Project number: 7123420

LOOPER

LEARNING LOOPS IN THE PUBLIC REALM

WP2. Data collection and visualization framework
T2.2. Set-up of the LOOPER database

Deliverable D 2.1

REPORT ON DATA COLLECTION PROCEDURE FRAMEWORK

ANNEX 3: INSTRUCTIONS FOR DATA COLLECTION AND DATA RECORDING

Version: 0.1

Date: 04 May 2018

Responsible partner: IUAV

Authors: Massimiliano Condotta (IUAV), Chiara Scanagatta (IUAV)

The project is supported by the Brussels Capital Region – Innoviris (Belgium), Ministero dell'Istruzione dell'Università e della Ricerca (MIUR) (Italy), the Economic and Social Research Council (UK) and the European Union.

DOCUMENT CHANGE RECORD

Version	Date	Status	Author	Description
0.1	04May2018	Draft	MC, CS	First internal draft

TABLE OF CONTENTS

LIST OF FIGURES 4
LIST OF TABLES 4
1. RULES AND SCHEMA FOR DATA COLLECTION AND DATA RECORDING 5
2. STATIONARY DATA 6
2.1. File name 7
2.2. “Measuring spot” table 7
2.3. “Data type” table 9
2.4. “Sensor type” table 11
2.5. “Sensor” table 12
2.6. “Campaign” table 14
3. UPLOADING DATA ON ONEDRIVE 16
3.1. Tables Upload 16
3.2. File Upload 17

LIST OF FIGURES

Figure 1: example of Common Framework for Verona Data Collection from D2.1 5

Figure 2: stationary data collection schema 6

Figure 3: example of “spot_id” and “sensor_id” 7

Figure 4: example of deleted row 16

Figure 5: example of added row 17

LIST OF TABLES

Table 1: “Measuring spot” table schema 8

Table 2: “Data type” table schema 10

Table 3: “Sensor type” table schema 11

Table 4: “Sensor” table schema 13

Table 5: “Campaignr” table schema 15

1. RULES AND SCHEMA FOR DATA COLLECTION AND DATA RECORDING

Some terms used in the next paragraphs and used in the LOOPER database derive from the "Common Framework for Data Collection" of D.2.1.

As said in 2.1 of D2.1, the diagram in Figure 1 explains the problem framing "logical model" shows the hierarchy of the connections between each urban problem to each single data to be collected.

