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

Version: 2.1

Date: 07 Oct 2020

Responsible partner: IUAV

Authors: Giovanni Borga (IUAV), Massimiliano Condotta (IUAV), Chiara Scanagatta (IUAV), Paolo Ruggeri (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	10Nov2017	Draft	GB	First internal draft
0.2	29Nov2017	Draft	GB, MC, CS, PR	Internal draft
0.3	19Dec2017	Draft	GB, MC, CS	First draft for content sharing and contribution request
0.4	20Dec2017	Draft	MC, CS, GB	First draft for content sharing and contribution request
0.5	24Jan2018	Draft	IK, MW, JA	Contributions
0.6	24Jan2018	Draft	MC, CS, GB	Receipt of contributions and integration
0.7	28Jan2018	Draft	JA	Contributions
1.0	31Jan2018	Final	MC, CS, GB	Receipt of contributions and integration
2.0	22Jul2019	Final	MC,CS	Final version with integration by VUB and UoM
2.1	07Oct2020	Final	CS	Final version with integrations

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1.Objectives of WP2

The objective of WP2 is to investigate how participatory sensing can complement deliberation to identify problems in the urban realm. It addresses specific challenges concerning technical feasibility, legal framework and visualization options of feedbacks for citizens.

The WP therefore defines the methodologies and tools for urban data collection, data processing and information visualization related to the specific topic of each Urban Living Lab.

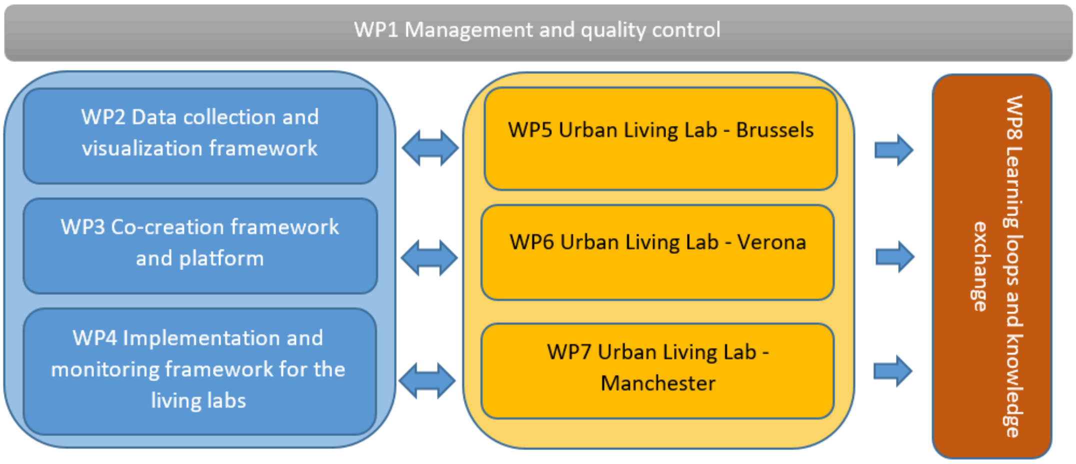


Figure 1: overall structure of the LOOPER work plan and WPs structure

WP2 contributes to the definition of the model of Stage 1 (identification of problems) and Stage 3 (data monitoring) of the LOOPER platform. A first preliminary version will be created and subsequently updated based on feedback after the first loop from the living labs.

Results of WP2 feed into WPs 3-4 (development of the LOOPER methodology and platform) and the living labs in WPs 5, 6 & 7.

