Introduction to Meter-Data

Training Document

Rev 0.1 - 28 Jun 2016

1 Introduction

This document explains

- the types of data collected as part of the Meter project;
- how these data are structured, stored and retrieved; and
- the tools used to manage the data.

This is a live document. For the latest version, please visit http://www.energy-use.org/files/MeterDataIntro.pdf.

2 What data are collected?

Meter collects three types of data:

1. Household and individual survey information
2. Individual activity information
3. Electricity readings at household level

Each of these data are explained in more detail here.

2.1 Household and individual survey information

Survey information is grouped into three tables:

Contact: This is personal information with restricted access (all other data will become public). It is provided by participants when registering via http://www.energy-use.org/hhq.php.

Household: Socio-demographic information about household composition (number of people, ages, housetype, appliances, electricity bills income). This table

Contact
- idContact INT(11)
- Name VARCHAR(45)
- Surname VARCHAR(45)
- Address1 VARCHAR(150)
- Address2 VARCHAR(45)
- Town VARCHAR(45)
- Postcode VARCHAR(8)
- email VARCHAR(45)
- phone VARCHAR(45)
- recorded TIMESTAMP
- Indexes

Household
- idHousehold INT(11)
- Contact_idContact INT(11)
- security_code VARCHAR(45)
- page_number INT(3)
- people INT(3)
- age_group1 INT(3)
- age_group2 INT(3)
- age_group3 INT(3)
- age_group4 INT(3)
- age_group5 INT(3)
- age_group6 INT(3)
- pet1 INT(3)
- pet2 INT(3)
- pet3 INT(3)
- pet4 VARCHAR(45)
- p6pm INT(3)
- house_type INT(3)
- house_type_other VARCHAR(45)
- rooms INT(3)
- own TINYINT(1)
- appliance_bo VARCHAR(45)
- appliance_b11 INT(3)
- appliance_b10 INT(3)
- appliance_b9 INT(3)
- appliance_b8 INT(3)
- appliance_b7 INT(3)
- appliance_b6 INT(3)
- appliance_b5 INT(3)
- appliance_b4 INT(3)
- appliance_b3 INT(3)
- 22 more...
- Indexes

Electricity
- idElectricity BIGINT(20)
- Meta_idMeta INT(11)
- dt DATETIME
- Watt FLOAT
- Indexes

Meta
- idMeta INT(11)
- Household_idHousehold INT(11)
- SerialNumber BIGINT(15)
- CollectionDate DATE
- DataType VARCHAR(3)
- Indexes

Activities
- idTimeUse INT(11)
- Meta_idMeta INT(11)
- dt_activity DATETIME
- dt_recorded DATETIME
- tuc INT(11)
- activity VARCHAR(45)
- location INT(3)
- enjoyment INT(1)
- Indexes

Individual
- idIndividual INT(11)
- Meta_idMeta INT(11)
- Gender INT(1)
- Age_range INT(1)
- Occupation INT(1)
- WorkingHours INT(1)
- WorkRegularity INT(1)
- WorkHome INT(1)
- BillContributer INT(1)
- Income INT(1)
- UseDishwasher INT(1)
- UseWashingMachine INT(1)
- UseTumbleDryer INT(1)
- UseShower INT(1)
- UseVacuumCleaner INT(1)
- TimeMobile INT(1)
- TimeScreens INT(1)
- TimeCooking INT(1)
- TimeExercise INT(1)
- TimeOfDiary INT(1)
- Rushed TINYINT(1)
- UnusualDay TINYINT(1)
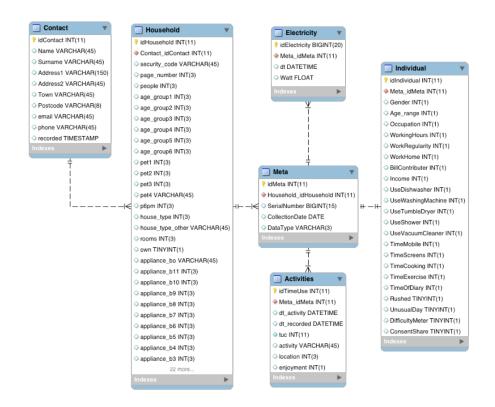- DifficultyMeter TINYINT(1)
- ConsentShare TINYINT(1)
- Indexes

Figure 1: Database schema for Meter Data

also keeps the 'date_choice' - the preferred date for the study, and 'status' (see Section 'Status progression').

