



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

## **Deliverable 7.2**

# **Data Management Plan**

Prepared by Johannes Reichl, Christopher Binder (EI-JKU)  
Version 2.0: 08<sup>th</sup> of October 2016



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 695945.



Document control information											
Title	Del 7.2 – Data Management Plan										
Editor	Johanes Reichl										
Contributors	Christopher Binder										
Dissemination Level	<input type="checkbox"/> <b>CO</b> Confidential, only for members of the consortium (including the Commission Services) <input type="checkbox"/> <b>RE</b> Restricted to a group specified by the consortium (including the Commission Services) <input type="checkbox"/> <b>PP</b> Restricted to other programme participants (including the Commission Services) <input checked="" type="checkbox"/> <b>PU</b> Public										
Reviewers	<table border="0"> <tr> <td><input type="checkbox"/> Baskent-Elek</td> <td><input checked="" type="checkbox"/> GreenPocket</td> </tr> <tr> <td><input type="checkbox"/> DTU</td> <td><input checked="" type="checkbox"/> IJsfontein</td> </tr> <tr> <td><input checked="" type="checkbox"/> EI-JKU</td> <td><input type="checkbox"/> RTDS</td> </tr> <tr> <td><input type="checkbox"/> ELDER</td> <td><input type="checkbox"/> Tecnalia</td> </tr> <tr> <td><input type="checkbox"/> ENAMO</td> <td><input type="checkbox"/> 220 Energia</td> </tr> </table>	<input type="checkbox"/> Baskent-Elek	<input checked="" type="checkbox"/> GreenPocket	<input type="checkbox"/> DTU	<input checked="" type="checkbox"/> IJsfontein	<input checked="" type="checkbox"/> EI-JKU	<input type="checkbox"/> RTDS	<input type="checkbox"/> ELDER	<input type="checkbox"/> Tecnalia	<input type="checkbox"/> ENAMO	<input type="checkbox"/> 220 Energia
<input type="checkbox"/> Baskent-Elek	<input checked="" type="checkbox"/> GreenPocket										
<input type="checkbox"/> DTU	<input checked="" type="checkbox"/> IJsfontein										
<input checked="" type="checkbox"/> EI-JKU	<input type="checkbox"/> RTDS										
<input type="checkbox"/> ELDER	<input type="checkbox"/> Tecnalia										
<input type="checkbox"/> ENAMO	<input type="checkbox"/> 220 Energia										
Status	<input type="checkbox"/> Draft <input type="checkbox"/> WP Manager accepted <input checked="" type="checkbox"/> Co-ordinator accepted										
Action requested	<input type="checkbox"/> to be revised by Partners involved in the preparation of the deliverable <input type="checkbox"/> to be reviewed by applicable PEAKapp Partners <input type="checkbox"/> for approval of the WP Manager <input type="checkbox"/> for approval of the Project Coordinator										
Requested deadline for Action	31/10/2016										



# Content

- 1 Introduction and Motivation..... 5**
- 2 Data Management Plan..... 5**
  - 2.1 Administrative Information of the DMP..... 5
  - 2.2 Data Collection..... 5
  - 2.3 Documentation and Metadata ..... 6
  - 2.4 Ethics and Legal Compliance..... 6
  - 2.5 Data Sharing ..... 7
- 3 Annex ..... 8**
  - 3.1 Variable List ..... 8

## 1 Introduction and Motivation

PEAKapp targets the development of an app to trigger lasting energy savings through behavioural change and continuous engagement, and to enable increased consumption of clean and low priced electricity from the spot market for household customers. The analysis of the impact of the app needs detailed research data about energy consumption, app usage, and contract and tariff information.

Deliverable 7.2 develops a data management plan which outlines most important points of how to handle the research data from the data generation process until data sharing. The subsequent chapters detail about the data management plan.

