Personal Energy Administration Kiosk application:
an ICT-ecosystem for Energy Savings
through Behavioural Change, Flexible Tariffs and Fun
Contract No 695945

Deliverable D3.2
Data Collection Guidebook

Prepared by Gerold Muggenhumer (EI-JKU)
Version 1.0: 6th of August 2017

Document control information

Title	D3.2 Data Collection Guidebook
Editor	Johannes Reichl (EI-JKU), Gerold Muggenhumer (EI-JKU)
Contributors	IJS, GreenPocket, ENAMO, 220E
Dissemination Level	☐ **CO** Confidential, only for members of the consortium (including the Commission Services) ☐ **RE** Restricted to a group specified by the consortium (including the Commission Services) ☐ **PP** Restricted to other programme participants (including the Commission Services) ☒ **PU** Public
Reviewers	☐ Baskent-Elek ☒ GreenPocket ☐ DTU ☐ IJsfontein ☒ EI-JKU ☐ RTDS ☐ ELDER ☐ Tecnalia ☒ ENAMO ☐ 220 Energia
Status	☐ Draft ☐ WP Manager accepted ☒ Co-ordinator accepted
Action requested	☐ to be revised by Partners involved in the preparation of the deliverable ☐ to be reviewed by applicable PEAKapp Partners ☐ for approval of the WP Manager ☐ for approval of the Project Coordinator
Requested deadline for Action	

Versions

Version	Date	Change	Comment/Editor
VFinal	13/07/2016	Final version	Integration of comments from reviewers

Table of Contents

1   Introduction

Within the framework of the 'PEAKapp' project, a software tool is developed which collects load profiles of households' electricity consumption as well as data on the recruited individuals' application usage behavior.The 'PEAKapp' project develops a software tool that collects load profiles of housholds' electricity consumption as well as data on how the recruited individuals behave using the application. The aim of the project is to provide a user-friendly access to the energy consuming behavior of the individuals. In particular, the app should sensitize the participants to their energy consumption and hence foster energy savings due to energy saving measures.

In this deliverable, the focus will be set on the data acquisition process necessary to enable a comprehensive and holistic analysis. For this purpose, personal anonymized data is collected from various sources. The gathered data will then be analysed and evaluated in the field test as described in deliverables D1.3 and D1.5.

How will ethical issues be managed?

Personal data, which will be collected within this project, has to be accepted by the National Data Protection authority (for more details see Data Protection). The identity of the participants will be protected by anonymization of the data. The anonymization process guarantees that no particular individual can be identified. Statistics and tables of quantitative research will be published in a manner such that it will not be possible to identify any person. Furthermore, users will be informed comprehensively and have to agree to the data collection for scientific purposes with their signature. The legal experts of this project will guarantee that the recruiting process, including the information for the participants about data protection issues, fully complies with national and EU laws.

2   Data Collection

This section gives an overview about the collected data. In the field test, data is collected from various sources. All collected data will be described in the following sections. In addition, section 3.1 gives a comprehensive overview about the collected data during the field test.

2.1   Client Data

During the registration process, every user has to provideinformation about her/his household with regard to household characteristics and the appliances connected to the electricity grid. These informations are essential for several purposes. One feature in the app is benchmarking one individual users' energy consumption with other app users. To prevent discrimination in terms of comparing a 'rich' household with a 'poor' one, an algorithm guarantees that households with similar characteristics (size, appliances, heating system, etc.) are compared.Moreover, for research purpose, it is important to know what the main drivers of the energy consumption of the individual households are. For example, households heating with heat pumps or electric heating systems tend to have a higher energy consumption than households heating with other sources. Apart from that, for the energy consumption pattern, it is essential to know whether an individual user has a photovoltaic system or an electricity storage battery, as this leads to considerable reductions in the individual energy consumption.

In this process, the users also have to confirm the decleration of consent, meaning that the energy supplier is empowered to retreive the energy consumption data quarter-hourly. The users also confirm, that the Energy Institute is authorised to use this data for research purposes in anonymized form.

Information about the collected data

- Measurement frequency: data transmission once at the registration process

- Data source: via Typeform[1]

- Receiver: Energy supplier, Energy Institute

2.2 Application Usage Data

One key element of this project is to foster the energy saving behaviour of individuals by using the PEAKapp application. To maintain a comprehensive and robust analysis of the app usage behavior, personal usage data, like the number of log-ins orthe average usage duration of the application is needed. The Google Analytics™ interface provides the data in anonymized form (individual ID's). To get access to this data, an approved account is necessary. The data will be gathered with the Excel™ add-in 'Analytics Edge™' from Analyticsedge.com via Oauth2 authentication. Beside the variables in section 3.1, Google also collects locational data in accordance with the NUTS 1geocode. Therefore, there is no exact data on the customer's location available to the consortium. Only researchers of the Energy Institute at the JKU Linz and technicians from GreenPocket have access to this data, but without knowing the individuals' name or contact data. The data will be collected on a monthly basis as well as for the number of access per date of the access. The data also include the number of postings to social media platforms, but without any content. Information on the operating system and the software is necessary for trouble shooting purposes, as the developers of the application need to know by which operating system the reported bugs occurred.

