cited as

"D3.6 Reference Data Sets" of the HORIZON 2020 project domOS, EC Grant Agreement No 894240.

Disclaimer

This document's contents are not intended to replace the consultation of any applicable legal sources or the necessary advice of a legal expert, where appropriate. All 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 user, therefore, uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors' view.

Table of contents

Executive Summary **4**

1. Introduction **4**

2. Overview of the Reference Data Sets **4**

2.1. Reference Data Set from Sion Demonstrator 4

2.2. Reference Data Set from Aalborg Demonstrator 7

2.3. Reference Data Set from Skive Demonstrator 12

3. Conclusion **13**

Executive Summary

This document presents the three reference data sets from the domOS project. The data sets have been pseudonymised and can be shared for research purposes. It is not possible to share the data sets as open data as it is not possible to do anonymisation of them. Metadata descriptions of the data sets are available on Zenodo.org.

1. Introduction

In the domOS project, smart building technologies are demonstrated in five different sites. In those demonstrators, data is collected from sensors and meters in the buildings. The collected data is of vital importance to the demonstrators to operate their offered smart services. The data can also be useful to the wider research community when evaluating or simulating future smart services. To enable this, the domOS project has prepared reference data sets. As described in "D1.2 Data Management Plan" and "D3.4 Rules and Methods for Data Anonymisation and GDPR Compliance", the data sets collected by the demonstrators contain personal information and it is not possible to do anonymisation of them. Therefore, the prepared data sets have been pseudonymised (for example, by replacing exact addresses with IDs). As described in "D3.4 Rules and Methods for Data Anonymisation and GDPR Compliance", not every demonstrator could share pseudonymised data and three reference data sets have therefore been prepared. They are from the Sion (CH), Aalborg (DK), and Skive (DK) Demonstrators.

Given that the reference data sets contain personal data and have been pseudonymised, but not anonymised, they cannot be shared as open data. Instead, metadata descriptions are made available in the open data repository Zenodo from where it is also possible for researchers to request access to the data. Such requests will be evaluated on a case-by-case basis and require a formal, written agreement to ensure GDPR and consent forms compliance. Even though the reference data sets are not shared as open data, they still fulfil the FAIR principles as they are **Findable** and have DOIs and machine-readable metadata, **Accessible** as it is described how to get access and there is access to metadata, **Interoperable** as descriptions of the data are provided, and **Reusable** as metadata descriptions are available and procedures for getting access to the data are established. See "D3.4 Rules and Methods for Data Anonymisation and GDPR Compliance" for more details.

The deliverable "D3.6 Reference Data Sets" thus consists of the metadata descriptions available on Zenodo together with this short overview of the reference data sets.

2. Overview of the Reference Data Sets

2.1. Reference Data Set from Sion Demonstrator

DOI: 10.5281/zenodo.8238570

Zenodo Record: https://zenodo.org/record/8238570

Description: This dataset includes data about the electricity consumption of 15 single-family residential households all located around Sion (Switzerland), in the OIKEN service area (https://oiken.ch). All the

households are equipped with an individual heat pump providing energy for both space heating and domestic hot water.

The ongoing data collection began in January 2021 and includes the households' total electricity consumption provided by the Smart Meter. It also includes the specific consumption of the heat pump and the production of the solar panels (where applicable) provided by dedicated sub-meters.

Metadata:

Buildings			
Type	**Number**	**Size**	**Heating system**
Single-family residential households	15	100 – 300 m^2 of living area	Individual heat pump

Data			
Household total electricity consumption	**Heat pump electricity consumption**	**Solar panels electricity production (where applicable)**	**Weather info**
Active and reactive power for both import and export (aggregate of all three phases)	Active power (single-phase measurements)	Active power (single-phase measurements)	Possibility to correlate with Sion weather data provided by MeteoSwiss
Unit: Watts / VAR Sampling interval: 5 seconds	Unit: Watts Sampling interval: 15 seconds	Unit: Watts Sampling interval: 15 seconds	-

Building Metadata				
Attribute	**Description**	**Type**	**Unit**	**Range of values**
BuildingID	Unique Identifier (UID)	String	-	"000" – "300"
BuildingArea	Total heated area	Integer	m^2	100 - 300

Consumption Data				
Attribute	**Description**	**Type**	**Unit**	**Range of values**
Time	Measurement data and time in UTC	ISO 8601 timestamp YYYY-MM-DDT hh:mm:ssZ	-	Since 2020-12-31T23:00:00Z (corresponds to 2021-01-01 T00:00:00Z CET)
Power	Measured active or reactive power	Integer	Active: Watts Reactive: VAR	0 - 100000

Data Structure

The data is organized in JSON files, there's one file per household and per day. Each JSON file has the following seven attributes:

- "Active Import", "Active Export", "Reactive Import", "Reactive Export", "Heat Pump L1", "Heat Pump L2" and "Heat Pump L3"

Where applicable (households equipped with solar panels), each JSON file has the following three extra attributes:

- "Solar Panels L1", "Solar Panels L2" an "Solar Panels L3"

Each of these attributes contains two arrays:

- "Time" and "Power"

The "Time" array contains the timestamps (YYYY-MM-DDThh:mm:ssZ format), and the "Power" array contains the corresponding values of the active or reactive power (integer).

