2022

green energy lab.at

Open Data Platform
Final Report

Foto: Pecold/Shutterstock

Foto: Manuel Schmid/Shutterstock

Foto: shutterstock, © metamorworks

This project was part of the Flagship Region Green Energy Lab and supported with the funds from the Austrian Climate and Energy Fund and implemented in the framework of the RTI-initiative “Flagship Region Energy”.

klima+ energie fonds

FTI Initiative Energy Model Region

Publishable final report

Programme control:
Climate and Energy Fund

Programme management:
The Austrian Research Promotion Agency (FFG)

Final Report
created on
20/07/2022

Green Energy Lab Open Data Platform (GEL ODP)

Project number: 868693

Call	2. Call FTI Initiative Energy Model Region
Project Starting Date	01/11/2018
Project Ending Date	31/04/2022
Total duration of project (in months)	42 Months
Project holder (Institution)	TU Wien – Institute of Energy Systems and Electrical Drives (TUWESEA)
Contact person	Georg Lettner
Postal address	Gußhausstraße 25 – 29 / E370-03, 1040 Vienna, Austria
Telephone	+43 1 58801 370376
Fax	-
E-mail	lettner@eeg.tuwien.ac.at
Website	https://www.gel-odp.com/

GEL ODP

Green Energy Lab - Open Data Platform

Authors:
Christoph Loschan (TUW-ESEA, Energy Economics Group)
Georg Lettner (TUW-ESEA, Energy Economics Group)
Christoph Maier (TUW-ESEA)
Wolfgang Gawlik (TUW-ESEA)
Christian Lechner (EVN)
Thomas Schlager (EVN)
Eva Fleiß (University of Graz - ISIS)
Philip Ohnewein (AEE INTEC)
Christian Pfeiffer (Forschung Burgenland)
Aline Leiner (ms.GIS)
Lukas Künzel (ms.GIS)
Johann Schrammel (AIT)
Lisa Diamond (AIT)
Normann Lindner (twingz)

1 Table of content

2 Abstract

The transition from a centralised energy system on the basis of fossil fuels to a decentralised, renewables-based system requires the large-scale implementation of technological innovations and measurement systems. Such intelligent metering devices (smart meters) enable customers to analyse their energy consumption and to participate pro-actively in the electricity market. This can be done, for instance, through flexible electricity prices and adapted electricity procurement. For such an implementation, however, a data transfer and a subsequent visualisation of the data is necessary. In this manner, the measurement data can be made available to the stakeholders involved: end customers, energy suppliers, grid operators, but also research companies. A better knowledge of the consumption and generation of the customers enables distribution grid operators to better understand the condition of the grid and free grid capacities can subsequently be made available. This enables an improved integration of consumers with high peak loads such as heat pumps or electric vehicles.

In the GEL OpenDataPlatform (GEL ODP) project, a publicly accessible web platform for the energy sector was developed to provide easy access to and an overview of relevant data and interrelationships in the energy system.

The online platform enables customers to view their energy consumption and generation of a photovoltaic system online in a clear and concise manner. The greatest importance was attached to compliance with the Basic Data Protection Regulation. Although all information is geographically assigned to an area, it can no longer be traced back individually by combining several households. The smallest possible resolution always comprises at least five households, which means that data protection is implemented in the best possible way in combination with local weather data that is as accurate as possible.

This weather information is used to calculate the expected heat demand with the assistance of an intelligent predictive algorithm. From that, a temporal flexibility with regard to the electricity demand can be derived, which covers both the heat demand and the electricity demand. It can also be used in the future, e.g. for more cost-effective electricity procurement. This promotes the active participation of customers in the electricity market.

In addition, algorithms developed particularly for this research project make it possible to disaggregate the measured total electricity demand and assign it to specific appliance classes. The resulting classification of total consumption by individual consumers enables customers to identify high-consumption devices. On the one hand, this creates a better understanding of electricity costs, and on the other hand, reasonable measures to increase energy efficiency can be identified.

3 Introduction

The predominance of our fossil fuel-based energy system demands a transition to a sustainable energy system. A successful transition should include the prompt and widespread adoption and diffusion of technical innovations, which also triggers an individual and thus social change in energy consumption. Technical innovations in the course of the digitalization can support an energy transition due to the ability of connecting different stakeholders. The importance of an increased acceptance of smart meters is the high potential to serve as a digital enabler in order to collect information about electricity flows and energy use. A better understanding of electricity flows in the low voltage grid level is essential to all stakeholders in the energy system, pointing out the system operators (transmission – TSO and distribution - DSO), the energy companies and the consumers. From a system operator's point of view, a better forecast reduces risks related to congestions in the grid area. Even more, the system operators are responsible for distributing the energy with the purpose of supplying the market and for achieving grid-balance at all times. The costs of balancing measures are charged to the balance group that was responsible for the respective fluctuation. In the event of unforeseen fluctuations in generation (e.g. through power plant failures, changing renewable energy generation) or deviations from the expected consumption level, the energy balance in the grid must be guaranteed by the control area manager. This is done through the connection or disconnection of generation units (e.g. backup power plants). Finally, the additional cost of imbalance and ill-managed grids are billed by the energy company to the costumers. From the consumer perspective (households or SMEs) at present there is no insight in the energy consumption on the level of individual appliances. Especially households with solar PV production, battery storage units, heat pumps or other appliances which operate automatically, there is no way to track any changes in the overall consumption. Especially for the non-professional user, it is important to relate energy consumption to individual appliances where they can change their consumption pattern. Hence, these predictors allow energy companies a better day ahead prediction which in turn can be leveraged to save millions of energy costs.

A major challenge is to get people involved in new ODP Findings in the field of open government ODP's show that it is difficult to trigger off the civil society providing data and make their benefit clear. This challenge relates to an often-lacking stakeholder-orientated approach of ODPs, which facilitates a goal-oriented feedback to users and thus a decreasing user satisfaction. Moreover, user acceptance is a major challenge in respect to people involvement and participation. Spatially remote, invisible climate damage is hardly conceivable for consumers and even less perceivably linked to individual behaviour. Regarding an ODP investigation, several stakeholder groups have to be considered within a participation process and the discovery of the ODP has to be goal-targeted, possible in an effortless and useful way for different stakeholders. Apart from overall awareness building through stakeholder participation, ODPs may also be an enabler for more efficient everyday data-driven energy feedback and management systems. By exploiting larger data quantities, persuasion strategies can be applied in a targeted manner to influence user behaviours. Also, the opportunities of disaggregated consumption on the device-level for generating more meaningful behaviour change recommendations are essential. This also involves the aspect of experienced privacy issues that are tackled by an aggregation and anonymisation approach. To foster user engagement the arising challenge is to come up with suitable means for data-driven, pre-prepared and easily customizable engagement programs.

3.1 Project goals and methodological approach

The main goal of the GEL ODP project [1] was to develop and implement an ODP for the energy sector in order to provide easy access and overview of relevant data and interdependencies of a current and future integrated energy system. The GEL-ODP shall be usable for all relevant stakeholders (e.g. end-users, research institutions, system operators, energy suppliers, start-ups, technology providers, public bodies, policy makers, etc.), which are part of the Green Energy Lab (GEL) cluster. Furthermore GEL-ODP is able to easily integrate various data-sources, e.g. from sub-projects. An additional aim of the platform is to provide all end customers insight into their energy consumption or efficiency data and allow a comparison with similar households (same size of PV, heat pump, number of residents) and tailor-made recommendations for energy relevant measures. Moreover, the platform should serve as the basis for professional users of start-ups, SMEs, utilities, system operators, public bodies, research institutes etc. to get input for new business ideas and opportunities.

The participation of end-users represents a goal to find models of participation that leads to an adoption of technologies to generate energy data, such as smart meters and open data platforms. In particular, to support the provision and exchange of end-user energy data and to develop such participation models for respective stakeholder groups.

Based on a stakeholder analysis, participation models are developed for the respective stakeholder groups to get in contact and ensure participation in the GEL ODP project. EVN equiped electrical heating Systems such as heat pumps and boilers or classic ripple control with the “optimization assistant” (joulie OA [2]) to measure the electrical consumption of those devices. Furthermore, EVN equiped households with PV production and electrical heating system such as boilers with the joulie optimization assistant to measure PV production, the electrical consumption of the devices and the overall load of the household. Additionally, the load of electrical vehicles and the overall load of households can be measured. The installed joulie system enables “Demand Side Management” for the equipped devices and will provide input for the e.g. simulation of the electric system or other external algorithms that are connected via an API to the ODP.

Furthermore, the project aims to provide disaggregated end-user data to the GEL-ODP. Hence, the energy data streams of households were processed via the twingz disaggregation services. The result is the identification of the end-users most electricity consuming appliances. This result allows the development of the consumption pattern and can be provided to the consumers via Apps (e.g. EVN App) and other platforms (e.g., sub-project Smart Grid Control market and energy communities platform), enabling them to understand their household’s energy costs and system impact.

Another goal was the development of consumption patterns and predictive models based on these patterns. The better understanding of the electricity flows in the low voltage grid level is based on knowing the appliances consuming and generating electricity, their activity track and their behaviour under certain circumstances. Therefore, flexibility potential must be identified in the energy data streams of end-users, to generate usage pattern models. This enables the prediction of the behaviour of clusters of end-users and single appliances. Based on this, demand and supply on the low voltage level can be foreseen, resulting in the identification of the related flexibilities.

The project demonstrated the opportunities enabled by GEL-ODPs that explicitly include energy consumers via live interaction. The project investigated suitable data-driven engagement programs, which react to certain events or developments in the consumption data and which trigger dedicated persuasive strategies.

3.2 Project structure

The project GEL ODP received funding from the Austrian Research Promotion Agency (FFG) under the 2nd Call FTI Initiative Energy Model Region programme.
It started in November 2018 and was completed in April 2022. It was divided into seven work packages. The first contains the project management, followed by six task that are explained in the next section.
A consortium, representing a broad spectrum, with extensive experience in numerous research projects, conducted the research project:

- TU Wien – Institute of Energy Systems and Electrical Drives (TUWESEA) has extensive know-how in the area of modelling of energy systems and networks in various preliminary projects, including the integration of electro mobility, dynamic simulations of complex energy systems as well as optimization and energy storage.
- EVN AG (EVN) is a successful global player for years and supplies the majority of customers in lower Austria with electricity. Furthermore, EVN actively optimises and forecasts a broad portfolio
- of assets.
- The AIT Austrian Institute of has a strong track record in the development and execution of national and international research and cooperation projects related in the domain of user-centered smart grid services.
- AEE - INTEC (AEE) has been working for more than 20 years on the development of solar thermal components and systems and on the optimization of industrial processes and energy systems.
- twingz development Gmbh (twingz) has several patents on the Disaggregation Technology are ready for filing. The market success is proven by numerous proofs of concept.
- ms.gis Informationssysteme GmbH (msGIS) provides Software as a service application for municipalities and governmental institutions, which helps to speed up data generation and maintenance processes within their organisations. The reliable geospatial Internet of Things platform allows big data collection and analysation in real time.
- Karl-Franzes-Universität Graz - Institut für Systemwissenschaften, Innovations- und Nachhaltigkeitsforschung (KFU-SIS) has extensive experience in interdisciplinary projects addressing stakeholder engagement, perceptions of renewable energy technologies, energy efficient behaviour as well as the translation of self-reported behaviour data into a serious game.
- Forschung Burgenland GmbH (FB) possesses expertise in the area of quantitative and qualitative social scientific methods, especially participative designs.