To comply with national and EU legislation, the collected data will be stored at the university's server located at the Johannes Kepler University Linz, where only researchers of the Energy Institute who are directly involved in the project will have access to.

Information about the collected data

- Measurement frequency: every app access

- Data source: Google Analytics

- Receiver: Energy Institute, Greenpocket GmbH

2.3 Energy Consumption Data

The basic prerequisite to use PEAKapp is that the consumers have so-called smart meters. The smart meter enables the transmission of the housholds' energy consumption data to the DSO[2] in very short time intervals. This collected data is essential to foster energy savings, as the consumers need to know their own energy consumption in order to react to their consuming behaviour. These data is stored directly at the DSOs' server. Before registering to PEAKapp, every user has to permit that the recorded data is available to the energy retailer in a 15-minute granularity. For PEAKapp, there is a stand-alone server also located at the DSO in the demilitarized zone. Therefore, the data transfer from the DSOs' database to the PEAKapp application server is done within the DSOs local area network. As there is an ongoing development and enhancement process, Greenpocket GmbH also has access to the PEAKapp application server for maintenance purposes.

During the project, researchers from the Energy Institute will evaluated the users' energy consumption data. According to this data, the users get different energy cost saving messages. These combined informations give

---

[1] https://energieinstitut-linz.typeform.com/to/yoaD4h

[2] Distribution System Operator

important insights on how the users behave at different energy prices and therefore deliver very important results for the projects' purpose on determining energy saving potentials due to certain cost reduction measures. In addition, constant information on energy prices in the app should also sensitize users to consume energy at times when the energy price is low.

Information about the collected data

- Measurement frequency: ¼ hourly

- Data source: Smart Meter data

- Receiver: DSO, Energy supplier, Energy Institute

3 Annex

3.1 Varialbe List

The following table contains the variables, which will be generated in the PEAKapp project. The term user is defined as follows:

Index	Symbol	Unit	Description
User identifier	i	a user	The index of a user, i.e. i=1 refers to the user number 1
Months after release	m	number of months	Counts the months since the release of the app, with m = 1 as the starting point.
Months, after the download of the app	t	number of months	t gives the time in months after the download of the app. Example: User has downloaded the app in January (month of download = January), than for this user, January t = 1, February is t = 2, ...

*The PEAKapp **user** references one account to the PEAKapp system, where one account is presenting one household.*

Variable Name	Var ID	Unit of measurement	Unit of observation	Description
Client Data				
Number of Residents	Res_i	# of residants	per user	This variable contains the number of residents living in the household of each user.
Living area	LA_i	square metre	per user	This variable contains the information about the living space of the household in m².
House type	HT_i	type of the residence	per user	This variable contains the information about the type of house, eg. single-family house, multi-family house,…
Heating System	HS_i	type of the heating system	per user	This variable contains the type of heating system of the users' household. Choosable types: heatpump, electric heating, biomass, firewood, natural gas, district heating, oil heating, unknown.
Warm Water Generation	WWG_i	type of the warm water generation system	per user	This variable contains the type of warm water generation unit. Choosable types: heatpump, electric boiler, flow heater, solar heat, biomass, firewood, natural gas, district heating, oil heating, unknown.
Number of electronic household appliances connected	$HA_{i,x}$	# of appliances	per user	This variable contains the number of electrical household appliances of type used per household. Chooseable types: fridge, freezer, washing machine, dishwasher, tv, computer, dryer, swimming pool, aquarium, waterbed, sauna, air condition, electric vehicle.
Electrical Power and Storage Equipment	$EPS_{i,x}$	kW	per user	This variable contains the information about the availability of an electrical production and storage equipment and the corresponding power in kilowatts.
App Installation				
Number of downloads at month m	D_m	# in M1, # in M2, …	# of users	This variable contains the number of downloads per month.
Number of users in month t after the month of the download	$D_{t,m}$	# in M1, #in M2, …	# of users	This variable contains the number of downloads in month t after the month of the download

Dates of the downloads per user	DD_i	Date	per download per user	This variable contains the date of download.
OS Type	OS_i	Type, e.g. Android	per download per user	This variable contains the Type of Software per user i with the download date
App Version	AV_i	Version, e.g. PEAKapp v1.05	per update per user	This variable contains the new versions/updates per user i.
App usage in general				
Active users in month m (m=months since app release, m=1: month off app release)	A_m	# in M1, #in M2, …	# of users	This variable contains the number of active users in month m
Number of log-ins per user per /month	NL_i	# in M1, #in M2, …	per user	This variable contains the number of log-ins per user i in month m.
Active users in month t after the month of the download	A_t	# in M1, #in M2, …	# of users	This variable contains the number of active users in the app in month t after the month of the download of the app.
Session Duration of users in month m	SD_m	Minutes	# of users	This variable contains the session duration of the app users in month m.
Session Duration of users in month t after the month of the download	SD_t	Minutes	# of users	This variable contains the session duration of the app users in month t after the month of the download.
Date, Daytime of each app use	DU_i DTU_i	Date, daytime	per access of the app	This variable contains the Date/Daytime of the log-in into the app per user i.
Consumption information				