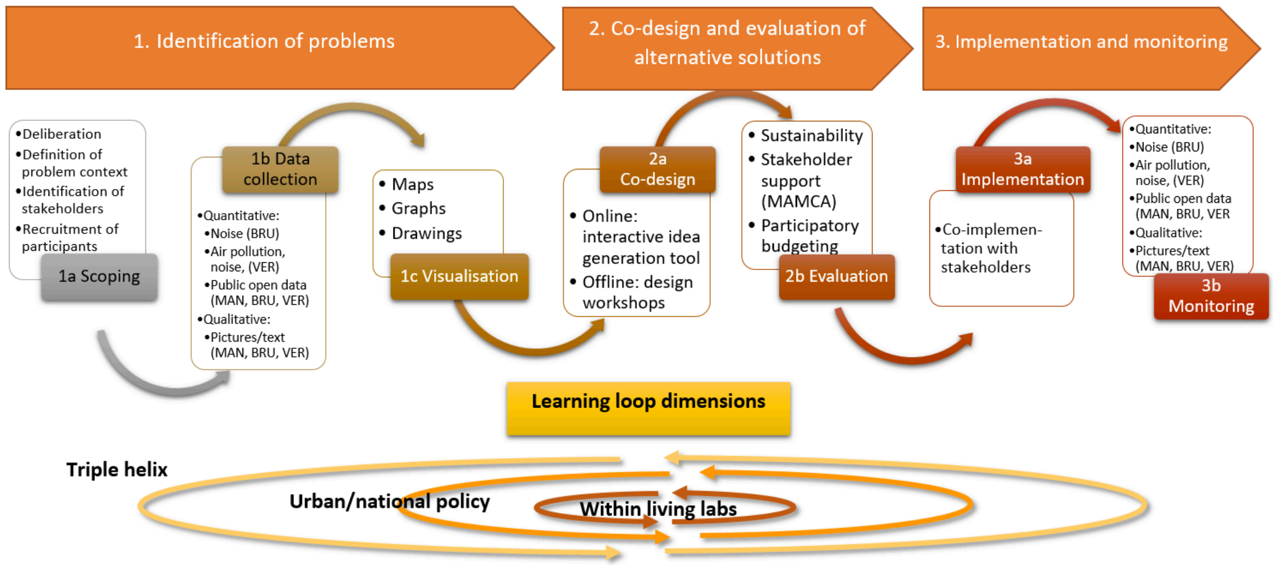


Figure 2: the concept of the LOOPER platform

1.2.Specific objectives of D.2.1

The Deliverable 2.1 is a report on the activities of the first task of WP2: "T.2.1 Set-up of procedure and framework for distributed data collection". In this task, the following aspects of the LOOPER platform are studied and then defined:

- Issues to be investigated in the three different Urban Living Labs
- List of innovative portable devices and static official measurement systems to be used in the three different Living Labs.
- Procedure and a common framework for the collection of data;
- Conceptual schema of the interaction between the different functionalities
- Preliminary functions and aspects of the data visualization dashboard.

All the activities are coordinated by IUAV.

1.3.Target of D.2.1

The results of D.2.1. feed several other tasks of the LOOPER project and are addressed towards different stakeholders of the project. Deliverable results will be used mainly in the following activities:

- T2.2: Set-up of the LOOPER database;
- T2.3: Creation of participatory methods and information dashboard;
- T2.4: Development of the legal framework;
- T3.1: Development of the methodology for the co-design of alternatives;
- T4.1: Development of guidelines for implementing the urban living labs;
- T4.3: Set-up of the LOOPER Platform;
- T5.1: Inception of living lab and scoping of problems;
- T5.2: Participatory data collection and visualization;
- T6.1: Inception of living lab and scoping of problems;
- T6.2: Participatory data collection and visualization;
- T7.1: Inception of living lab and scoping of problems;
- T7.2: Participatory data collection and visualization.

Stakeholders directly addressed by this Deliverable are:

- Project partners;
- Project co-operation partners;
- Implementation partners;
- ULL leaders and managers.

1.4.Structure of this document

The objective of this document is to give all LOOPER project partners the tools to properly collect the data they need, and to create and implement the visualization of all data in order to have a clear and complete vision of the situation.

To make the document as clear as possible it has been divided into six chapters, each of which deals with a specific topic. Please consider that each topic is specific but also linked to all of the others.

The first chapter of this document concerns the **PROBLEM FRAMING**.

This section has the intent to describe the logical model that relates the ULLs to the problems that they have to face, and to give a generic description of the type of data that each ULL has to collect (both quantitative and qualitative).

The logical model shows how the hierarchy of urban problems and data to be collected is connected. This means that it shows the steps in between the identification of problems inside the ULL, the definition of those problems, the problem framing, the list of problems that have been found and its needed data.

The list of data to be collected will sometimes coincide in the three ULLs, and in other cases will be different. Because of this it is described for each ULL which data are needed in order to give to the participants the tools to make the best analysis of possible mitigation solutions. Every ULL have its own logical model with a definition of problems and data to be collected that have to be clarified before the beginning of the monitoring phase.

The second chapter analyzes the creation of a **COMMON FRAMEWORK OF DATA TYPOLOGY**.

The intent of this section is to explain which data typology classification has been chosen and that will be used to divide data both in a logical and in a graphical way. This helps to better understand which kind of data are we collecting and why are we collecting it.

The goal of the common framework is in fact to develop a valid conceptual data model for the LOOPER database and platform. This method of data grouping will be used to upload the data (that will need in some cases a "manual" upload and in others an "automatic" upload) into the platform.

The categories of data (official data from stationary survey; participatory data collected with portable measurement tools; participatory qualitative data) that have been identified will be better explained mentioning the type of data that will enter that category, the type of visualization that will be needed and which kind of data will be shown.

The third chapter shows the **LOOPER PLATFORM DATA FLOW**.