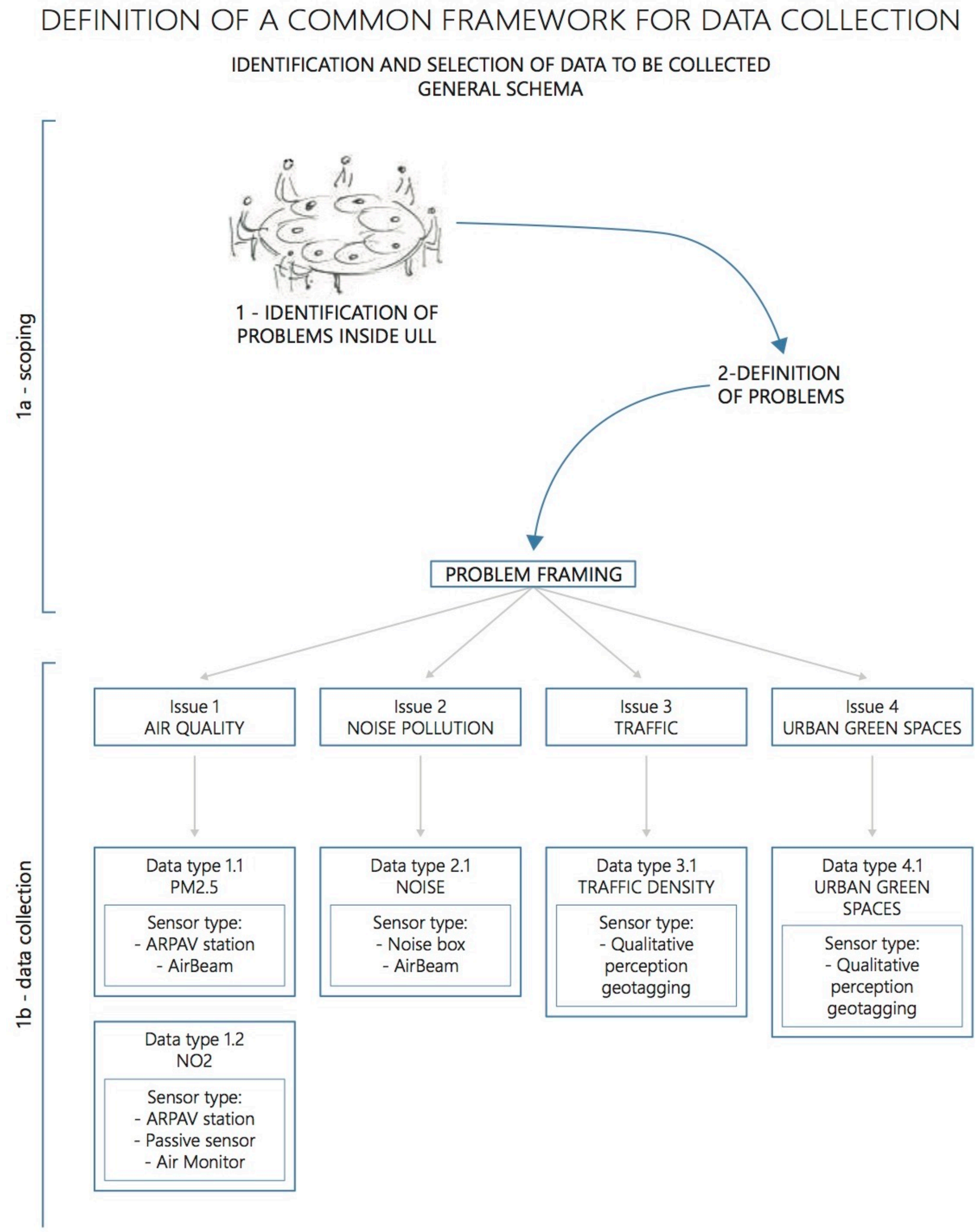

DEFINITION OF A COMMON FRAMEWORK FOR DATA COLLECTION
IDENTIFICATION AND SELECTION OF DATA TO BE COLLECTED
GENERAL SCHEMA
1a - scoping
1 - IDENTIFICATION OF PROBLEMS INSIDE ULL
2-DEFINITION OF PROBLEMS
PROBLEM FRAMING
1b - data collection
Issue 1 AIR QUALITY
Issue 2 NOISE POLLUTION
Issue 3 TRAFFIC
Issue 4 URBAN GREEN SPACES
Data type 1.1 PM2.5
Sensor type: - ARPAV station - AirBeam
Data type 2.1 NOISE
Sensor type: - Noise box - AirBeam
Data type 3.1 TRAFFIC DENSITY
Sensor type: - Qualitative perception geotagging
Data type 4.1 URBAN GREEN SPACES
Sensor type: - Qualitative perception geotagging
Data type 1.2 NO2
Sensor type: - ARPAV station - Passive sensor - Air Monitor

Figure 1: example of Common Framework for Verona Data Collection from D2.1

2. STATIONARY DATA

Here it will be explained more in detail how to name and organize the data that have been collect and needs to be inserted in the database. It is important to understand which is the logic behind the different tables of data that have been implemented in the database.

First, it must be clear since the beginning which kind of data are to be collected. This because the type of data that needs to be collected also influences the sensors that have to be used.

The following sections will show:

- File name: how to name the files of the data collected
- Measuring spot: how to fill the table with the measuring spots that have been chosen for the co-monitoring
- Data type: how to list all data that will be collected
- Sensor type: how to name and organize the type of sensors that will be used
- Sensor: how to list all the sensors that will be used
- Campaign: how to organize data from the co-monitoring campaign

File name here has been inserted as a first section, but please note that it is the result of the data collected and uploaded in the tables of the database.

The “Measuring spot” and “Campaign” table are linked to each other because in the “Campaign” table a data from the “Measuring spot” table is required. Both these tables are not linked to the “Data type”, “Sensor type” and “Sensor” tables.