## 2 Data Management Plan

The data management plan (DMP) presented in this deliverable contains the data generation process, data documentation, ethical and legal compliances and data sharing process. This documentation is necessary to make it possible for secondary users to understand the data, to ensure ethical standards of research and for internal information.

### 2.1 Administrative Information of the DMP

<b>Project Name:</b>	PEAKapp
<b>Project Identifier:</b>	695945
<b>Principal Investigator / Researcher:</b>	Dr. Johannes Reichl
<b>Researcher ID:</b>	orcid.org/0000-0003-1054-9454
<b>Project Data Contact:</b>	Dr. Johannes Reichl
<b>Date of first Version:</b>	07. August 2016
<b>Date of first Update:</b>	tbd
<b>Responsibility of the Data Management:</b>	Energieinstitut at the Johannes Kepler University Linz
<b>Data owner:</b>	Energieinstitut at the Johannes Kepler University Linz

### 2.2 Data Collection

This section gives a description of the data and the data generation process. It is explained what type of data will be collected and the process how it will be collected.

The PEAKapp project develops a software tool collecting load profiles of households' electricity consumption, transforms this rather technical data into user friendly aggregates, and feeds the gathered information back to the households in a way that is engaging and motivating for energy efficiency increases. By matching the households' electricity consumption data with the app usage data it will be analysed if the tool stimulates energy efficient behaviour and decisions. To collect the necessary information from the households PEAKapp carries out field tests in four European countries, Austria, Estonia, Sweden and Latvia.

#### What data will be collected?

In this project data from the following categories is collected:

- **App usage data:** Data which collects types of interaction of the user with the app which allows analysing the consumer interaction with the app, (e.g. session duration, number of bets,...). For more information see in the Annex Variable List.
- **Client data:** Personal information about the app users, descriptive data of the households (e.g. household size,...) To protect the identity of the app user this data will be anonymised.
- **Contract data:** Information about energy contract between the household and the energy retailer. It gives information about the type of tariff which the household has, duration of the contract, contract number and tariff data.
- **Meter data:** Smart meter data which shows energy consumption of the household in continuous time intervals (i.e. load profiles).
- **Survey data:** Field test consumers will be asked to complete small online surveys sent to them through the app. These surveys are designed to better understand which actions have been taken by the respondents to reduce their energy consumption. To protect the identity of the app user this data will be anonymised

Furthermore the datasets will be merged. For detailed information about the variables see the Annex 3.1.

#### How will the data be collected or created?

The data collecting process runs automatically via a specified software interface when users are participating in the field tests.

## 2.3 Documentation and Metadata

This section contains the documentation that will accompany the data to help secondary users to understand the data.

#### What documentation and metadata will accompany the data?

Following information about the dataset will be provided:

<b>Where can the data be found?</b>	The data will be provided in an open source repository, after approval of the consortium. Most likely the publication of the data will be done after the end of the project runtime.
<b>Who created or contributed the data?</b>	220 Energie OÜ, GreenPocket, IJsfontein, ENAMO, EI-JKU
<b>Data Title and date of creation</b>	In 2018, date and title will be defined in an update of the DMP
<b>Requirements to access the data</b>	Project members can access the data every time without any requirements. External people can access the data after the publication.

Necessary information for secondary users to understand the data is provided in the Annex Variable List

## 2.4 Ethics and Legal Compliance

This section deals with ethical and legal compliance issues, like the consent for data preservation and sharing, protection of the identity of participants and how sensitive data will be handled to ensure it is stored and transferred securely.

Data protection and good research ethics are major topics for the consortium of this project. Good research ethics meet all actions to take great care and prevent any situation where personal information could get misused. This is what the consortium wants to guarantee for this project.

Research data which contains personal data will just be disseminated for the purpose for which it was specified by the consortium. Furthermore all processes of data generation and data sharing have to be documented and approved by the consortium to guarantee highest standards of data protection.