Individual: Collected on the study day via the booklet or app. Covers individual information, such as age, occupation, working hours, use of appliances. . .

2.2 Individual activity information

Recorded via booklet or app. Each entry contains:

Time (dt_activity): Freely chosen by participant (unlike in conventional Time-use research, where 10 minute windows are prescribed). Certain times of particular interest will be suggested in the app (e.g. 5:30pm, 6pm, 8am). 'dt' stands for DateTime in the format '2012-11-13 23:59:59'.

Time recorded (dt_recorded): The time when the entry was made. In case of the app, entries can be back-dated (what did you do 20 minutes ago?). Booklets may be coded up days later. Entries made via the web-interface later still. This time stamp helps to identify 'how devoiced' the entry is from the actual event.

Time use code (tuc): Numeric code based on an extended version of the Harmonised European Time-Use Survey codes (eurostat (2014)). See field 'ID' in activities.json.

Activity: Plain text description of the activity. See field 'title' in activities.json.

Location: Numeric code for location:

1. home
2. travelling
3. work
4. public place
5. outdoors
6. other
7. not specified

Enjoyment: Numeric code for enjoyment:

1. not at all
2. not very much
3. neutral
4. somewhat
5. very much
6. not specified

2.3 Electricity readings

Collected via current clamp connected to an Android phone using DMon software. Recordings are taken every second and stored in a csv file with two columns: a DateTime stamp '2012-11-13 23:59:59') and the reading in Watt. The file name refers to the Meta table 'idMeta' value followed by underscore and a sequential number for each time readings were started for this device with this 'idMeta'. Often the xxxx__2.csv file contains the 'real' readings, because the phone was started up for testing before shipment. A xxxx_03.csv file is created when starting up the phone to process the data.

When uploading to the database this file is copied to the server first to speed up the database processing. Each reading is a row with an ID (idElectricity) the reference to the Meta entry (Meta_idMeta), the DateTime stamp and the reading value in Watt.

3 Database structure

The Meta table is central to the data structure as shown in Figure 1. The name is perhaps an unfortunate left over from early trials with data structure. It could equally be called the 'devices', 'instruments' or 'study events'.

Each row in this table has a unique ID (idMeta) and represents either a booklet or a phone that has been sent out. The entry is created as part of the equipment set up. Each household gets one entry for the electricity recorder (eMeter) they receive and one for each booklet or activity app. DataType for an eMeter is 'E', whereas a booklet or device with the activity app is labelled 'A'.

When devices are returned the CollectionDate, which is 'NULL' until now is updated.

Individual and Activity data share the same idMeta value, because they come from the same instrument and can thus be linked. Individual/Activity data and Electricity readings can be linked to their Household via the Household_idHousehold reference in their Meta entry and vice versa: To find all Activities for a given Household with the ID 1234 one could use the following SQL statement:

```
SELECT idMeta FROM Meta WHERE Household_idHousehold = 1234 AND
DataType = 'A';
```

This will return one idMeta for each individual which has completed a study day. The activities can then be retrieved for each of these returned values. If 6789 was returned, the activity record can be accessed with:

```
SELECT dt_activity,activity FROM Activities WHERE Meta_idMeta =
6789;
```

This will produce a list of DateTime values and the description of the activity.

4   Further information

The data handing is principally conducted with the Meter Interface. This repository includes a dummy database file in /dbDummy which can be used for experimenting. To set this up you require a working MySQL version on your machine. Create a user and grant privileges to the data.

To gain access to the live Meter database on http://www.energy-use.org you will need to be granted access.

eurostat. 2014. "Harmonised European Time Use Surveys." 2008 guidelines. Office for Official Publications of the European Communities.