To the contrary “Data type”, “Sensor type” and “Sensor” are strictly connected one to the other and have to follow the order shown in Figure 2 to be filled:

1. Data type needs to be collected. This will result in the creation of the “Data type” table and the “datatype_id”;
2. “Sensor type” table has to be created and “datatype_id” has to be used to fill in the table about sensor type. This table will create “sensortype_id”.
3. The “sensortype_id” will be used to complete the sensor table which will result in the creation of a “sensor_id”.

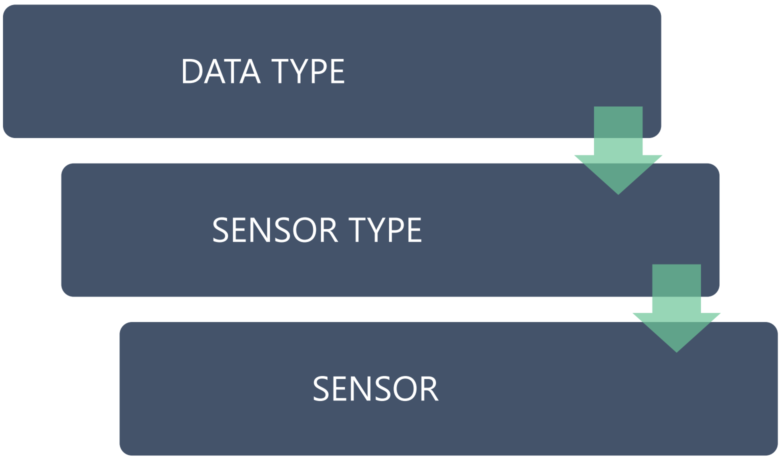


Figure 2: stationary data collection schema

2.1.File name

The file with the data of each different measure must be named with a specific code name. It must be used the following schema:

“spot_id”_”sensortype_id”_”sensor_id”_”start_date”-“end_date”

Example:

VER006_NO2-ARPAV-STATION_1VER_16february2018-23april2018

In the following paragraphs are described all the parameters and the information that have to be collected during monitoring campaign.

2.2.“Measuring spot” table

It must be created a table that list all the “measuring spots” of the ULL. A “measuring spot” (see Figure 3) is a specific place where the measurements are carried out. They are spots inside the city that have been identified as sensible places for a specific issue.