Each house has its own directory (name: BuildingID), and the name of the JSON file contains the day in YYYYMMDD format followed by "RawData.json".

Example of file hierarchy

```
040
├── 20210101RawData.json
├── 20210102RawData.json
│   ...
└── 20221231RawData.json
123
├── 20210507RawData.json
├── 20210508RawData.json
│   ...
└── 20230801RawData.json
```

Example of a JSON file (including solar production)


```
{
    "Active Import":
        {
            "Time": ["2020-12-31T23:00:04Z", "2020-12-31T23:00:09Z", ..., "2021-01-01T22:59:59Z"],
            "Power": [379, 379, ..., 3381]
        },
    "Active Export":
        {
            "Time": ["2020-12-31T23:00:04Z", "2020-12-31T23:00:09Z", ..., "2021-01-01T22:59:54Z"],
            "Power": [0, 0, ..., 0]
        },
    "Reactive Import":
        {
            "Time": ["2020-12-31T23:00:04Z", "2020-12-31T23:00:09Z", ..., "2021-01-01T22:59:54Z"],
            "Power": [0, 0, ..., 3150]
        },
```

```
    "Reactive Export":
        {
            "Time": ["2020-12-31T23:00:04Z", "2020-12-31T23:00:09Z", …, "2021-01-01T22:59:55Z"],
            "Power": [311, 311, …, 0]
        },
    "Heat Pump L1":
        {
            "Time": ["2020-12-31T23:00:01Z", "2020-12-31T23:00:16Z", …, "2021-01-01T22:59:47Z"],
            "Power": [0, 0, …, 1181]
        },
    "Heat Pump L2":
        {
            "Time": ["2020-12-31T23:00:01Z", "2020-12-31T23:00:16Z", …, "2021-01-01T22:59:47Z"],
            "Power": [0, 0, …, 938]
        },
    "Heat Pump L3":
        {
            "Time": ["2020-12-31T23:00:02Z", "2020-12-31T23:00:16Z", …, "2021-01-01T22:59:47Z"],
            "Power": [0, 0, …, 934]
        },
    "Solar Panels L1":
        {
            "Time": ["2020-12-31T23:00:14Z", "2020-12-31T23:00:29Z", …, "2021-01-01T22:59:59Z"],
            "Power": [0, 0, …, 0]
        },
    "Solar Panels L2":
        {
            "Time": ["2020-12-31T23:00:14Z", "2020-12-31T23:00:29Z", …, "2021-01-01T22:59:59Z"],
            "Power": [0, 0, …, 0]
        },
    "Solar Panels L3":
        {
            "Time": ["2020-12-31T23:00:14Z", "2020-12-31T23:00:29Z", …, "2021-01-01T22:59:59Z"],
            "Power": [0, 0, …, 0]
        }
}
```


2.2. Reference Data Set from Aalborg Demonstrator

DOI: 10.5281/zenodo.8289790

Zenodo Record: https://zenodo.org/record/8289790

Description: This dataset includes data about 147 family homes of which 144 are flats in 12 building blocks and the rest are single family buildings. All buildings are owned by a building association and are heated via the local district heating company. The buildings have been renovated and are now prepared for low temperature district heating. Within the area, new private homes (mostly blocks and flats) are being built or renovated. The district heating supply to the 12 building blocks is supplied from a local mixing loop separating the district heating transmission from the local distribution. The heating installations in each block can be controlled as well as the mixing loop to the area. Energy and indoor climate data is supplied from selected apartments and from each heating central. The district heating to

the three single family buildings is not supplied with district heating from the local mixing loop, but from the transmission line directly. In each building, datalogging and control on the heating installation are possible. Furthermore, data from indoor climate sensors as well as sensors on doors and windows is available. The ongoing data collection has been started in several steps. On the three single family houses, data collection started in February 2020. For the apartment buildings data collection began in June 2021 and for the central mixing loop it began in March 2022.