3.3 Structure of this report

This report summarizes the results of the project GEL ODP. It is structured in the following manner:

- An introduction and overview is given in chapter 3 while the content of the other six work packages is presented in chapter 4 to 7. Each of these chapters contains one or two work packages.
- Chapter 4 gives a detailed description of the used data sources and the subsequent handling of the generated data.
- Chapter 5 describes the main functionalities and user roles of the developed Open Data Platform.
- In Chapter 6 the external applications that are executed and visualised on the ODP are presented and described.
- The User Profiling and Engagement as well as the evaluation of the ODP are described in Chapter 7.
- Following this the conclusion and an outlook is made in Chapter 8.

4 Data structure and implementation

4.1 Data Sources

4.1.1 Joulie Optimization Assistant

The joulie optimisation Assistant (OA) is an easy-to-use system that connects several devices behind the meter and enables end customers to do PV-optimization and connects devices to a Virtual Power Plant (VPP). The vendor and a by EVN contracted party for this product is tiko Energy Solutions AG. This system provides the hard- and software to connect electrical devices in private households in order to measure their consumption and control them within certain boundaries. In the ODP project the joulie delivered measurement data (active power in 5 min resolution) to the platform. 112 customers with 298 measured devices were equipped with the system and provide continuous data over a long time period of 1 to 3 years.

For the end-customer the joulie OA provides the possibility to see how much energy the individual devices in his home consume and helps to identify potential to increase efficiency. In addition to this visualization feature, the devices can also be controlled automatically by the system. Most of these control mechanisms practically mean, that the consumption of a device is shifted to another point in time when it is more efficient – for various reasons described in the following.

The individual devices within the household are typically connected with a dedicated piece of hardware, such called K-Box. Up to 20 K-boxes can be connected to a central gateway, the M-Box. In order to minimize wiring efforts, the communication among the devices is set up via power line communication (PLC) along the power lines that are already present in the house. The M-Box serves as the communication hub for all devices within the house to the central backend, where most storage and processing of data takes place. The central backend is integrated into EVNs infrastructure. Finally, the end-user can interact with the system via a web-portal or an app on his mobile device – the frontend applications.

While there are many similar energy management systems on the market today, this has a special feature that sets it apart: it enables EVN to also do active flexibility management and connect the devices to a Virtual Power Plant (VPP).

The system consists of a communication device the so-called M-Box and a measurement device with integrated actor (potential free contact and future interfaces e.g.: Modbus). The M-Box is the gateway. It connects all other devices within the home via PLC (power line communication). The connection to the backend can be established via 3G mobile communication or through connection to a local internet router. The K-Box is a three-phase smart meter that measures the consumption of either the home (serially connected behind the utility-meter) or an individual producer/consumer in the home. It is equipped with an additional two-state relay that serves to switch potential-free contacts on suitable devices.

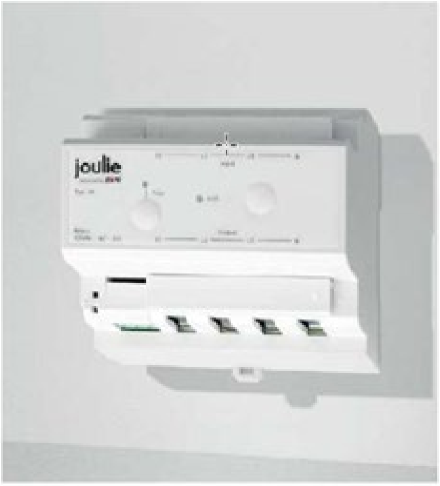


Figure 3: Current hardware joulie OA

Visualization is one of the core functionalities of the system and it is provided for all supported devices. The power consumption of all devices that are connected is continuously measured and the data together with several other information and other adjustment options (e.g. PV-Optimization) is displayed in the portal and app for the end customer. There are several backend systems to manage and support the installations and customers. The system overview is shown in Figure 4.

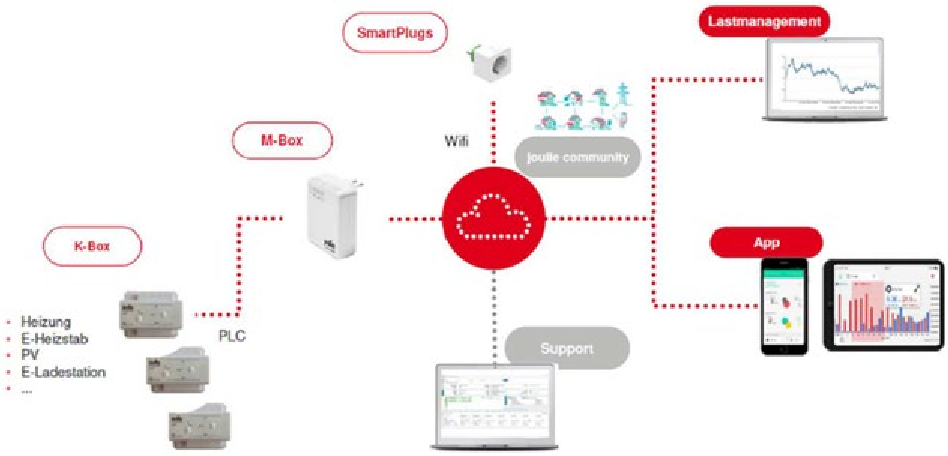


Figure 4: Overview of joulie OA and system, provided by tiko Energy Solutions AG

A core function is the connection to the VPP. Decentralized assets are aggregated and activated as a pool in order to replicate the characteristics of a conventional power plant. In case of a VPP with heating assets like heat pumps and boilers, an extremely fast reaction time is possible because the switching action is almost instantaneous. The control signal is implemented by switching on/off the individual assets. That means that the VPP shifts the energy consumption of the heating loads in such a manner as to withhold enough power and be able to follow the control signal. Figure 5 depicts the operating principle of this VPP.

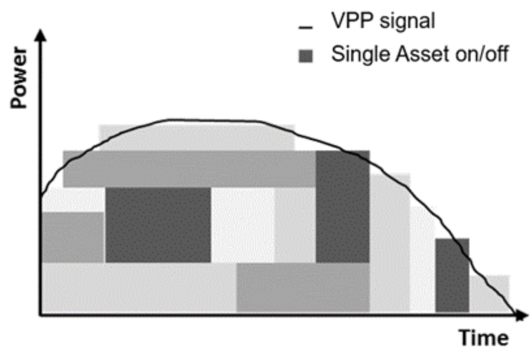


Figure 5: Depiction VPP signal implementation

The VPP feature did not play an active role of the ODP project but was offered to all customers and did additionally provide value and knowledge gain within the project. Furthermore, their potential was enhanced by the ODP.

4.1.2 Twingz eCoach

In the context of the project twingz has processed metering data streams of 110 households which were equipped with twingz eCoaches. The eCoach is a plug-and-play device that allows users to better understand their energy consumption in the home or office. Energy companies, in combination with these analytics tools, get a clear overview of their customers by using continuous anonymized data streams. These can be used to improve prediction of the entire customer portfolio. For energy consumers, the e-Coach can connect to all types of electricity meters, from mechanical meters to smart meters.

- Users can view their consumption at the appliance level via the Android / iOS app. They can make their homes energy efficient by setting their own budgets and goals and applying the advice provided.
- Organizations can monitor real-time data from anonymized users. The eCoach generates deep insights into consumers' energy consumption routines and enables end users to customize services to meet their needs.
- Device-level energy consumption information enables consumers to understand where and how they can save on their energy costs.
- The Twingz API can be integrated with third-party products like the ODP

Twingz has also prepared the processing of metering data streams generated/provided by EVN's Joulie Box. While the eCoach/twingz generated data streams are automatically processed by the twingz Disaggregation Analysis SaaS via existing APIs/Interfaces, the EVN-related data streams are delivered to twingz via the ODP Platform and after processing the results are provided back to the GEL-ODP.

To optimize the quality of the analysis of the household-related minute - and 15-minute interval data - it is mandatory to recognize and verify consumer appliance activity, before feeding it to the self-learning disaggregation system. For this purpose, twingz is using a dedicated tool, the twingz Disaggregation IDE (Integrated Desktop Environment), with which the disaggregation system can be trained.

4.2 Data Clusters

Prosumer and consumer data of 112 households (298 measuring points in 04/2022) are measured by EVN's Joulie Box. The aggregated energy data in 5 minutes temporal resolution are stored and used in the GEL-ODP data platform. The data is provided by EVN as a CSV download via an SFTP server approximately one month afterwards and then integrated on the ODP. In addition to a holistic measurement of each pro/consumers energy balance at the interconnection point measurements of specific devices are done. The data type corresponding to the interconnection point measurement differs based on the used appliances. As thermal energy data are of high relevance to investigate the consumption behavior of hot water and heating, additional measurements of the electrical power consumption of the installed heat pumps have been proceeded. The complete overview of the various devices/data types are listed in Table 1.

Table 1: Measurements and data types of EVN's Joulie Box

Abbreviation	Temporal resolution	Value range	Description
HOM	5 minutes	Positive (energy procurement)	Interconnection point, without PV system installed (sum of power demand)
HMP	5 minutes	Positive (energy procurement), Negative (PV feed-in)	Interconnection point, with PV system installed (sum of power demand or feed-in, if PV generation is higher than demand)
HMBP	5 minutes	Positive (energy procurement), Negative (PV feed-in)	Interconnection point, with PV and BES system installed (sum of power demand or feed-in, if PV generation is higher than demand)
HPU	5 minutes	Positive (power consumption)	Heat pump power demand
PVO	5 minutes	Positive (power generation)	PV system power generation
BOI	5 minutes	Positive (power consumption)	Boiler or heating element
CAR	5 minutes	Positive (power consumption)	Vehicle charging infrastructure
NSH	5 minutes	Positive (power consumption)	Night storage
DEH	5 minutes	Positive (power consumption)	Direct electrical heating
UNK	5 minutes	Positive (power consumption)	Other / unknown device

In the context of the project, twingz has set up metering data streams which have been equipped with twingz's eCoach. All of these households are non-prosumer households, hence, there were no PV or other energy-producing devices installed which influences the overall household's electrical energy consumption and the energy procurement. To get high-resolution data the eCoach, as the energy data stream generating unit, is reading/measuring the household's overall energy consumption in near real-time resolution and aggregates the measured data to a data stream with a one-minute temporal resolution. To reduce the amount of data (and the memory requirements) the data can be down-sampled to a 15 minutes temporal resolution to make processing and further calculations more efficient.
After the measured data has been uploaded to the twingz Cloud it is not only stored but also automatically disaggregated. The disaggregated end-user data provides further information about the energy consumption of a household, split up onto the level of single appliances and devices. In the end, the sum of all calculated appliances is the aggregated energy consumption of the household. Figure 1 shows the typical data report which is generated by the disaggregation process.