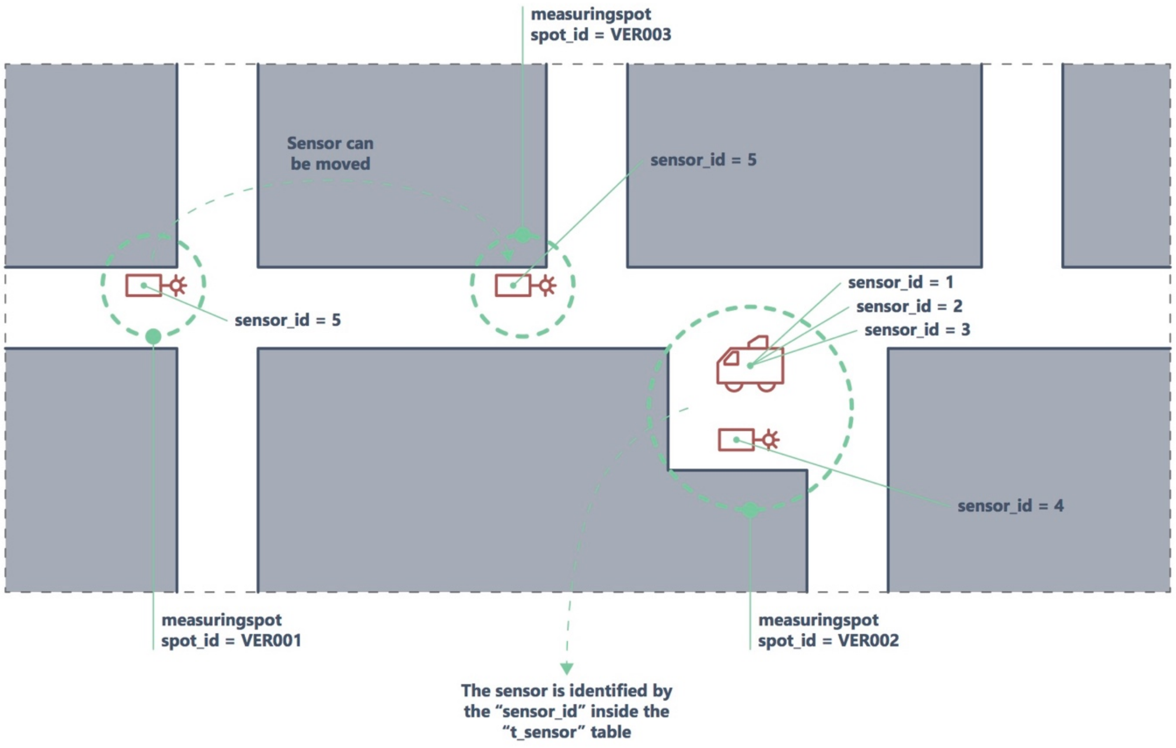


Figure 3: example of “spot_id” and “sensor_id”

The table must have 5 columns filled with the following data:

- “geo_lat” and “geo_lon”:
  They indicates the geographical position of the spot.
  These data are collected with the World Geodetic System 84 (WGS 84) using coordinates in decimal degrees (i.e. Google Maps).

- "name":
  It is the name of the measuring spot.
  You should use here a name that easily indicates the position inside the city. You can use a free text inside here.
  N.B. This text will be visualized in the Platform.
- "spot_id":
  It is a unique code for the identification inside the database of the measuring spot.
  It must be unique and has to follow the rule:
  - the first three letters indicate the City (VER for Verona, MAN for Manchester and BRU for Brussels);
  - then a three-digit number.
- "note":
  It is a free field if you want to add notes.

The table schema is the following:

geo_lat	geo_lon	name	spot_id	note
45.40843	10.98517	Chiesa San Giovanni Battista	VER001	
45.41976	10.96926	Chiesa Santa Maria Assunta	VER002	
45.42076	10.97793	Fiera via Scopoli	VER003	
…..	…...	…….	…….	

Table 1: "Measuring spot" table schema

2.3. "Data type" table

The Data Type table lists all kind of "Data type" that are used in the data collection process of each Looper Living Lab.

Referring to the problem framing schemas of D.2.1 (paragraph 2.2), for each issues (a specific problem that will be analyzed in the Living Lab of a city) we are going to collect some data: a "Data type" is a measure of an indicator that is related with the specific issue. For example some "Data type" can be: concentration of PM2.5, concentration of NO2, use of public spaces, urban noise, ...

The table must have 5 columns filled with the following data:

- "datatype_id":
  It is a unique code for the identification inside the database of the type of data.
  It must be a single word with no spaces, there can be dots (i.e. PM2.5) but no other symbols.
- "datatype":
  It is an extended version of the datatype_id which should help participants to understand the type of data. I.e. the datatype field of a pollutant will be its name instead of its formula.
  N.B. The text in the "datatype" field will be visualized in the platform.
- "measunit":
  This field must be filled with the measuring unit of the data type.
- "issue":
  The issue column refers to the issues listed in Figure 1: example of Common Framework for Verona Data Collection from D2.1
- "note":
  It is a free field if you want to add notes.

This is the schema of the "Data Type" table:

datatype_id	datatype	measunit	issue	note
NO2	Nitrogen dioxide (NO2)	ug/m3	Air pollution	
NH3	Ammonia (NH3)	ug/m3	Air pollution	
CO	Carbon Monoxide (CO)	mg/m3	Air pollution	
NOISE	Noise	dB(A)	Noise pollution	
PM2.5	Particulate Matter 2.5	ug/m3	Air pollution	
GREENSPACES	Urban green spaces	qualitative	Urban green spaces	
MOBILITY	Streets and mobility	qualitative	Traffic	
ENVIRONMENT	Quality of the urban environment	qualitative	Urban green spaces and Traffic	

datatype_id	datatype	measunit	issue	note
ELEMENTS	Urban elements	qualitative	Urban green spaces and Traffic	

Table 2: "Data type" table schema

2.4.“Sensor type” table

The Sensor Type table lists all kind of “Sensors” that are used in the Data collection process of each Looper Living Lab.

Referring to the problem framing schemas of D.2.1 (paragraph 2.2), for each Data we are using some instruments to measure it. In some cases we use also different instruments to measure the same “datatype_id”. Each different kind of instrument represent a “Sensor Type”. For example the “datatype_id” PM2.5 can be measured with two different “Sensor Types”: a portable sensor or with a stationary official measuring station, etc.