Here are taken into account the already mentioned topics of problem framing and common framework which need to merge to create the schema of how the LOOPER platform, and LOOPER website, is going to work.

The platform will be based on the idea of having a section for each step of the learning loop approach, and different pages for each ULL.

In the data flow, it is possible to see:

- How the website and platform are linked;
- How the platform and the external geotagging tool are linked;
- Which sections will be found in the platform;
- Which data will be found in each section;
- Which tools will be found in some sections;
- How the data collected will be uploaded in the dashboard section.

The LOOPER platform data flow will be implemented during the work of the ULLs in order to add the data and the tools that will be needed for the co-creation and co-evaluation process.

The fourth chapter is meant to give a list of possible **TOOLS FOR DATA COLLECTION**.

This list collects a sample of tools that can be used to collect both quantitative and qualitative data that the ULLs will need for their work.

These tools can be from an official agency or from participants; with high costs or low cost; already made or self-made.

The list of tools is divided in three categories (two categories show tools for quantitative data collection and one is about qualitative data collection):

- Tools for stationary surveys;
- Tools for mobile measurements;
- Tools for qualitative data collection (feedback and report).

For each tool, there is a description of the type of data that it collects, how it works, how to build it (if it has to be assembled), which operating system is needed (if referring to apps) and some other technical data (like measurement precision, image of components or interface, etc.).

The fifth chapter shows the results of the ULLs **DATA COLLECTION PRELIMINARY SURVEY**.

This section shows the dataset characterization sheet and the results of the data collection preliminary survey.

This data is needed to setup the common framework. Here you will find the preliminary survey filled by each ULL in which it was asked which issues are investigated and which data will be needed for each issue.

The second paragraph of this section explores more specific information about each problem using a predetermined set of questions that have to be answered by each ULL:

- What: object of the survey;
- How: collection mode, tools and devices;
- When: time information and accuracy;
- Where: location information and accuracy;
- Who: user/operator information.

The sixth chapter explains the **COMMON FRAMEWORK FOR DATA MODELLING AND PROCESSING**.

Talking about data modelling and processing, the common framework is a set of rules and protocols needed to manage the data collected by the ULLs in a unified way.

To better explain this common framework, this topic has been divided into three paragraphs:

- Conceptual data model;
- Data model details tables;
- Protocols and procedures.

The conceptual data model is the part of the framework that defines how the data will be stored and related. This data model is formed by four main entities, several minor entities and some lookup tables.

Each of these entities and tables is explained in more detail in the data model details tables paragraph, that follows the conceptual data model description, and will give information such as: keys; field name; field data type; field description and notes.

In the end, the protocols and procedures paragraph describes how these are, and have to be, defined to ensure the effective data flow and process from data collection storages to LOOPER platform.

2. PROBLEM FRAMING

This section describes the structure of data collection in relation to problems that the Urban Living Labs will work with. It is a preparatory guide finalized at the development of a common framework for data collection (T2.1).

2.1.Problem framing logical model

The diagram of the problem framing "logical model" shows the hierarchy of the connections between each urban problem to each single data to be collected. See section 2.2.3 for the problem framing logical model of the Verona case study.