	timestamp	SB kWh	SB count	SB duration	DU kWh	DU count	DU duration	FR kWh	FR count	FR duration	WM kWh	WM count	WM duration	DR kWh	DR count	DR duration	DW kWh	DW count	DW duration	OV kWh	OV count	OV dur
2	17.10.2020 00:00	3,719	0	0	8,22	0	0	0,072	0	0	1,477	1	110,5	0	0	0	0,567	1	62,938	0	0	
3	18.10.2020 00:01	3,737	0	0	10,105	0	0	0,053	0	0	0	0	0	0	0	0	2,097	2	138,015	0	0	
4	19.10.2020 00:01	3,724	0	0	6,368	0	0	0,081	0	0	0	0	0	0	0	0	1,01	2	130,233	0	0	
5	20.10.2020 00:00	3,754	0	0	19,887	0	0	0,052	0	0	0,765	1	44,167	0	0	0	1,652	1	76,989	0	0	
6	21.10.2020 00:00	3,73	0	0	10,565	0	0	0,056	0	0	0,729	1	44,167	0	0	0	1,923	2	126,3	0	0	
7	22.10.2020 00:01	3,742	0	0	11,469	0	0	0,052	0	0	0	0	0	0	0	0	0,663	1	36,82	0	0	
8	23.10.2020 00:00	3,731	0	0	4,414	0	0	0,106	0	0	0	0	0	0	0	0	2,786	2	140,682	0	0	
9	24.10.2020 00:00	3,715	0	0	4,477	0	0	0,054	0	0	0	0	0	0	0	0	1,734	1	72,97	1,666	1	
10	25.10.2020 00:00	3,9	0	0	6,851	0	0	0,585	0	0	1,13	1	81	0	0	0	1,422	2	89,696	0	0	
11	26.10.2020 00:00	3,698	0	0	9,505	0	0	0,353	0	0	0	0	0	0	0	0	0,696	1	36,82	0	0	
12	27.10.2020 00:00	3,711	0	0	10,128	0	0	0,13	0	0	0	0	0	0	0	0	2,581	2	124,805	0	0	
13	28.10.2020 00:00	3,732	0	0	15,179	0	0	0,106	0	0	0	0	0	0	0	0	1,498	1	72,97	0	0	
14	29.10.2020 00:01	3,698	0	0	10,929	0	0	0,143	0	0	0	0	0	0	0	0	1,492	1	68,367	0	0	
15	30.10.2020 00:01	3,735	0	0	6,687	0	0	0,141	0	0	0	0	0	0	0	0	1,438	1	57,167	0	0	
16	31.10.2020 00:01	3,733	0	0	7,43	0	0	0,127	0	0	1,669	2	146,917	0	0	0	1,587	1	71,167	1,504	1	
17	01.11.2020 00:01	3,73	0	0	12,625	0	0	0,13	0	0	0,577	1	59,517	0	0	0	3,591	3	213,515	0	0	
18	02.11.2020 00:00	3,73	0	0	10,697	0	0	0,066	0	0	0	0	0	0	0	0	2,992	3	196,742	0	0	
19	03.11.2020 00:01	2,012	0	0	3,508	0	0	0,164	0	0	0	0	0	0	0	0	0	0	0	0	0	
20	04.11.2020 16:53	1,11	0	0	7,667	0	0	0,023	0	0	0	0	0	0	0	0	0	0	0	0	0	
21	05.11.2020 00:00	3,712	0	0	14,967	0	0	0,112	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	06.11.2020 00:01	3,155	0	0	12,139	0	0	0,154	0	0	0,387	1	88	0	0	0	0,763	1	36,82	0	0	
23	07.11.2020 08:44	2,344	0	0	10,56	0	0	0,021	0	0	1,144	1	81	0	0	0	3,978	3	193,376	0	0	
24	08.11.2020 00:00	3,726	0	0	9,421	0	0	0,226	0	0	0,564	1	59,517	0	0	0	1,315	1	67,156	0	0	
25	09.11.2020 00:01	3,74	0	0	23,938	0	0	0,072	0	0	0	0	0	0	0	0	1,862	1	67,294	0	0	
26	10.11.2020 00:00	3,731	0	0	13,357	0	0	0,137	0	0	0	0	0	0	0	0	0	0	0	0	0	
27	11.11.2020 00:00	3,749	0	0	20,235	0	0	0,034	0	0	0,544	1	59,517	0	0	0	1,256	1	67,258	0	0	
28	12.11.2020 00:00	3,744	0	0	15,902	0	0	0,069	0	0	0	0	0	0	0	0	0,596	1	36,82	0	0	
29	13.11.2020 00:00	3,41	0	0	12,285	0	0	0,034	0	0	0	0	0	0	0	0	1,946	2	143,867	0	0	
30	14.11.2020 00:01	2,03	0	0	9,68	0	0	0,119	0	0	0,576	1	59,517	0	0	0	0	0	0	0	0	
31	15.11.2020 07:35	2,559	0	0	18,575	0	0	0,025	0	0	0,682	1	59,517	0	0	0	0,542	1	73,425	0	0	
32	16.11.2020 00:00	3,735	0	0	9,758	0	0	0,111	0	0	0	0	0	0	0	0	0	0	0	0	0	
33	17.11.2020 00:01	3,707	0	0	16,547	0	0	0,142	0	0	0	0	0	0	0	0	0	0	0	0	0	
34	18.11.2020 00:00	3,747	0	0	12,226	0	0	0,145	0	0	0	0	0	0	0	0	0	0	0	0	0	
35	19.11.2020 00:00	3,754	0	0	16,731	0	0	0,094	0	0	0	0	0	0	0	0	2,768	2	137,367	0	0	
36	20.11.2020 00:01	3,731	0	0	12,686	0	0	0,111	0	0	0	0	0	0	0	0	2,199	1	73,273	0	0	
37	21.11.2020 00:01	3,754	0	0	17,623	0	0	0,013	0	0	0	0	0	0	0	0	1,349	1	67,156	0	0	
38	22.11.2020 00:01	3,732	0	0	11,907	0	0	0,04	0	0	0	0	0	0	0	0	1,226	2	135,291	0	0	

report_twingz_B3486C4F-AF3C-4A9

Bereit

Figure 1: eCoaches data which has been disaggregated by twingz

This disaggregation process is carried out continuously every 24 hours since the calculations are more accurate if an entire day is considered. The different appliances which can be detected by the algorithms are listed in Table 2.

Table 2: Appliances, which can be detected by the disaggregation process

Abbreviation	Description
SB	Stand by
DU	Day use
FR	Fridge
WM	Washing Machine
DR	Dryer

DW	Dishwasher
OV	Oven
HP	Heat Pump
CC	Car Charger
CO	Cooking
HE	Heating
MW	(system use)
UN	Unlabeled
AC	Air Condition
NO	(system use)
WD	Washer Dryer

Each appliance is classified by the amount of energy that is used in total and the duration per day. Besides, if the appliances have not a continuous power consumption but specific runtimes, their number is counted.

Table 3: Parameters per appliance

Type	Description
kWh	The day's consumption aggregate of the specific label of energy usage
count	In case it is a cyclic energy usage, the number of cycles is indicated
duration	The length of usage in minutes, aggregated for the specific day.

Beyond the data listed in Table 3, the proportion of a specific appliance in relation to the overall energy consumption is calculated. After this disaggregation process has been carried out, the data results are transferred back to the GEL-ODP and can be further analyzed or visualized.

Additionally, to the measured data which is provided by EVN and other publicly available data can be imported and visualized on the ODP. Energy market prices in Austria can be analyzed to further adapt energy usage of a household and provide market flexibilities. In a future stage, RES-E production, such as PV or wind generation, as well as the overall energy demand could be included to the ODP. Furthermore, this additional data makes a specific household comparable with the region or the whole country. In addition, weather data is visualized on the platform to be able to draw conclusion between the weather and RES-E generation.

4.3 Data privacy

Within the scope of the task, the requirements for the platform were specified and the appropriate software technologies were selected based on the platform requirements. The Fiware Framework is recommended by the EU for Smart City applications. It was evaluated within the ODP project to the test the Fiware framework for the energy use case. The Fiware framework offers so called "building blocks" with functions such as context, time series, identity management as well as data models. The ODP uses data models, which comply with the NGSI specifications.

To display spatially referenced data anonymously on a map, a GeoHash method is used. GeoHash is a method to encode a geographical position (Latitude and Longitude) into a short sequence of letters and numbers. It will not follow any given political or existing human known border and is randomly organized across the surface of the earth. As a GeoHash is a multi-level ordering mechanism, it enables to aggregate geographical positions into higher, thus more inaccurate, levels to provide anonymity. By this procedure it is possible to represent personal energy data, anonymized. Platform users have the possibility to enter their address in the web application. This is indexed using the spatial indexing method "H3" [3]. This method was introduced by Uber and is open sourced. It uses hexagons as a geographical reference of the GeoHash method. They are displayed on a map with varying degrees of accuracy.

The platform was developed in consideration of the need to fulfill the requirements:

- Preferred use of open source elements
- Interoperability
- Scaling for processing big data
- Provide long-term storage of data
- Enable processing of real-time data
- GDPR/DSGVO compliant

The users of the platform should be able to have different roles/represent different target groups. Users should have a personal dashboard for the visualization of their energy data and spatial reference of the energy data. To provide data privacy and anonymity, the ODP platform displayed energy data on the map if at least 5 end customers where inside a GeoHash.

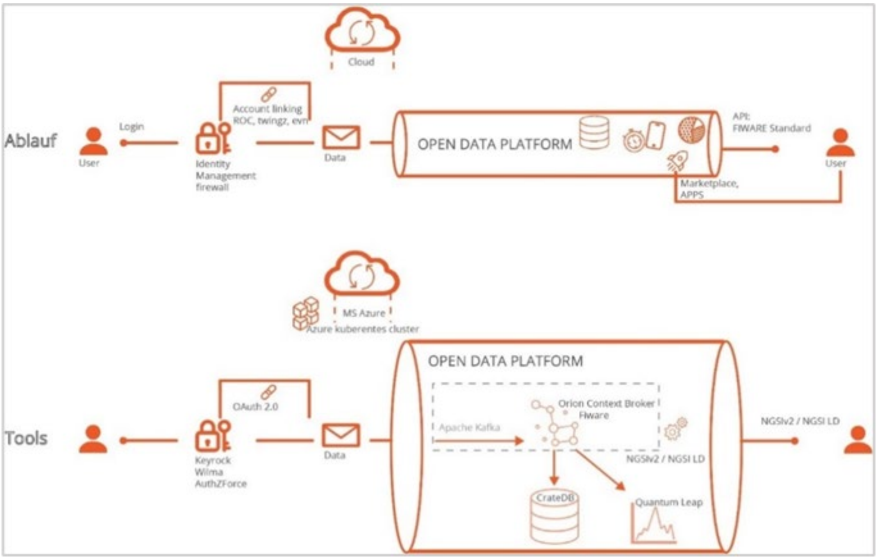


Figure 1 ODP-workflow and ODP-tools

The figure above shows the data flow and data linking method how end users can connect to their data respecting general rules of data privacy and protection methods. No automatic data transfer between EVN and ODP was implemented due to security reasons. EVN provides the data manually as CSV. ms.GIS imports it into the platform with a script. End users can connect to their data in the ODP by entering there hardware key of the EVN system, which only they knew and is directly one to one linked to a single user.

However, to fulfill the goal and to test real time data processing, ms.GIS implemented an interface to selected smart home test households from their partner company roc-connect.

5 Open Data Platform

5.1 Stakeholders and Key users

The stakeholders were defined as two groups:

- Key users, who provide most of the functionalities like research institutes and to research projects affiliated municipalities and/or companies.
- Customers, who provide data.

These stakeholders are implemented as roles into the ODP, with different functionalities, as listed in *Table 4*.

Table 4 ODP user roles

Role	Organization	Function
Private user	ODP-user	Sees his own data in the webapp sees public data* in the webapp Can download his own data Can share his own data** Can download shared data** in the webapp EVN user can link to, share and download his data
Key user	specific organization e.g. EVN, twingz, AEE	sees all (user) data they have contributed in the webapp Can download the raw data via the webapp Cannot share end-user data with all for the webapp Can share raw data for research***
	research	download raw data

Organizations have access to their contributed data and can download this raw data, while research organizations have access to all raw data as long as permission is given.
Private users have access to their own data, can see public data, aggregated data of other customers and data of customers who have specifically granted sharing their data with selected customers.