The table must have 3 columns filled with the following data:

- “sensortype_id”:
  It is a unique code for the identification inside the database of the type of data.
  This code can have more words but with no spaces, the words must be divided by an hyphen.
  The first word of the “sensortype_id” must be the “datatype_id”
- “type”:
  It indicates the name of the kind of sensor that is used.
  N.B. The text in the “type” field will be visualized in the platform.
- “datatype_id”:
  It is the same code used in the 2.3 “Data type” table.

This is the schema of the “Sensor Type” table:

sensortype_id	type	datatype_id
CO-ARPAV-STATION	CO rilevazioni ARPAV	CO
NO2-ARPAV-STATION	NO2 rilevazioni ARPAV	NO2
NH3-ARPAV-STATION	NH3 rilevazioni ARPAV	NH3
PM-ARPAV-STATION	PM2.5 rilevazioni ARPAV	PM2.5
NO2-PASSIVE-SENSOR	NO2 radiello	NO2
NO2-AIR-MONITOR	NO2 air monitor	NO2
PM-AIRBEAM	PM2.5 airbeam	PM2.5
NOISE-BOX	Noise box	NOISE
GREENSPACES-CROWDMAPPING	Crowdmapping	GREENSPACES
MOBILITY-CROWDMAPPING	Crowdmapping	MOBILITY
ENVIRONMENT-CROWDMAPPING	Crowdmapping	ENVIRONMENT
ELEMENTS-CROWDMAPPING	Crowdmapping	ELEMENTS

Table 3: “Sensor type” table schema

2.5.“Sensor” table

The sensor table is a list of all sensors (instruments, devices etc..) that you use to make measurements. You can have more sensors for each “Sensor Type”.

The table must have 6 columns filled with the following data:

- “sensor_id”:
  It is a unique code for the identification inside the database of the sensor.
  It must be unique and has to follow the rule:
  - first a one-digit number:
  - then the first three letters of the City (VER for Verona, MAN for Manchester and BRU for Brussels);

  NB: Table can be integrated at any time, but the “sensor_id” name MUST remain the same
- “name”:
  It shows the name that have been given to the sensor. A number can be added if there are more sensors of the same kind.
- “stationary”:
  It indicates if the sensor is stationary or not.
  NB: the “t” stands for “true”, while the “f” stands for “false”
- “owner_id”:
  It indicates who is the owner of the sensor (i.e. national agency, living lab, etc.)
- “sensortype_id”:
  It is the same code used in the 2.4 “Sensor type” table
- “note”:
  It is a free field if you want to add notes.

This is the schema to follow:

sensor_id	name	stationary	owner_id	sensortype_id	note
1VER	ARPAV Mobile Station 01 NO2	t	ARPAV	NO2-ARPAV-STATION	
2VER	ARPAV Mobile Station 01 NH3	t	ARPAV	NH3-ARPAV-STATION	
3VER	ARPAV Mobile Station 01 CO	t	ARPAV	CO-ARPAV-STATION	
4VER	ARPAV Mobile Station 01 PM	t	ARPAV	PM-ARPAV-STATION	
5VER	ARPAV Mobile Station 02 NO2	t	ARPAV	NO2-ARPAV-STATION	
6VER	ARPAV Mobile Station 02 NH3	t	ARPAV	NH3-ARPAV-STATION	
7VER	ARPAV Mobile Station 02 CO	t	ARPAV	CO-ARPAV-STATION	
8VER	ARPAV Mobile Station 02 PM	t	ARPAV	PM-ARPAV-STATION	