Metadata:

The data set contains W3C Web of Things (WoT) Thing Descriptions[1] (TDs) which are JSON files that describe what IoT devices can do. Further, the data set contains domOS Common Ontology (dCO) Building Descriptions (BDs) which are machine-readable, abstract descriptions of buildings and their properties As BDs are presented in "D2.3 Functional Specification of the IoT-Ecosystem" and "D3.5 domOS Common Ontology, Final Version", they will not be further described here. The Thing Descriptions and Building Descriptions are named as td-*ID*.json and bd-*ID*.json, respectively.

The rest of the data set consists of several groups of files as described below.

"buildingMetadata"

This group of files also contains information about the buildings.

The building metadata will be exported once a day but is expected to be pretty stable. Although the current process lacks automation through Neogrid's exporter, this functionality is expected to be implemented in the near future.

Attribute	Description	Type	Unit	Range of values
locationId	Location identifier	int	-	
buildingType	Apartment building, single family house	Text	-	Name
buildingArea	Total heated area	Int	m^2	
buildingApartments	Number of apartments	Int	-	

"IndoorClimate"

This group of files contain measurements related to the indoor climate in the buildings.

Attribute	Description	Type	Unit	Range of values
locationId	Location identifier	int	-	
zoneId	Zone identifier	Int	-	0 if no zones are defined
timestampEnd	Measurement Date and	Timestamp	-	[01-01-1970

[1] https://www.w3.org/TR/wot-thing-description10/

	Time: DD-MM-YYYY HH:MM:SS. All timestamps are in UTC-0 (and do not reflect daylight saving time)			00:00:00; 31-12-9999 23:59:59]
indoorClimate CO2	CO2 level data	Float	Ppm	[490; 10000]
indoorClimate RH	Relative Humidity	Float	%	[0; 100]
indoorClimate Temperature	Indoor air temperature	Float	°C	
indoorClimate WindowOpen	Windows and door sensors	Int	-	

"*districtHeatingControl*"

This group of files contain measurements related to the district heating control.

Attribute	Description	Type	Unit	Range of values
locationId	Location identifier	int	-	
zoneId	Zone identifier	int	-	0 if no zones are defined
timestampEnd	Measurement Date and Time: DD-MM-YYYY HH:MM:SS. All timestamps are in UTC-0 (and do not reflect daylight saving time)	Timestamp	-	[01-01-1970 00:00:00; 31-12-9999 23:59:59]
heatLoopControl tSupply	Forward temperature (heat loop)	Float	°C	[0; 120]
heatLoopControl tReturn	Return temperature (heat loop)	Float	°C	[0; 120]
heatLoopControl actualSetPoint	Set point (heat)	Float	°C	[-1; 120]
heatLoopControl motorPosition	Estimated motor valve position	Float	%	[0; 100]
energyMain tSupply	Forward temperature (main inlet)	Float	°C	[0; 120]
energyMain tReturn	Return temperature (main inlet)	Float	°C	[0; 120]
hotWaterLoopContro tSupply	Forward temperature (domestic hot water)	Float	°C	[0; 120]

hotWaterLoopControtReturn	Return temperature (domestic hot water)	Float	°C	[0; 120]
hotWaterLoopControl actualSetPoint	Set point (domestic hot water)	Float	°C	[-1; 120]
hotWaterLoopControl motorPosition	Estimated motor valve position	Float	%	[0; 100]

"*energyMeters*"

This group of files contain measurements from energy meters.

Attribute	Description	Type	Unit	Range of values
locationId	Location identifier	int	-	
zoneId	Zone identifier	int	-	0 if no zones are defined
timestampEnd	Measurement Date and Time: DD-MM-YYYY HH:MM:SS. All timestamps are in UTC-0 (and do not reflect daylight saving time)	Timestamp	-	[01-01-1970 00:00:00; 31-12-9999 23:59:59]
HEATING				
heatMeterPrimary volume	Volume	Float	m^3	[0; ∞]
heatMeterPrimary flow	Flow	Float	l/h el m^3/h	[0; ∞]
heatMeterPrimary energy	Energy	Float	kWh	[0; ∞]
heatMeterPrimary tSupply	Forward temperature	Float	°C	[0; 120]
heatMeterPrimary tReturn	Return temperature	Float	°C	[0; 120]
DOMESTIC HOT WATER				
hotWaterPrimary volume	Volume	Float	m^3	[0; ∞]
hotWaterPrimary flow	Flow	Float	l/h el m^3/h	[0; ∞]
hotWaterPrimary energy	Energy	Float	kWh	[0; ∞]

hotWaterPrimary tSupply	Forward temperature	Float	°C	[0; 120]
hotWaterPrimary tReturn	Return temperature	Float	°C	[0; 120]
COLD WATER				
coldWaterMain volume	Volume	Float	m^3	[0; ∞]

"*meteorologicalData*"

This group of files contain meteorological measurements.