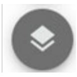
To implement a role-based user management the ODP uses the Fiware building block key rock, which is an open sourced standard Identity Management Tool. It follows the principle of applications, organizations and users and how those areas interact with each other. The user role concept follows the following principles:

- A user can belong to one or more organizations
- An organization can have one or more users
- An organization bundles access rights to applications
- Every registered ODP user can access all public data
- Each user can have different access rights on the platform (depending on the organizations they belong to)

5.2 Interface

In this task the goal was the implementation of the ODP web application. To do so ms.GIS provided a mockup for discussions the graphical user interface with the consortium. Furthermore, the ODP logo as well as some icons for the web application were designed. The website can be accessed via the following URL: https://www.gel-odp.com [1]. The web platform provides the pages and functions as listed in Table 5.

Table 5 Main functionalities of the ODP

Icon	Name	Function
-	Landing Page	Information about the platform, imprint, contact information and project partners
	Login/logout	Identity management
	Settings	Disclaimer, User profile, basic tour (first steps at the platform), contact information
	Layers	Presentation of energy and energy related data on the map
	Marketplace	Applications integrated on the platform. For research partners also, a data-marketplace for data download.
	Account Linking	Access specific data from another account through ODP (commonly known e.g. from spotify "login in with facebook account"). Within the project: EVN Joulie data and roc smart home
	My devices	Personal devices: device information, time series data (line chart), device disaggregation (pie chart), combine charts (view time series data from more than one device), add a device manually with csv upload, edit and delete devices
	Download	Download share and public time series data from a device category by clicking on one or more hexagon on the map

The landing page of the ODP can be seen in *Figure 2*, that leads to main menu. The user/privacy settings can be done as shown in *Figure 3*. The overview of the map layers and one example layer can be seen in *Figure 4*. *Figure 5* shows the configuration possibilities of such layers, while *Figure 6* provides the main information about it. *Figure 7* shows the ODP internal market place that allow to include external applications directly into the ODP. EVN joulie customers can directly connect their metering device into the ODP as shown in *Figure 8*. An overview of several visualization options can be seen in *Figure 9*.