sensor_id	name	stationary	owner_id	sensortype_id	note
9VER	NO2 Air Monitor 01	t	ULLVR	NO2-AIR-MONITOR	
10VER	PM AirBeam 01	f	ULLVR	PM-AIRBEAM	
11VER	PM AirBeam 02	f	ULLVR	PM-AIRBEAM	
12VER	Noise box 01	t	ULLVR	NOISE-BOX	
13VER	Noise box 02	t	ULLVR	NOISE-BOX	
14VER	Noise box 03	t	ULLVR	NOISE-BOX	
15VER	Noise box 04	t	ULLVR	NOISE-BOX	
16VER	Noise box 05	t	ULLVR	NOISE-BOX	
17VER	Noise box 06	t	ULLVR	NOISE-BOX	
18VER	Crowdmapping web app	f	ULLVR	GREENSPACES-CROWDMAPPING	
19VER	Crowdmapping web app	f	ULLVR	MOBILITY-CROWDMAPPING	
20VER	Crowdmapping web app	f	ULLVR	ENVIRONMENT-CROWDMAPPING	
21VER	Crowdmapping web app	f	ULLVR	ELEMENTS-CROWDMAPPING	
.....	.....	.....	.....	.....	
100VER	NO2 radiello 01	t	ULLVR	NO2-PASSIVE-SENSOR	
.....	.....	.....	.....	.....	

Table 4: "Sensor" table schema

2.6.“Campaign” table

The Campaign table lists all the monitoring “campaigns” that have been done in each Looper Living Labs. A “campaign” is each single measurement activity.

The table must have 7 columns filled with the following data:

- “campaign_id”:
  It is a unique code that will be automatically generated by the database
- “start_date” and “end_date”:
  They indicates the starting date and the ending date of the co-monitoring campaign.
  These data are collected using the YYYY-MM-DD schema.
- “spot_id”:
  It is the same code used in the 2.2 “Measuring spot” table
- “sensor_id”:
  It is the same code used in the 2.5 “Sensor” table
- “camp_type”:
  Only with the following labels can be used:
    - Official data and measurements
    - Participatory sensing
    - Stakeholders reports
- “file_name”:
  It is the same code used in the 2.1

This is the schema to follow:

campaign_id	start_date	end_date	spot_id	sensor_id	camp_type	file_name
	2018-02-16	2018-04-23	VER006	1VER	Official data and measurements	VER006_NO2-ARPAV-STATION_1VER_16february2018-23april2018
	2018-02-16	2018-04-23	VER006	2VER	Official data and measurements	VER006_NH3-ARPAV-STATION_2VER_16february2018-23april2018
	2018-02-16	2018-04-23	VER006	3VER	Official data and measurements	VER006_CO-ARPAV-STATION_3VER_16february2018-23april2018
	2018-02-26	2019-03-09	VER013	12VER	Participatory sensing	VER013_NOISE-BOX_13VER_26february2018-09march2018
	2018-01-31	2018-02-21	VER003	12VER	Participatory sensing	VER003_NOISE-BOX_13VER_31january2018-21february2018

campaign_id	start_date	end_date	spot_id	sensor_id	camp_type	file_name
	2018-02-26	2018-03-09	VER014	13VER	Participatory sensing	VER014_NOISE-BOX_12VER_26february2018-09march2018
	…..	…..	…..	…..	…..	

Table 5: "Campaignr" table schema

3. UPLOADING DATA ON ONEDRIVE

Once the data are collected you will be required to upload it on the OneDrive LOOPER Team Shared Folder following the steps described in the next sections.

There will be only one table per type (measuring spot, data type, sensor type, sensor, campaign), this means you MUST upload data on the tables that will be found in:

LOOPER Team Shared > WP2_Data collection and visualization > Monitoring data collection > Tables

To facilitate the first upload, already known data from the Verona LLL have been uploaded and can be used as blueprints.

3.1.Tables Upload

To upload data in the tables please follow the instruction from 2. STATIONARY DATA of this document.

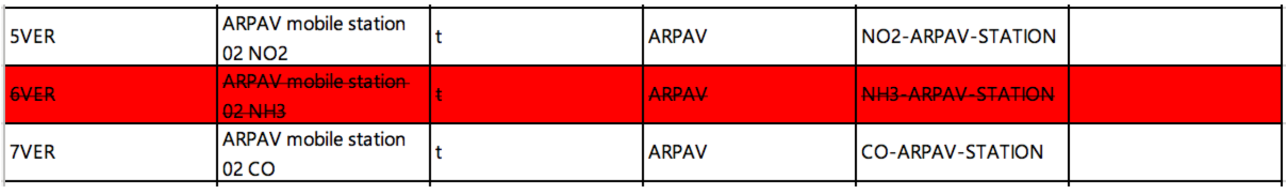
In the following sub sections it will be explained how to make modifications to the tables.