Attribute	Description	Type	Unit	Range of values
locationId	Location identifier	int	-	
zoneId	Zone identifier	int	-	0 if no zones are defined
timestampEnd	Measurement Date and Time: DD-MM-YYYY HH:MM:SS. All timestamps are in UTC-0 (and do not reflect daylight saving time)	Timestamp	-	[01-01-1970 00:00:00; 31-12-9999 23:59:59]
meteorologicalData temperature	Temperature	Float	°C	[0; 100]
meteorologicalData humidity	Humidity	Float	%	[0; 100]
meteorologicalData_ windSpeed	Wind Speed	Float	m/s	[0; 100]
meteorologicalData_ windDirection	Wind Direction	Float	°	[0; 360]
meteorologicalData clodiness	Cloudiness	Float	%	[0; 100]

Data Structure

File format: JSON

Available files: A new file is available for each hour. To be sure all data have been received from one hour, data will be exported after the first measurement from the next hour has been received. The file is named with the location, date, and time.

Measurement granularity: One measurement file pr. location (including all buildings) pr. group pr. hour.

File content: JSON structure including data and units from all buildings.

Number of records: 15 apartment buildings, 1 mixing loop, 3 single family buildings (tbd)

Example from the file Afd._50_-_Ravnkildevej_energyMeters_yyyymmdd_hhmm.json


```
{
    "locationId": "2521",
    "zoneId": "0",
    "timestampEnd": "2020-04-23 00:00:00",
    "heatMeterPrimary_volume": 4828.872083333333,
    "heatMeterPrimary_volumeUnit": "m³",
    "heatMeterPrimary_energy": 253914.5416666667,
    "heatMeterPrimary_energyUnit": "kWh",
    "heatMeterPrimary_tSupply": 71.5,
    "heatMeterPrimary_tSupplyUnit": "°C",
    "heatMeterPrimary_tReturn": 27.0,
    "heatMeterPrimary_tReturnUnit": "°C",
    "heatMeterPrimary_flow": 0.05791666666666667,
    "heatMeterPrimary_flowUnit": "m³/h",
    "heatMeterPrimary_power": 2.8041666666666667,
    "heatMeterPrimary_powerUnit": "kW",
    "hotWaterPrimary_volume": 2282.120416666667,
    "hotWaterPrimary_volumeUnit": "m³",
    "hotWaterPrimary_energy": 68027.16666666667,
    "hotWaterPrimary_energyUnit": "kWh",
    "hotWaterPrimary_tSupply": 76.125,
    "hotWaterPrimary_tSupplyUnit": "°C",
    "hotWaterPrimary_tReturn": 52.0,
    "hotWaterPrimary_tReturnUnit": "°C",
    "hotWaterPrimary_flow": 0.10208333333333335,
    "hotWaterPrimary_flowUnit": "m³/h",
    "hotWaterPrimary_power": 2.9375,
    "coldWaterMain_volume": 893.0,
    "coldWaterMain_volumeUnit": "m³"
}
```


2.3. Reference Data Set from Skive Demonstrator

DOI: 10.5281/zenodo.8296898

Zenodo Record: https://zenodo.org/record/8296898

Description: The Skive Living Lab consists of 12 residential single-family houses and is founded on two different platforms: SUNTHERM SmartHeat heating systems, where NEOGRID has a very detailed data level, and 'legacy' systems, where existing heating systems are retrofitted with a NEOGRID gateway to collect data.

The heating distribution systems have been upgraded to enable control of the indoor climate on room level. Either by replacing traditional thermostats with Smart Thermostats, or by interfacing with under floor heating controllers. Some buildings are fitted with both Smart Thermostats and underfloor heating.

Metadata:

Building Type: Residential

Heating system: Heat pumps, supplying central heating.

- Data collection and control is provided via Neogrid gateway and platform.
- Wireless sensor and energy data from the single-family houses is collected and controlled via Neogrid gateways and platform.
- Data from Neogrid can be supplied in 5 min resolution.
- Datalogging on oldest sites goes back to 2019 – and as of spring 2023 all 12 demonstration sites are collecting data.

The available data is similar to what is described for the Aalborg Demonstrator and the metadata descriptions will not be repeated here.

3. Conclusion

This document has presented the three reference data sets from the domOS project and their metadata. The data sets originate from the Sion (CH), Aalborg (DK), and Skive (DK) domOS Living Labs. The data sets can be found on Zenodo and researchers can request access to them. The data sets are owned by the respective demonstration leaders, and they will therefore handle such requests. The data sets fulfil the FAIR DATA principles.