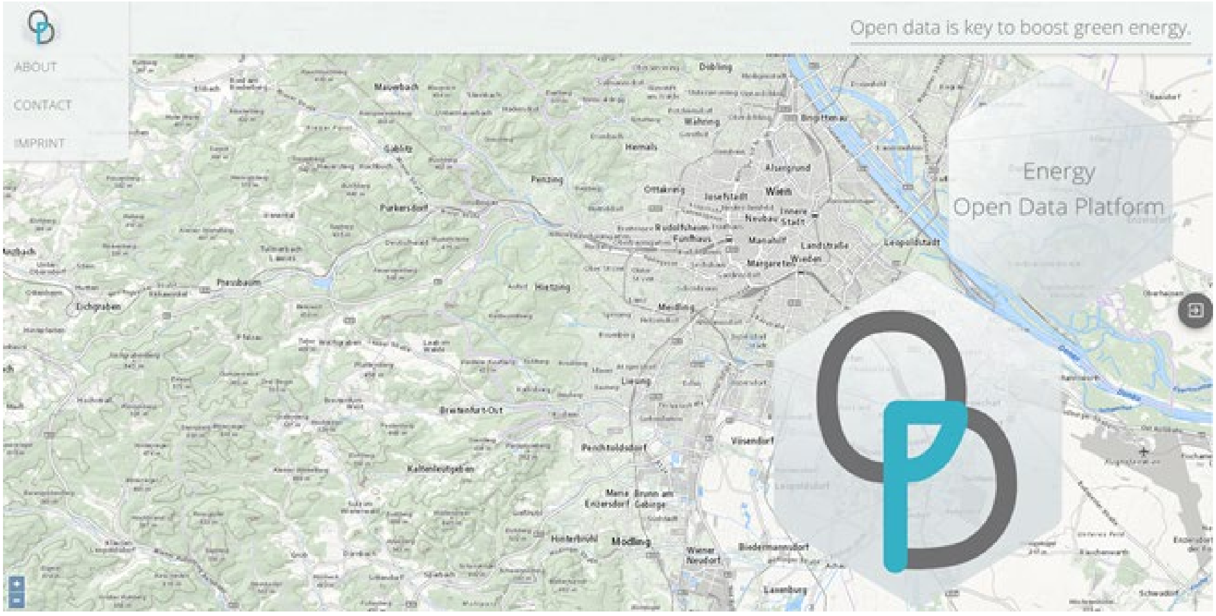


Figure 2 Landing page

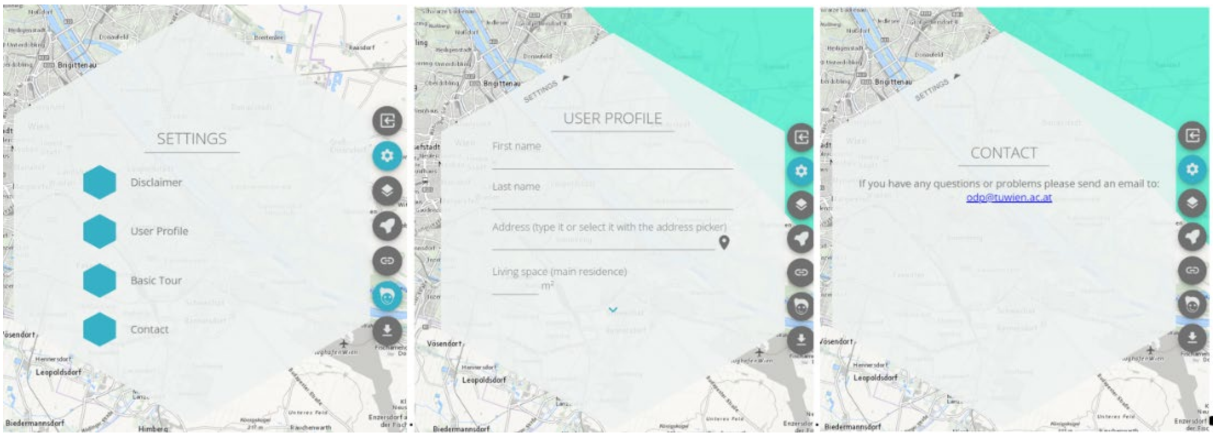


Figure 3 Settings

Figure 4 Map layer overview; example: weather data

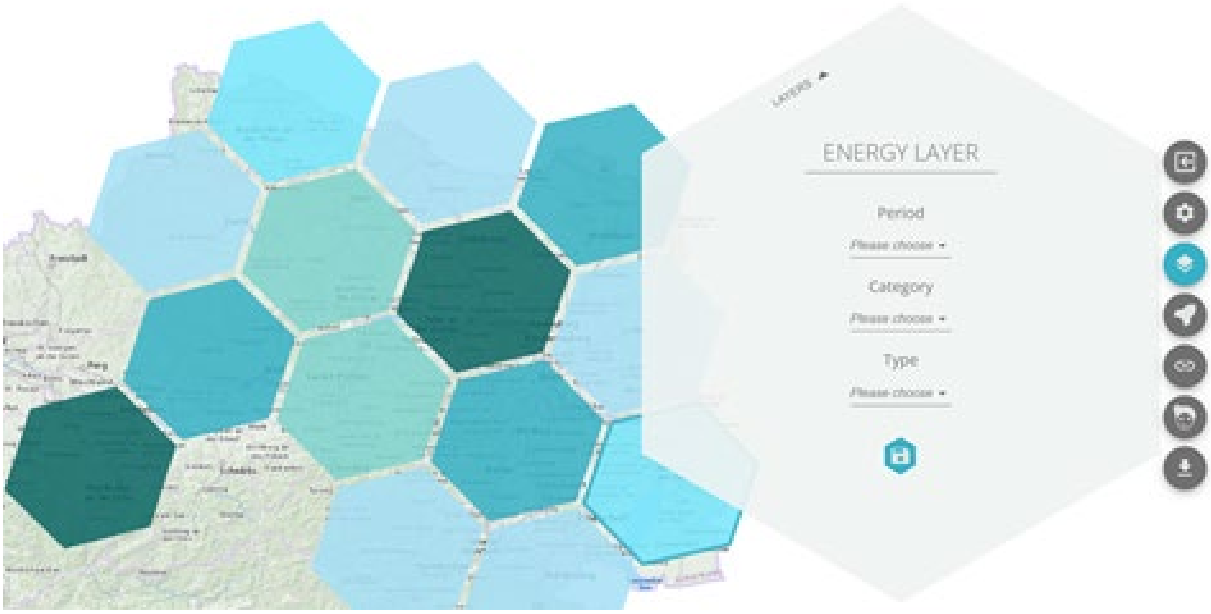


Figure 5 Map layer configuration

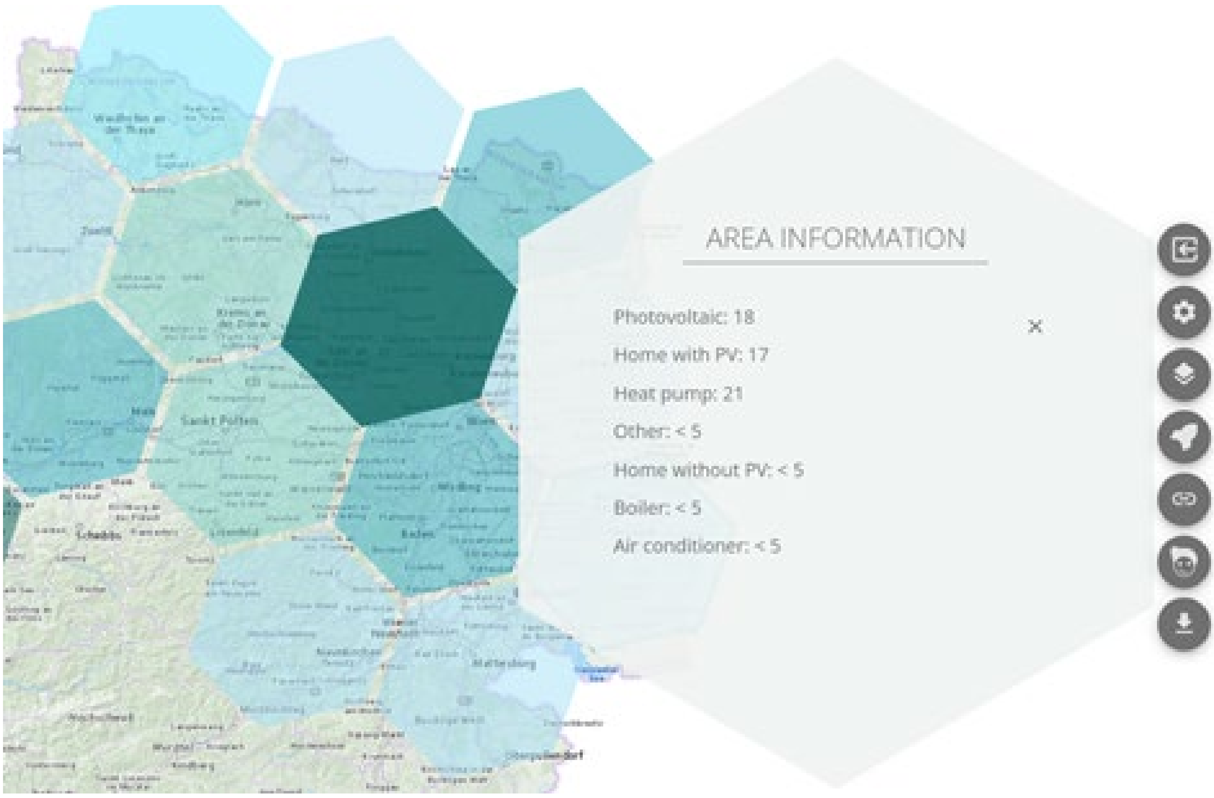


Figure 6 Map layer details

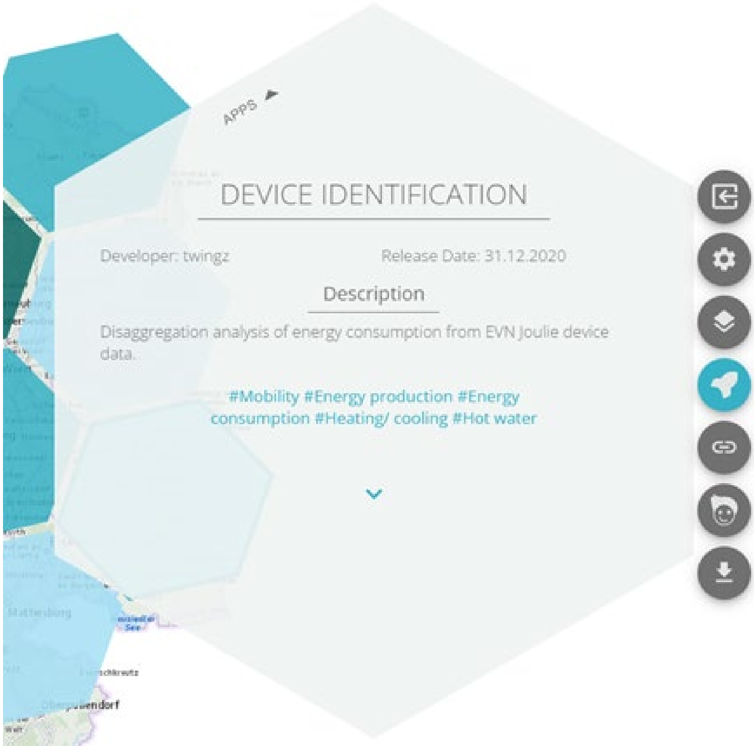


Figure 7 Market place application details

Figure 8 Account linking

Figure 9 My devices: overview, time series, combine charts, device disaggregation

6 Applications

6.1 Disaggregated Prediction Model

In this chapter the general concept and architecture of the disaggregation solution and the applied methods are described. This includes data collection methodologies and the improvement of the 1min disaggregation parameters, to the 15min disaggregation algorithm that is used in the project. The disaggregation solution is a mixture of java and python modules. For scalability and maintainability reasons a pure python-based solution (pydisag) is preferred in the future. It is proposed to rebuild the disaggregation functionality in python stepwise with a focus on 15min disaggregation first. The very first step is the construction of a test rig which allows the comparison of different model types and candidates. A quality module provides unified methods for calculating quality parameters from ground truth and the single model predictions. In the subsequent steps, additional modules for the python-based implementation of the disaggregation are built (Figure 10). A voter and merger module can process inputs of many models to build a final appliance activity proposal. The energy extraction module takes the results and generates the energy distribution prediction of whole days. The activity extraction module does the same for appliance activities.

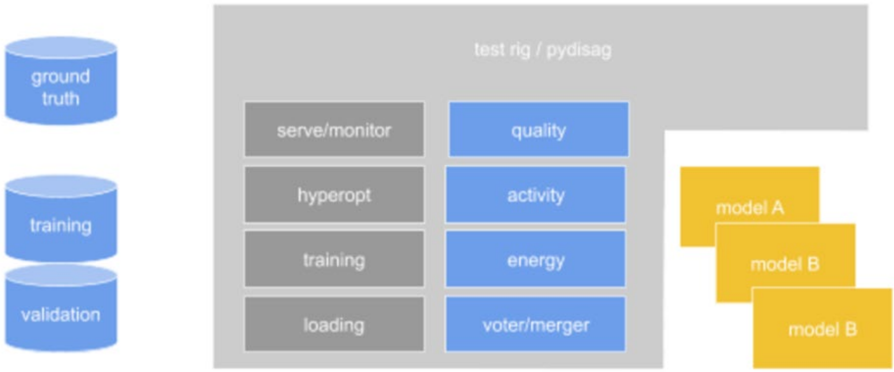


Figure 10 Model Architecture

Further support modules are planned for loading of ground truth and training data, training and retraining of models, hyperparameter optimization and model serving and monitoring. Pydisag shall be deployable via Jenkins in a docker container and the serving module shall access the twingz central database for selection of households to be disaggregated.

6.2 Heat Demand Prediction

In this chapter, AEE INTEC developed and implemented a forecasting methodology aimed at the operational needs of utilities and local energy communities. More specifically, AEE worked on a data-driven prediction software module that is based on measured heat pump electricity data. The goal of this module was to provide short-term forecasts of the electricity load attributable to domestic heat pumps. As a distinctive feature, these heat pumps are spatially distributed over a relatively large area (Lower Austria), so that local weather influence has to be considered in order to generate accurate forecasts. For data protection reasons, the exact location of individual heat pumps is not known; the given location information is restricted to the averaged location of several heat pumps aggregated over some geographical area. For this aggregation, the open source H3 framework [3] has been used, resulting in a minimum of five customers being aggregated in a hexagon. For more details on data privacy, see Chapter 4.3 of this report. Heat pump consumption data have been made available as electric power in 5-minutes resolution by project partner EVN and have been collected from the ODP server via a dedicated API. The heat pump forecasts generated by AEE are calculated in a SaaS architecture and sent back to the ODP via a targeted forecasting API. The number of included heat pump customers has increased over time and included more than 70 heat pumps. The envisaged business need has been elaborated jointly with the project consortium and the utility EVN, in particular. Forecasting the future heat pump load is interesting for utilities because heat pumps are nowadays a widely used space heating technology and account for a significant share of the total electric load. Space heating applications are often operated with a thermal storage, e.g. thermal mass of the heated building or a hot water storage tank. Making use of these storage capacities can enable electric load shifting or peak shaving in a sector coupling approach, but accurate load forecasts are an essential requirement to take advantage of this load shifting potential.

Weather is known to have a major effect on the analysed heat pump consumptions, so weather forecast data were required as a model input. The management of weather data included retrieving forecasts for all involved heat pump locations (in terms of centres of the H3 region hexagons). The fact that several heat pumps are aggregated in a potentially large H3 region introduces a location mismatch between heat pump location and weather forecast reference location, additionally to the usual weather forecasting error. In case of local weather phenomena like local fog banks, this mismatch can become substantial, but since exact locations are unknown, the mismatch error cannot be quantified. Because the exact heat pump locations were unknown during development, it was unfeasible to quantify the location mismatch effect on the forecasting error. Weather forecasts were retrieved and stored from the weather service provider openweathermap [4] for all hexagonal H3 regions. Figure 11 shows an example forecast of the heat pump electricity consumption over a 48-hour horizon and the heat pumps spread over a large geographical area. *Figure 12* shows the errors spread across several cross-validation intervals over a 48-hour horizon.

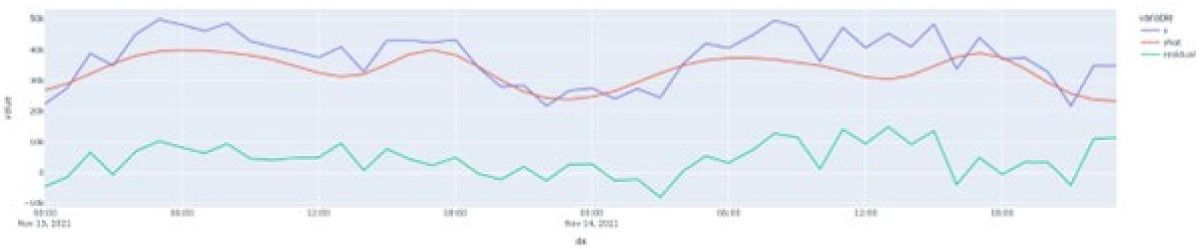


Figure 11 Example forecast of hourly heat pump consumptions

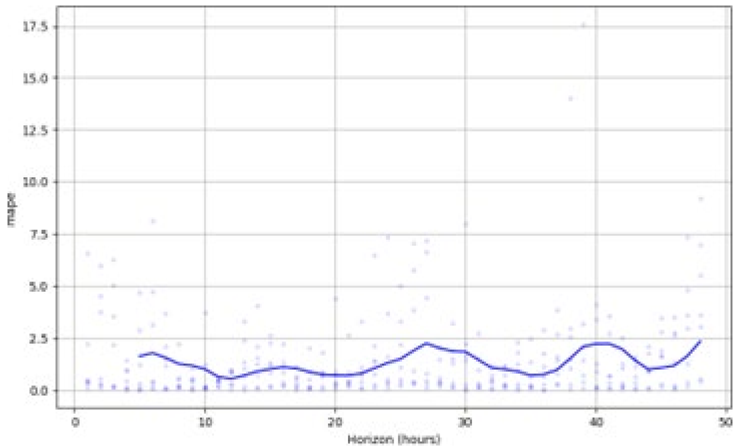


Figure 12 MAPE error spread across several cross-validation intervals

For data preparation, a standard data processing architecture has been set up, based mainly on the Python packages pandas and statsmodels and a binary data store. Data preparation steps were implemented, including data series selection, data filtering as well as treatment of outliers and imputation of missing data. Weather data were retrieved as hourly forecast values from openweathermap [4] via API calls and stored locally. Unfortunately, the openweathermap service does not include solar radiation forecasts – a presumably important model input. However, this situation is expected to change in the near future, with the Austrian national meteorological institute ZAMG planning to provide weather forecast data [5] (including solar radiation) as open data, following the new European legal framework for public sector data (PSI, Public Sector Information [6]).

In time series forecasting theory, two general approaches can be distinguished, alongside with two factions of researchers:

- machine learning approaches
- statistics-based approaches

While there is some disagreement, even acknowledged by researchers from both communities, as to which approach is generally superior, major review publications such as the M5 accuracy competition [7] suggest that both approaches include algorithms that can produce high quality results. However, forecasting algorithms have different degrees of algorithmic maturity, and time series forecasting is still a very active field of research. Especially machine learning based forecasting techniques often have intractable time and space complexities, see [8]. However, also statistical techniques such as ARIMA / SARIMAX can be intractable assuming a large number of model parameters. The forecasting model chosen by AEE INTEC for the described task is based on a GAM-type ensemble model. Overall, to the best of the authors' knowledge, there exists no mature algorithm for time series forecasting, especially when exogeneous inputs (such as weather in this case) are considered.

6.3 User and Demand Clustering

Within the scope of the project, AIT investigated whether and in what form the temporal power consumption patterns of users can be used to improve communication and optimize proactive and automatically generated behavioral recommendations.

The starting point is that in the communication with the end customer, constantly changing history or forecast curves represent an obstacle. This communication can be simplified by an abstraction to typical history curves. The cluster analysis starts with the 15-min electricity consumption values per household, collected with smart meters. For the clustering process, the data (if available for a sufficiently long time) of the last twelve months is used. On the one hand, this ensures that seasonal changes can be considered, and on the other hand, only relatively recent patterns flow into the clustering, to ensure a certain topicality and relevance of the patterns.

The next step is data cleaning, i.e. the input data is checked for completeness and plausibility, and problematic data records are removed. The analysis is always based on 24h periods, which start at 4h00 in the morning. By this choice of the observation period, the transition from one pattern to the next falls into a range with, according to experience, very low activity. This facilitates the meaningful sequencing together of different temporal profile progressions.