### **How will you manage any ethical issues?**

Personal data which will be collected within this project are to be confirmed by the National Data Protection authorities, while the approvals of the respective authorities was outstanding at the time of the submission of the DMP. The identity of the participants will be protected by anonymization of the data. The anonymization process guarantees that no particular individual can be identified anymore. 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 this scientific purpose 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. To make sure that the recruitment process of research participants is transparent and traceable, all actions which are made to identify/recruit research participants will be documented in a recruitment plan.

## **2.5 Data Sharing**

This sections deals with the issue how the data will be shared and under which conditions.

### **How will the data be shared?**

The results of the research activities within PEAKapp, like consumer behaviour studies, will be published in peer reviewed journals according to the knowledge management strategy (e.g. open access).

For verification and potential re-use, the datasets forming the basis of these publications will be made available as supplemental information in the journals database (if requested by the editor) only after a comprehensive anonymization of the data **not allowing the identification of individuals or households**. The data transfer and system communication can be realized through various ways of data exchange. Data which are crucial to commercial exploitation and whose release would infringe trade secrets of the developed technology will not be made accessible to the public.

### **Are any restrictions on data sharing required?**

The publishing platform and publishing date will be decided by the consortium.

## 3 Annex

### 3.1 Variable List

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

*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
<b>Household characteristics</b>				
Number of residents in the household	$R_i$	# of residents	Per user	This variable contains the number of permanent residents in household $i$ in month $m$ .

Type of building	$H_i$	Indicator	per user	This variable contains the type of building user $i$ lives in, either a single family house, a terraced house, or an apartment.
Living space	$SQ_i$	# of square meters	per user	This variable contains the number of square meters of user $i$ 's dwelling.
Main heating system: electric	$MH_i$	Indicator	per user	This variable contains whether user $i$ 's main heating system is electric.
Hot water system: electric	$MW_i$	Indicator	per user	This variable contains whether user $i$ 's hot water system is electric.
Air conditioner	$AC_i$	Indicator	per user	This variable whether household $i$ has air conditioning.
Tumble Dryer	$TD_i$	Indicator	per user	This variable whether household $i$ has a tumble dryer.
PCs, laptops, tablets, game console	$PC_i$	# of devices	per user	This variable contains the number of respective devices of user $i$ .
Swimming Pool	$SP_i$	Indicator	per user	This variable whether household $i$ has a swimming pool.
Aquarium	$AQ_i$	Indicator	per user	This variable whether household $i$ has an aquarium.
Stand-alone deep freezer	$DF_i$	# of devices	per user	This variable contains the number of respective devices of user $i$ .
Waterbed	$WB_i$	Indicator	per user	This variable whether household $i$ has a water bed.
Sauna	$SA_i$	Indicator	per user	This variable whether household $i$ has a sauna.
Photovoltaic panel	$PV_i$	$kW_{peak}$	per user	This variable contains the $kW_{peak}$ installed capacity of user $i$ .

Home battery storage	$HB_i$	storage capacity in kWh	per user	This variable contains the storage capacity in kWh of user $i$ .
Electric car	$EC_i$	# of vehicles	per user	This variable contains the number of respective vehicles of user $i$ .
Electric motorbike	$EM_i$	# of vehicles	per user	This variable contains the number of respective vehicles of user $i$ .
Electric bicycle, segway, etc.	$EB_i$	# of vehicles	per user	This variable contains the number of respective vehicles of user $i$ .
<b>App Installation</b>				
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.
Number of electronic appliances connected	$NA_i$	# of appliances	# of appliances per user	This variable contains the number of appliances per user $i$ .
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.
<b>App usage in general</b>				
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.
<b>Consumption information</b>				
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.
<b>Price-related information</b>				
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.
<b>Serious Gaming</b>				
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	$ST_i$	? tbd	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.
<b>Social Networking</b>				
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
<b>Push messages</b>				
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
<b>Others</b>				
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 $i$ .