Number of accessing the normal consumption information per month (i.e. all pages related to Cluster 1 as in the GA)	$NNC_{i,m}$	# in M1, # in M2, …	per user	This variable contains the number of accesses into the normal consumption information per user i in month m.
Date, Daytime of accessing Cluster 1 pages	DNC_i $DTNC_i$	Date, daytime	per access of Cluster 1 pages	This variable contains the Date, Daytime of accessing Cluster 1 pages per user i.
Duration of accessing Cluster 1 pages	$SDNC_i$	Minutes	per access of Cluster 1 pages	This variable contains the session duration of accessing Cluster 1 pages per user i in month m in minutes.
Number of accessing a Benchmark per month (i.e. all pages related to Cluster 2 as in the GA)	$NBM_{i,m}$	# in M1, # in M2, …	per user	This variable contains the number of accesses in the Benchmark per user I month m.
Date, Daytime and Type of requesting a benchmark.	DBM_i $DTBM_i$ TBM_i	Date, day time, type	Date, daytime, type	This variable contains the Date, Day time and type of requesting a benchmark per user i. The Type of a benchmark is its respective timeframe, i.e. whether a monthly benchmark was requested, or a weekly benchmark, or….
Duration of requesting a benchmark	$SDBM_i$	Minutes	per request of a Benchmark	This variable contains the session duration of requesting a benchmark per user i in month m in minutes.
Price-related information				
Widget for retrieving price information	PI_i	1 / 0	per user	This variable contains the number of users who use a widget for retrieving price information with 1 = user uses widget and 0 = user don´t uses widget.

The finally applied prices/discounts	FP_i	user price per hour during the field test (around 8600 values)	one file for the Austrian field test	This variable contains the end-user price/discount per hour during the field test.
Serious Gaming				
Number of accessing the serious game per month (i.e. all pages related to Cluster 4 as in the GA)	$NSG_{i,m}$	# in M1, # in M2, …	per user	This variable contains the number of accesses into the serious game per user i in month m.
Number of total points scored per user, per month	$NTP_{i,m}$	# in M1, # in M2, …	per user	This variable contains the number of total points scored per user i in month m.
Number of total points scored by tweaking per user, per month	$TTP_{i,m}$	# in M1, #in M2, …	per user	This variable contains the number of total points scored by tweaking per user i in month m.
Number of accessing the hints	$NH_{i,m}$	# in M1, #in M2, …	per user	This variable contains the number of accesses of the hints per user i in month m.
Date, Daytime of using the serious game	DSG_i	Date, daytime	Date, daytime	This variable contains the Date/Daytime of using the serious game.
Duration of using the serious game	$SDSG_i$	Minutes	per user	This variable contains the duration session of using the serious game per user i in month m.
All stats earned/achieved/… in the serious game	ST_i	Date, daytime	per user	This variable contains all stats earned/achieved/... in the serious game per user i in month m.
Date of setting a bet	DB_i	Date, daytime	per user	This variable contains the date when a bet was placed per user i.
Bets in month m	B_m	# in M1, # in m2, …	# of users	This variable contains the number of bets in month m.

Bets in month t after the month of the download	B_t	# in M1, # in M2, …	# of users	This variable contains the number of bets in t month after the month of the download.
Social Networking				
Monthly number of facebook posts done from PEAKapp	$FP_{i,m}$ $FP_{i,t}$	# in M1, # in M2, …	per user	This variable contains the monthly number of facebook posts done from PEAKapp
Date and Type of facebook post	DFP_i TFP_i	Date and type	per facebook post	This variable contains the Date and Type per facebook posts. Type can be a consumption information, a benchmark, money saved through dynamic prices over a certain time period, a game score, etc
Push messages				
The sent push messages	PM_i	text	per message per user	This variable contains the sent push messages per message/user i
The date and time of the sent push messages	DPM_i $DTPM_i$	Date and time	per message per user	This variable contains the date and time of the sent push messages per message/user i
Others				
Total Target Group	N	Number of People	Customer of Energy supplier with Smart Meters within a certain cohort	This variable contains the total number of the target group of the energy supplier.
Call Center calls	CC	Quantitative: Amount of calls Qualitative: Type of calls (complains, questions, improvement suggestions,…)	# of call center call	This variable contains the total number of call center calls, which are related to the PEAKapp project.
Meter Data	SMD_i	Load profile	per user	This variable contains the smart meter data per user i.