The next step of the analysis is to smooth the temporal consumption curves. This smoothing removes smaller (and less relevant) deviations of the consumption curves and makes it easier to identify the underlying patterns. The smoothing mechanism used is a moving average, where the curves are smoothed over 9 data points (corresponding to a 2-hour period). This parameterization to 9 data points has been shown to be most appropriate in preliminary tests.

Based on this smoothed data, a distance matrix is calculated, i.e. a distance measure is determined for each combination of available 24h periods. We use the so-called 'dynamic time warping' [9] - [10] to calculate this distance. In contrast to other distance measures (e.g. cosine, euclidean), this method allows to evaluate similar patterns with a small distance, even if they occur slightly shifted in time, compressed or stretched.

The final step is to cluster the individual 24h periods based on the distance matrix using cluster analysis. The number of target clusters is determined individually for each household in the range between 3 and 7 clusters, the exact number is determined automatically by the software using the Silhouette method.

As a result, typical temporal power consumption curves are available for each household. An analysis of data series from more than 100 households over several years has shown that this approach produces realistic and useful results. To illustrate this the following graph shows example results of this clustering process for a single household. In this case the process identified eight clusters with different temporal characteristics. One can identify different typical patterns – frequently consisting of a morning an evening peak – that can be used for improved communication with the end users and for making their consumption patterns more actionable.

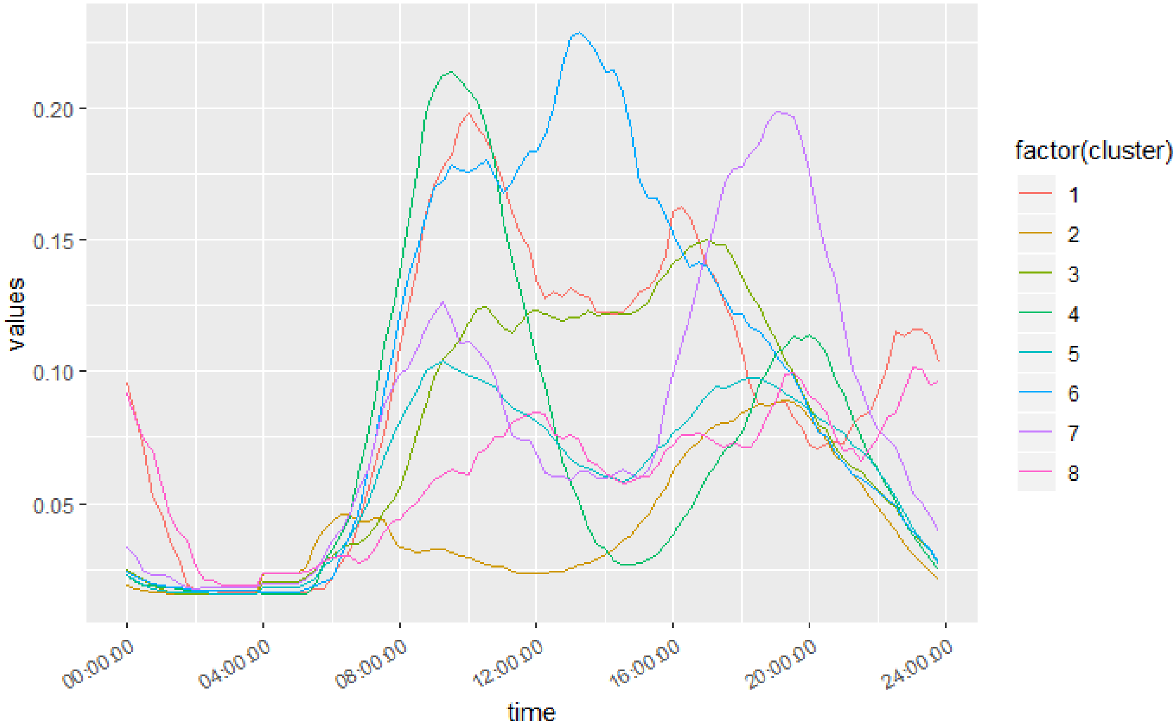


Figure 13: Clustering example for a single household

A first application scenario addresses the provision of more meaningful energy feedback to the end users. A second relevant application scenario is the improved forecast of future consumption by use of the identified patterns. As described above the combination of predicted consumption patterns and day-ahead price or sustainability information can guide the users' behavior. A logical next step is the automatically analyzation of this day-ahead information, identify possibilities for high impact and proactively inform the users regarding selected possibilities for changing behavior with high impact.

6.4 Data quality and error detection

A significant problem in power generation data sets are measurement errors because they are associated with reduced data quality. For this purpose, power measurements of 90 PV units and 92 heat pumps of 112 households in lower Austria were analyzed. The data was measured by the joulie OA from 14.10.2019 to 31.03.2021. A total of 9 480 708 measurement data of photovoltaic plants and 9 560 988 of heat pumps were evaluated.

Different errors, like missing values, zero-watt values, and wrong ordered values, occurred in the data sets. The algorithms to find these errors and replace them with approximated values were programmed in python. Different approaches were applied depending on the type of device from which the data originated. The focus was on photovoltaic systems and heat pumps as they contributed most to the household's generation and load.

An equation has been established to calculate the missing or wrong measured values of the photovoltaic system. The required radiation and weather data are retrieved from online databases. Furthermore, the nominal power of the different photovoltaic plants is needed. The 99.99 quantiles of the measured data are used to approximate these values. The orientation of the photovoltaic modules was estimated, considering the time at which the maximum power value of each day was reached. Linear regression was used for this purpose. For calculating power values, a standard module inclination angle of 35 degrees is used. Temperature and precipitation data are consulted to determine if snow on the photovoltaic modules led to lower power values in winter.

This methodology was implemented as a python PV model with the following characteristics:
If new measurement data of a PV plant are available, all timestamps are ordered correctly, and missing values are marked. Afterward, the nominal power and the orientation of the photovoltaic plant are estimated. If measurement data of a plant is present over one year for the first time, an overall evaluation starts. The expected power values of each timestamp over the entire time are calculated. In addition, calculations of the power values with a module inclination of 30 and 40 degrees and an orientation of ±30 degrees are performed. If one of these calculations is closer to the correct measured power values, the inclination and orientation are saved in a csv file. A normal evaluation is performed if measured power values are available over a shorter period than one year or an overall evaluation has already taken place. Power values are calculated with the standard module inclination angle of 35 degrees and the estimated orientation or with the stored values in the csv file. Measurement errors are marked and replaced by the calculated power values.

A different method is used for the heat pumps. To estimate the power consumption of a heat pump with missing values, the measured load profile of each day is analyzed and linked to the outside temperature, retrieved by an online database. To characterize each load profile, five specific power values were defined:

- The average power consumed each day.
- The number of measured power values from zero watts in percent.
- The percentage of power values that lies between zero and five percent of the maximum measured power value.
- The percentage of measured values between five and seventy percent of the maximum measured power value.
- The percentage value reflecting how many of the values are above seventy percent of the maximum measured power value.

The python heat pump algorithm orders all timestamps correctly and marks missing values. The maximum measured power value is saved, and the average daily temperatures are queried via an online database. All days without missing values are analyzed based on five specific power values. These values are saved and categorized into different temperature ranges. If there are missing values during a day, the five specific power values are used to find a typical load profile for the temperature on the considered day. Hence, the average of the five values is calculated in each temperature range. The average power consumed on the day plays the most important role.

The first calculations have shown that the power value calculations for the PV plants worked fine most of the time. Comparing the calculated and correctly measured power values revealed that occasionally large deviations occurred due to insecurities.
The most important ones are:

- The irradiation on the modules is not measured directly on-site but instead, averaged irradiation values from an online irradiation database are used, which can vary significantly in some cases compared to the real conditions.
- The values needed for the calculation, such as the nominal power, the orientation, and inclination of the PV systems, are not known precisely and are estimated.
- Shadow losses cannot be considered due to a lack of local information

Figure 1 shows the power production profile of a PV plant on a sunny winter day. Short-term changes in radiation at the location caused by clouds cannot be detected by the averaged radiation data of the online database. The deviations were even more frequent in unsettled weather conditions. In addition, power generation at the plant was measured for a longer time on this evening, while the radiation data in the database had already dropped to zero. Another large discrepancy between measured and calculated power values sometimes occurred in the morning, as shown with the red curve. The unrealistic shape of the power production profile happened due to a division of a cosine term when calculating the direct irradiation on the solar modules at a low elevation angle.

Numerous evaluations and analyses were carried out to improve the calculations. To avoid over-calculated power values at low elevation angles, a root function has been introduced, which is used instead of the cosine term up to an elevation angle of ten degrees. The improved power values calculations in the morning can also be seen in purple in Figure 1. In addition, two different models were tested when calculating diffuse irradiation. Analyses have shown that the power in the winter months was overestimated, so monthly correction terms were introduced. To take the degradation of the PV modules into account, a linear decrease of 0.5 % of the nominal power per year was assumed.

In summary, it turned out that measurement errors could be found much more easily if calculated power values were available to compare measured and calculated values. The adapted and improved photovoltaic model works well, and the overall evaluations improve the calculated values and reduce the uncertainty when estimating the module orientation and inclination. Finally, the photovoltaic model was successfully tested with additional new measurements of the considered plants until 25.01.2022.

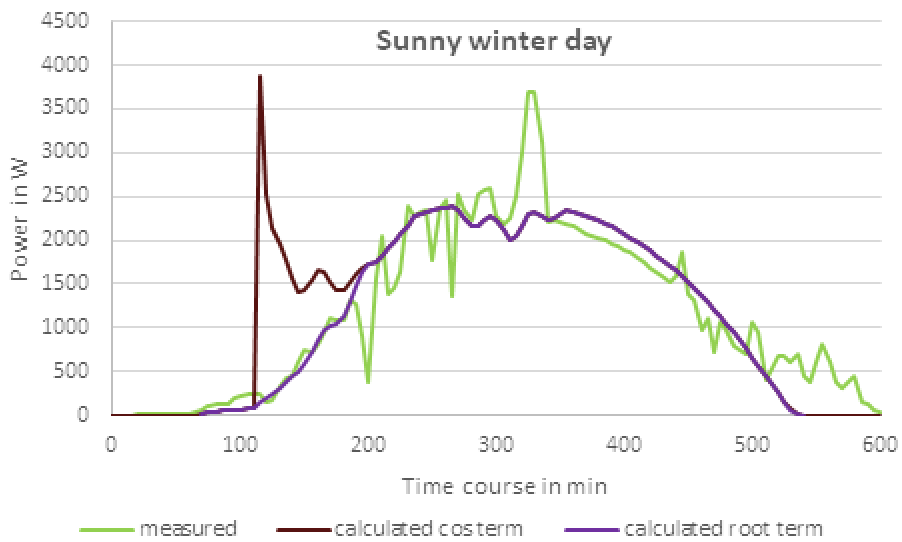


Figure 14 Power production profile of a photovoltaic plant on a sunny winter day with start time six o'clock with measured and calculated power values with cosine term and root term

The heat pump measurement data were also analyzed and evaluated in detail. The results show a clear correlation between the outdoor temperature and the consumed power of the heat pumps. The calculated five specific power values are saved into different temperature ranges. A typical load profile curve of a heat pump found on a day when there were missing measurement values after 1000 minutes is shown in Figure 2. Even if the time course of the two profiles does not match exactly, the five power values match well, and the modeled curve can be used from the time of the missing measurements onwards.