3.1.1.Tables changes

Please notice that you MUST NOT change any data, you can only add rows from the tables.

This is extremely important and has to be kept in mind, otherwise there can be problems with the database.

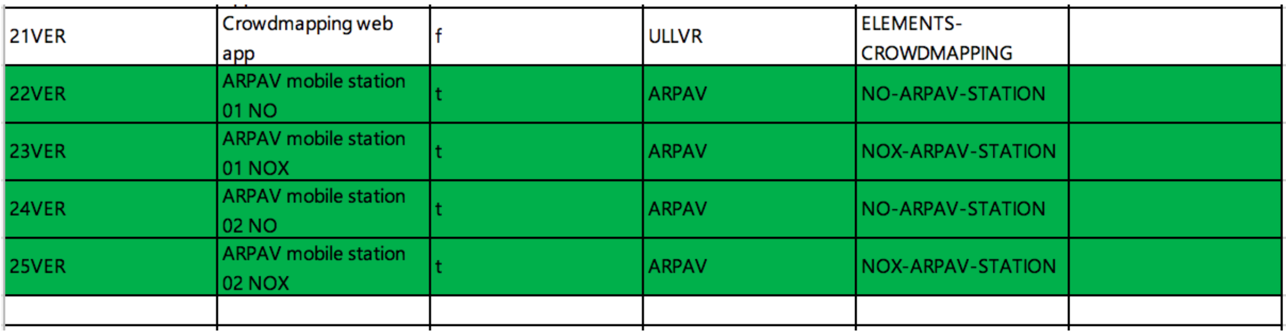
If you need to delete a row from a table on the folder, you have to fill the row in red (Figure 4). To delete a row you DO NOT HAVE to delete the information in the row, you ONLY HAVE to turn it red. In addition you also have to select the text and apply the strikethrough formatting to it (shortcut for Windows: CTRL+5; shortcut for Mac: CMD+SHIFT+X).



5VER	ARPAV mobile station 02 NO2	t	ARPAV	NO2-ARPAV-STATION	
~~6VER~~	~~ARPAV mobile station 02 NH3~~	~~t~~	~~ARPAV~~	~~NH3-ARPAV-STATION~~	
7VER	ARPAV mobile station 02 CO	t	ARPAV	CO-ARPAV-STATION	

Figure 4: example of deleted row

In the same way, if you want add a row you to add it after the last filled row, you have to fill it with the data required and in the end you also have to fill the row in green to declare that it is a new set of data (Figure 5).



21VER	Crowdmapping web app	f	ULLVR	ELEMENTS-CROWDMAPPING	
22VER	ARPAV mobile station 01 NO	t	ARPAV	NO-ARPAV-STATION	
23VER	ARPAV mobile station 01 NOX	t	ARPAV	NOX-ARPAV-STATION	
24VER	ARPAV mobile station 02 NO	t	ARPAV	NO-ARPAV-STATION	
25VER	ARPAV mobile station 02 NOX	t	ARPAV	NOX-ARPAV-STATION	

Figure 5: example of added row

Nota Bene: a set of linked data used in a deleted row cannot be used again (i.e. the sensor_id "6VER" from Figure 4 once has been deleted cannot be used anymore).

3.2.File Upload

After filling all the tables you will be able to compose the file name for your data. After you are able to gain the file name the first thing to do is to rename the linked document (.txt, .xls or .xlsx) so that it gains an univocal reference and it is ready to be uploaded.

After naming the file, in the Monitoring data collection folder (LOOPER Team Shared > WP2_Data collection and visualization > Monitoring data collection), you will find other three sub-folders, besides the "Table" folder, where the file can be uploaded.

As can be seen from the folders name, data will be collected and stored separately for each city. The three folders are:

- WP5_data BRU;
- WP6_data VER;
- WP7_data MAN;

and in each of them all data collected in a LLL must be uploaded. No other subfolders have to be created as file names already give information about the type of pollutant, the campaign period and the sensor that have been used.

Once data have been uploaded to the folder they MUST NOT BE CHANGED as they will be uploaded to the database.

Nota Bene: The prefix (001, 002, etc.) that can be seen for files in the folder WP6_data VER is an automatically generated prefix which will be added once files are uploaded to the database. You MUST NOT add it by yourselves.