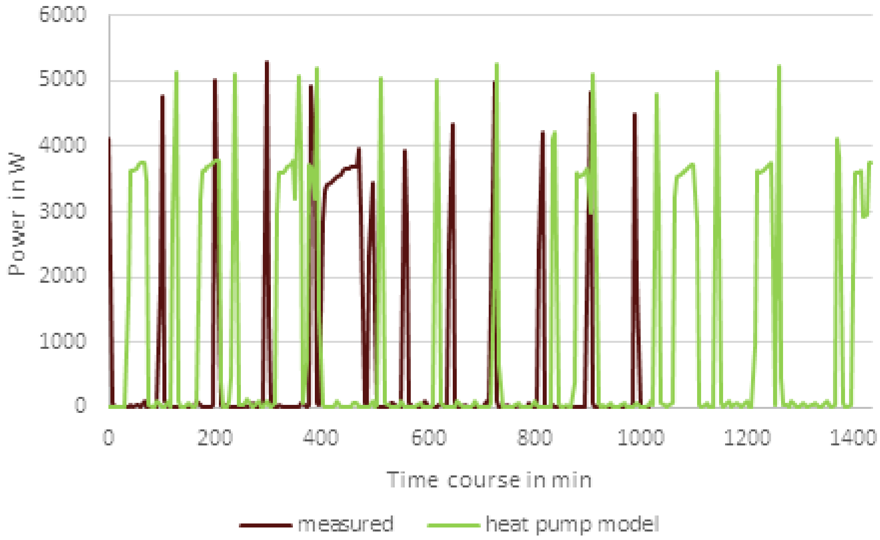


Figure 15 Comparison between measured and modeled load profile over a whole day, with the occurrence of missing data from minute 1000 onwards

It turned out that the best results were achieved by storing the five power values in twelve temperature ranges at intervals of 2.5 °C. The more measurement data were available in the different temperature ranges, the better the heat pump model worked. A distinction between days during the week and the weekend did not prove helpful, as the calculated performance values did not differ. In the end, the model was successfully tested with measurement data from the heat pumps until 25.01.2022.

7 User Profiling and Engagement

7.1 Stakeholder integration and participation models

This chapter highlights both targeted participation models and relevant stakeholders in the project region that refer to the development of the GEL-ODP.

Luyet et al. [11] proposed a framework for stakeholder participation that is structured as process. This framework is based on a literature review regarding stakeholder participation in the environmental field, whereas the framework design was derived from analyzing and discussing literature findings of good and bad procedures in stakeholder participation. The framework is considered as system with processes including inputs and outputs. Depending on the project context and stakeholder characteristics, the degree of participation is identified and subsequently the participation technique is selected. This process for stakeholder participation includes six steps: stakeholder identification, stakeholder characterization, stakeholder structuration, choice of participatory techniques, implementation of participatory techniques, and evaluation. Holifield & Williams [12] applied the framework and proposed an extension by summarizing the processes in three clusters. First, they include the stage of stakeholder recruitment as part of the identification and characterization process. Then, they summarize the stakeholder structuration, technique choice and implementation in the context of stakeholder integration. Finally, in addition to evaluation they suggest monitoring as an important aspect and label both steps as sustaining participation. For the purpose of the GEL-ODP project, the suggested participation model of Holifield & Williams (2019) was applied.

Identified stakeholder parties were private households (e.g., EVN customers, other Austrian households), science and research (e.g., universities, R&D institutions), local entities (e.g., municipalities, agencies), regional entities (e.g., federal government, Green Energy Lab), national entities (e.g., Climate and Energy Fund), community initiatives, NGOs (e.g., dsb Data Protection Authority, ARGE Daten privacy service), energy companies (e.g., utilities, grid operators), other companies (e.g., start-ups, smart system providers), ODP operators, and lawyers. Depending on the participation goal and the stakeholder group, different participation techniques were chosen, i.e., quantitative interviews (multi criteria analysis), qualitative interviews, online surveys, eye-tracking analyses, and focus group discussions.

7.2 Stakeholder engagement

The objective was to identify motives, barriers and needs for participating and using an open data platform (ODP) providing – amongst others – energy data. Several topics are closely linked to the actual use of an ODP – which have been addressed in this task. This refers to lessons learnt from other – successfully implemented – ODPs. Moreover, an ODP – like the GEL-ODP – can only be implemented if individuals are willing to share data on their energy production and consumption. To do so, individuals must use Smart Home Energy Management Systems (HEMS), like for example the Joulie Box offered by EVN, which allows to collect and subsequently provide relevant data. Moreover, individuals' experiences with HEMS provide valuable insights regarding desired services, handling, and data visualization, which can be to some extent transferred to the design of the GEL-ODP.

Consequently, the following topics and research questions have been addressed to contribute to the design and development of the GEL-ODP:

1. What ODPs do already exist in respect to energy data? How do they look like?
2. What characterizes acceptance and experience with HEMS?
   a. What are drivers and barriers behind individuals' adoption of HEMS?
   b. What can be learned from user experiences with HEMS?
3. How willing are individuals to share (energy) data?
4. What are end-users' preferences regarding potential features of an ODP?

Stakeholders have been included during the development of the GEL-ODP throughout the project, applying different techniques of stakeholder participation. The empirical studies caried out in task 5.2 allowed a characterization of several important stakeholders (e.g., EVN customers, Austrian citizens – including adopters and non-adopters, and E-Coach-users) more precisely according to the proposed stakeholder participation framework in Deliverable 5.1. In addition, the studies applied different techniques for stakeholder integration, including focus group discussions, qualitative interviews, online surveys, and questionnaires as well as eye-tracking studies, techniques that aid not only to inform stakeholders but to get in direct contact with them and thus to ensure a dialogue and an exchange of opinions, preferences, and views.

7.2.1 Energy open data platforms – a description of the status quo

As our energy system shifts to a decentralized renewable system, digital technologies are playing an increasingly important role in management and control for grid stability and optimization of supply and demand. One of these digital innovations are online platforms with the potential to facilitate interaction and communication for the exchange of information. Platforms represent services (do not have physical infrastructure) that, among other things, optimize the use of goods (e.g., car sharing), promote new collaborative lifestyles, or create opportunities for innovation. Exchanges on platforms are network-based, fostering cooperation (e.g., communities can manage and offer local resources themselves). Energy platforms are also characterized as services facilitating exchange of distributed resources and connecting actors.

In addition to the focus on integration of the platform into the infrastructure and the associated opportunities for users, energy platforms can also be characterized as carriers of data and information – emphasizing access to energy data. One approach allowing for such an open access is „Open Data". Following Wiese et al. (2019, p. 402) „openness of code and data are identified as key requirements for energy system models to comply with scientific standards like improved reproducibility and greater scrutiny". Open access to data enables co-creation, productivity, public participation, and transparency, creating legitimacy for energy policy decisions. Originally, the term open data platform was used mainly in the context of e-governance and in the field of scientific claims (transparency in scientific procedures and results). In the field of energy data, however, an open data platform can help to ensure the participation of society as a whole in the energy transition.

Eventually, ***which open data platforms do already exist in respect to energy data and how do they look like*****?** Based on first literature research regarding the characteristics of energy open data platforms, the following energy open data platforms serve as examples:

Energydata.Info [13]: As well as sharing data, the platform also offers tools to visualize and analyze energy data. Although the World Bank Group has made available various dataset and apps, external users and organizations are encouraged to contribute. The concepts of open data and open-source development are central to the project, since the platform has been developed as public good. Energydata.info uses its own fork of the CKAN open-source data portal as its web-based platform. The Creative Commons CC BY 4.0 license is preferred for data but other open li-censes can be deployed. Users are also bound by the terms of use for the site.

Energy Research Data Portal of South Afrika [14]: The Energy Research Data Portal for South Africa is being developed by the Energy Research Centre, University of Cape Town, Cape Town, South Africa. Data of South Africa and certain other African countries are provided. The website uses the CKAN open-source data portal software. Various data formats are supported, including CSV and XLSX. The site also offers an API for automated downloads. As of March 2017, the portal contained 65 datasets.

Energi DataService [15]: Energi Data Service is a free and open data portal. Here, anyone can get data about the Danish energy system such as CO2 emissions and consumption and production data. Data Service makes data available that allow you to broaden your understanding of the Danish energy system and develop new services to benefit society in general and the green transition.

Open Power Systems Data [16]: Open Power System Data is a free-of-charge data platform dedicated to electricity system re-searchers. Collect, check, process, document, and publish data that are publicly available but currently inconvenient to use. The project is a service provider to the modeling community. The platform provides data on installed generation capacity by country/technology, individual power plants (conventional and renewable), and time series data. Data are open source and available under the MIT license on GitHub. They seek to publish all data under a Creative Commons Attribution license.

Enera [17]*:* The "Smart Data and Service Platform" enables seamless interaction between the smart grid and the smart market and supports the development and operation of innovative "Energiewende Apps" and new business models. Enera demonstrates how the energy system infrastructure can be innovated to be highly resilient despite the new requirements and the variety of technologies deployed simultaneously. In addition, enera demonstrates how markets and digitalization can significantly reduce grid expansion costs and create opportunities for innovative business models. Enera Smart Data and Service Platform (SDSP):

i. Organizes data flows
ii. Enables connection of analytics applications and services
iii. Optimizes the moderation of the overall system
iv. Opportunities to realize new business models based on digitalization ("Energiewende AppStore")

7.2.2 Acceptance and experience with Smart Home Energy Management Systems

To address the question – **What characterizes acceptance and experience with HEMS?** – and to identify drivers and barriers for HEMS adoption as well as learn from users' experiences, several empirical studies have been carried out:

i. A multicriteria-analysis, addressing benefits and disadvantages (n=14)
ii. A survey investigating drivers behind individuals' adoption decision (n=521)
iii. Two eye-tracking studies (n=13/n=11)
iv. A representative survey (n=626) followed by another eye-tracking study (n=50)

(i) The AHP analysis provided information on the importance of positive and negative factors influencing the decision to use a HEMS. The AHP analysis shows that advantages about HEMS were more important to the respondents than disadvantages. This result is not surprising, since the participants have already decided to adopt and use the HEMS (i.e., "Joulie-Box"). According to Diffusion of Innovation (DOI) Theory, the phase of innovation adoption is followed by the confirmation phase, in which the user wants to have the decision confirmed (Rogers, 2003). Accordingly, the criteria of the advantage dimension were also rated more important than the disadvantages in respect to the global priority weighting. Nevertheless, it is apparent that the criteria "investment costs" as a disadvantage was more important than the advantage criterion "entertainment".

(ii) The representative survey amongst Austrian citizens shows that a positive attitude towards HEMS as well as the subjective perception that HEMS are easy to use foster adoption. Moreover, we find that the possession of a PV power plant or an electric vehicle fosters adoption.

(iii) Summarizing two use-cases (Joulie OA and twingz eCoach), the primary motive for HEMS users is to be able to read the energy data in their own household. How detailed the energy data should be, however, depends on how much the people are already concerned with the topic of energy. A forecast function, the display of user-defined time periods, and the previous savings from a PV system are further detailed information that participants would like to see within HEMS. Regarding the usability of HEMS, it can be stated that users would like to have an easy-to-read start page with the most important information on energy consumption. Depending on the user's interest, it should be possible to find more detailed information on sub-pages. Furthermore, users should be able to easily switch between different time periods (day, month, year, minutes) in graphical views. An intro with explanations before the first use of the app is also recommended. In the graphic design, care must be taken to ensure that the texts stand out from the background with sufficiently high contrast. Click fields should also be made identifiable and preferred visualization forms are bar charts or pie charts.

(iv) Regarding visualization on the ODP, results from HEMS can easily be utilized for the ODP development. Since most of the participants had no particular knowledge about energy topics, simple energy charts, especially bar charts, should be used to a greater extent. Simple charts enable users to get familiar with the topic of energy, a result that is consistent with previous study findings. Therefore, the ODP should focus on traditional chart types. Contrary, onion or rose charts not only are less common but more likely to elicit negative emotions.

7.2.3 End users' willingness to share energy data

Based on three survey studies (n=548, n=604, n=1198), several conclusions can be made on **how willing are individuals to share (energy) data**. The first survey identified user segments based on their willingness to share energy data. The results identify a primary ODP target group of the Open Minded representing a relatively large segment of 28% in total. Their characteristics indicate that people living alone (can) devote their leisure time more intensively to the topic of energy and optimizing their own household. Expressed in absolute values, this means an impressive number of 2.363 ± 81 households that are very receptive to voluntarily provide their own energy data to an ODP. As the results indicate a notable sensitivity of willingness to share energy data for a personal added value depending on privacy concerns, possible interactions between privacy concerns and other barriers to data provision remain unclear. Adequate procedures for raising the users' awareness and targeted communication must be found to clarify possible misconceptions or fears among end users. Furthermore, appropriate incentives or added values for the provision of data must be designed. This will enable further valuable data to be utilized for the deployment of an integrated and sustainable energy system. Developers and operators of ODPs must therefore form a transparent and trustworthy platform and segment potential users to elicit which type of added value increases their willingness to share energy data most.

The second survey aimed at identifying perceived risk of different data, their sensitivity and willingness to share. The results show that energy data (i.e., heating energy consumption, washing machine power consumption and PV production data) are assigned to the same cluster. In comparison to the other cluster solutions this type of data are hardly associated with any type of risk and also the perceived sensitivity of data is low. Nevertheless, the willingness to share this data is not highest. One explanation for this may be that there is still no experience with sharing energy data and that there is therefore uncertainty, which on the one hand makes it difficult to assign risk types and on the other hand minimizes the willingness to share them. One exception is smart meter data, which was assigned to a cluster solution with data like data of birth, e-mail address and car plate number. Data in this cluster were associated with physical or social risk and the perceived sensitivity and willingness to share data was rated in the midrange.
The third survey addressed users' willingness depending on the operator (i.e., whether an energy supplier, public institution or association is responsible) to share energy data. Independent from the ODP operator, respondents' willingness to share electricity data is explained by trust in said operator, as well as data-driven innovation as a perceived advantage. Interestingly, the data-driven transparency for the collective is a perceived advantage when an energy supplier operates an ODP, whilst data-driven efficiency and data-transparency for oneself as an advantage explains willingness to share electricity data on an ODP operated by a public institution.

7.2.4 End users' preferences regarding ODP features

The question **What are end-users' preferences regarding potential features of an ODP?** was addressed in a two-step approach. Drawing upon results of a focus-group discussion with energy experts (n=7) and two qualitative interviews with HEMS-users, the overall idea of the GEL-ODP is appreciated, if the ODP is easy to use and provides attractive services, whereby the comparison with other households is highlighted as one promising option. Based on a representative survey (n=720) focusing on features of an ODP suggests that private households would prefer services allowing to optimize their own energy consumption, whilst data should be visible ideally for end-users themselves. However, the option to compare the households' energy data with those of other individual households or with aggregated data of other households appears to be an interesting feature which should be provided by the ODP. This is in line with experts' opinions, who also emphasize that the ODP could be one means to raise awareness for energy savings.

7.3 Data-driven engagement

One main goal of the work within GEL Open Data Platform was to develop strategies for long-term and sustainable user engagement, based on the processing of the available data in the platform. One of the goals of the project was to explore the possibilities of using existing data from and about the users, to make the interaction more pleasant and interesting.

Change of overall household energy consumption relies on the participation of household members. But to successfully conserve energy or shift loads away from problematic into more desirable time slots a certain understanding of the impact of different loads and behaviours is required. This is relevant because even if engagement is successful and people are motivated to approach their household energy consumption actively to curtail or shift, they need to have actionable knowledge how to do so effectively. If this is not the case, there is a high risk of either behaviours with less impact being chosen, which then lead to frustration due to limited results and consequential disengagement, or to uncertainty regarding the correct behavioural changes preventing behaviour overall.

Based on the available energy consumption data of the customers automated methods to better understand and classify the temporal patterns have been developed and implemented. This improved understanding of the customers energy relevant behavior now can be used to better guide the user interaction. Users can be better addressed with relevant and interesting information, and also the timing of the interaction can be matched to the users' needs and 'rhythm'.

A scale for measuring applied energy literacy was developed. The goal here was to create awareness, provide guidance for action, and to have a tool available for baseline survey & recording of changes after interventions. The scale was developed through extensive research, expert workshops, expert evaluation and a test run. The questionnaire correctly assigns 16 items with 4 levels of energy saving potential. It was validated in an evaluation study with N=627. The instrument gives insight into strengths and weaknesses in energy literacy

A collective engagement framework has been developed, which allows the integration of social identity theories, behavior change models & persuasion strategies into the engagement process. The development of the collective reinforcement mechanisms for digitally supported behavior change were guided by an expert study (N=6) in an iterative development of designs for 6 different engagement strategies. Additionally benefit-framing designs for engagement (versions for individual/collective engagement) were developed. An experimental online study to test the effectiveness of collective reinforcement mechanisms to increase engagement (N=499) has been conducted.

7.4 ODP Evaluation

The evaluation of the social dimensions regarding the ODP took part based on

i. A target-group survey of ODP participants (n=24),
ii. A representative survey (n=670)
iii. An eye-tracking study (n=42).

First, the target-group survey results indicate that most of the respondents are engaged in the ODP via the EVN joulie for six months or more (91%). Main benefit factors include – with about 80% agreement each – a better understanding of own energy data, more awareness to use own produced energy, and PV self-consumption optimization. Then, the representative survey participants evaluate the OPD as rather aesthetic, but perceive its benefit, usability and usefulness moderately. This yields to an overall moderate intention to use the ODP. However, there is a user cluster of 21% that intends to use the ODP and asked for platform registration. Finally, the eye-tracking study confirmed the aforementioned results. Although most participants could complete all usability tasks, their completion times were not satisfactory. Major optimization potentials lie in changing the website language to German, implementing a search function, strengthening contrasts and enlarging texts.

8 Conclusio and Outlook

Within the GEL ODP project an ODP for the energy sector in order to provide easy access and overview of relevant data and interdependencies of a current and future integrated energy system was developed. Different user roles were implemented to provide suitable functionalities for different stakeholders.
It was shown that it is essential to easily integrate various data-sources via publicly available APIs and to minimise entry barriers for costumers. Fur this purpose a standardized national-wide smart-meter rollout is very important to give every costumer the opportunity to use the ODP without the need for external hardware.
APIs are a key functionality to integrate external resource-intensive applications into the platforms. This allows the ODP to perform as a scalable platform for e.g. energy consumption prediction, clustering and disaggregation. This helps to minimise and optimise the own electricity consumption.

The developed clustering concept and algorithms to cluster energy consumption curves on a household basis worked reasonably well as shown with sample data from approximately 100 households. The algorithm could detect meaningful patterns for almost all consumption histories, and qualitative review of the resulting daily consumption curves showed plausible results. We think the approach shows very promising potential to be used in customer engagement and communication activities
The surveys found out that there is very high interest in sharing their energy data and in the optimisation of their own consumption. The main benefit for many costumers was a better understanding of own energy data, more awareness to use own produced energy, and PV self-consumption optimization.
Moreover, it was shown that the visualisation style is very essential to reach a high participation rate of the platform. Especially if the rollout of an ODP should happen nationwide this will be very important to get a high participation- and usage share.

Special thanks to the associated partners Energie- und Umweltagentur NÖ, Energie Agentur Steiermark and Green Tech Cluster for supporting the project. Especially the exchange with Energie- und Umweltagentur NÖ was very helpful.

9 Bibliography

[1] „GEL ODP project website,“ [Online]. Available: https://www.gel-odp.com/.. [Zugriff am 04 2022].

[2] „EVN Joulie OA,“ [Online]. Available: www.joulie.at. [Zugriff am 04 2022].

[3] „H3 - Spatial indexing methodology,“ [Online]. Available: https://h3geo.org. [Zugriff am 04 2022].

[4] „Open Weather Map,“ [Online]. Available: https://openweathermap.org. [Zugriff am 04 2022].

[5] „ZAMG data hub,“ [Online]. Available: data.hub.zamg.ac.at. [Zugriff am 04 2022].

[6] E. Commission, „Data strategy,“ [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/open-data . [Zugriff am 04 2022].

[7] S. Makridakis, E. Spiliotis und V. Assimakopoulos, „M5 accuracy competition: Results, findings, and conclusions,“ *International Journal of Forecasting,* 2022.

[8] H. Sak, A. W. Senior und F. Beaufays, „Long short-term memory recurrent neural network architectures for large scale acoustic modeling,“ 2014.

[9] Berndt und Clifford, 1994.

[10] Müller, 2017.

[11] V. Luyet und a. et, „A framework to implement Stakeholder participation in environmental projects,“ *Journal of Environmental Management,* pp. 213-219, 2012.

[12] R. Holifield und K. Williams, „Recruiting, integrating, and sustaining stakeholder participation in environmental management: A case study from the Great Lakes Areas of Concern,“ *Journal of Environmental Management,* pp. 422-433, 2019.

[13] [Online]. Available: https://energydata.info/ . [Zugriff am 04 2022].

[14] [Online]. Available: http://energydata.uct.ac.za/dataset. [Zugriff am 04 2022].

[15] [Online]. Available: https://www.energidataservice.dk/. [Zugriff am 04 2022].

[16] [Online]. Available: https://open-power-system-data.org/ . [Zugriff am 04 2022].

[17] [Online]. Available: https://projekt-enera.de/. [Zugriff am 04 2022].

10 Contact details

Georg Lettner
TU Wien – Energy Economics Group (EEG)
+43 1 58801 370 376
lettner@eeg.tuwien.ac.at
https://eeg.tuwien.ac.at/
https://www.gel-odp.com/

Christoph Loschan
TU Wien – Energy Economics Group (EEG)
+43 1 58801 370 334
loschan@eeg.tuwien.ac.at
https://eeg.tuwien.ac.at/
https://www.gel-odp.com/

Christoph Maier
TU Wien – Institut für Energiesysteme und Elektrischen Antriebe
+43 1 58801 370142
christoph.maier@tuwien.ac.at
https://www.ea.tuwien.ac.at/

Eva Fleiß
Karl-Franzens-Universität Graz
Institut für Systemwissenschaften, Innovations- und Nachhaltigkeitsforschung
eva.fleiss@uni-graz.at
https://sis.uni-graz.at

Philip Ohnewein
AEE INTEC Institut für Nachhaltige Technologien
p.ohnewein@aee.at
www.aee-intec.at

Christian Pfeiffer
Forschung Burgenland GmbH – Research Area Energy Transition
+43 5 7705 5433
christian.pfeiffer@forschung-burgenland.at
https://www.forschung-burgenland.at/

Christian Lechner
EVN AG
christian.lechner@evn.at
https://www.evn.at/

Thomas Schlager
EVN AG
thomas.schlager@evn.at
https://www.evn.at/

Lukas Künzel
ms.GIS GmbH
Lukas.Kuenzel@msgis.com
https://www.msgis.com/

Normann Lindner
twingz development GmbH
nl@twingz.com
https://www.twingz.com

Johann Schrammel
AIT Austrian Institute of Technology GmbH
johann.Schrammel@ait.ac.at
https://www.ait.ac.at