2.1.1. Logical model diagram

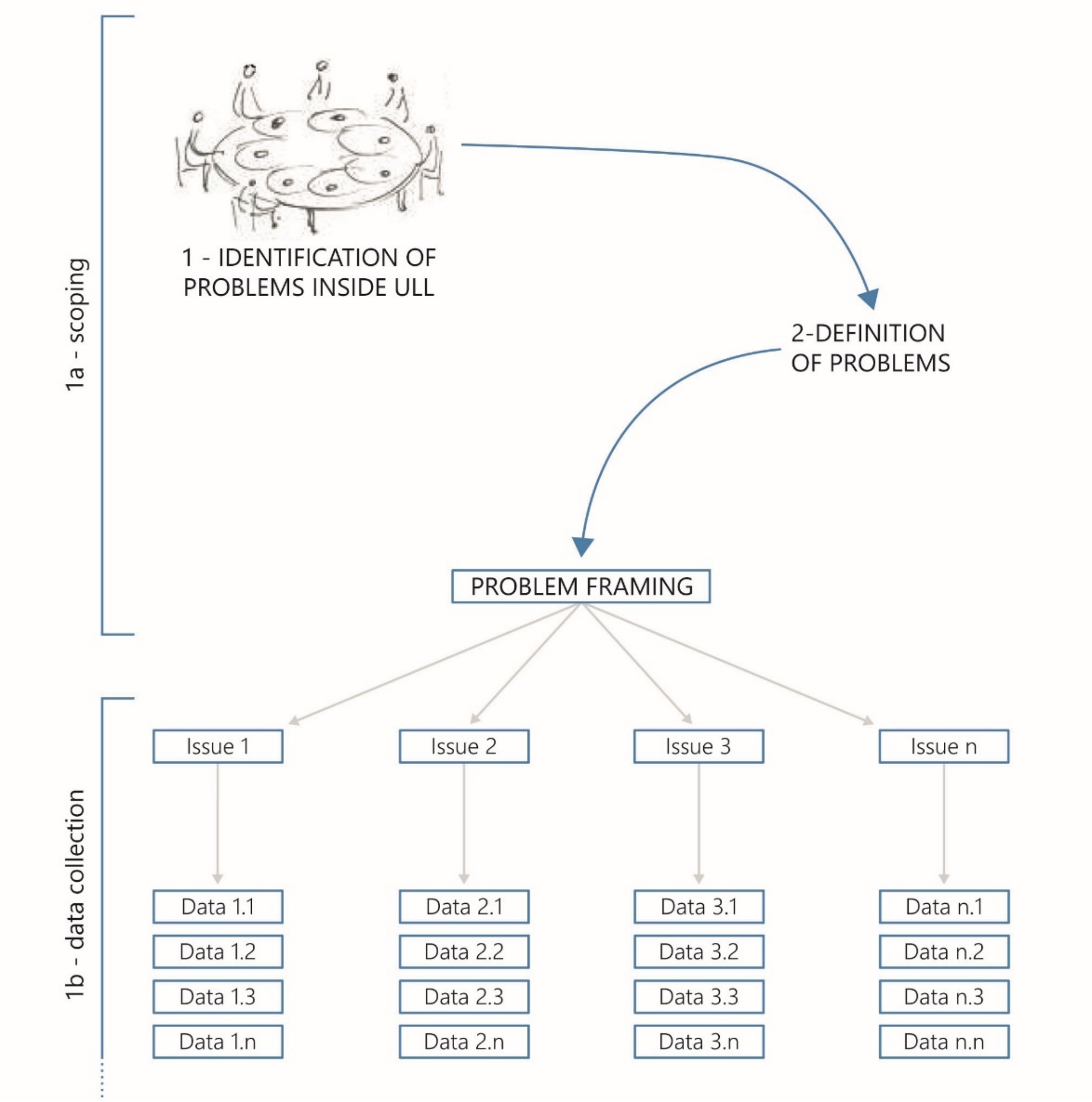


Figure 3: problem framing logical model diagram

2.2.Data to be collected from the Urban Living Labs

2.2.1. Brussels

Brussels is currently at the phase of the “identification of problems inside the living lab” therefore cannot confirm the data needs which are dependent on the final selected location of the ULL and the problems that the citizens propose.

In any case, it will most probably be a combination of quantitative data (noise and/or air quality and/or traffic) and qualitative data (pictures, sound, video).

Brussels expects the questions “Which data would we like to collect?” to only be treated by late march within the living lab. Until then, the below is only an informed estimate, rather than the actual choice of data to be collected:

Relevant government data comes from: https://data-mobility.brussels/fr/, formats available are HTML, WFS, GeoJSOM, CSV

Only one of the following topics will emerge as main theme of Living Lab:

- Safety on the way to school:
  - mapping of typical routes to school and highlighting of any spots perceived as dangerous
  - speed limits of streets (feed-in of government data)
  - Location of traffic lights, official crossings (feed-in of government data: http://opendatastore.brussels/fr/dataset/traffic-lights)
  - reported accident rates (feed-in of government data)
- Managing different uses along high street shared by bus, tram, cars, biker, parking, pedestrians
  - Current allocation of space along high street amongst the various means of transport and /or space dedicated to parking
- Rethinking the design of public space around a church and park located next to the high street
  - Analysis of pedestrian activity to reorganized calmed sections of streets around church and adjacent park
  - Parking occupancy rate (realtime in DatexII: http://opendatastore.brussels/fr/dataset/parking-occupancy)

Below therefore a hypothetical set of possible issues covered in Brussels:

- Travel behaviour and usage of public space
  - Usage of public space: Observations by ULL participants, recorded on map
  - Routes to school/cycling routes: Collected through interviews with passers-by, asked to draw their current trajectory onto a map.
- Traffic
  - Bicycle paths (gov data)
  - Location of Traffic lights (gov data)
  - Parking occupancy (realtime; static averages at 11am; static averages at 10pm)
  - Degree of congestion (gov data)
  - Accident rates (gov data)
  - Vehicle speed (fixed radar)
- Urban Space
  - Allocation of space to different means of transport (might be available a government data but could also require drawing shapes)
  - Space dedicated to parking (might be available a government data but could also require drawing shapes)

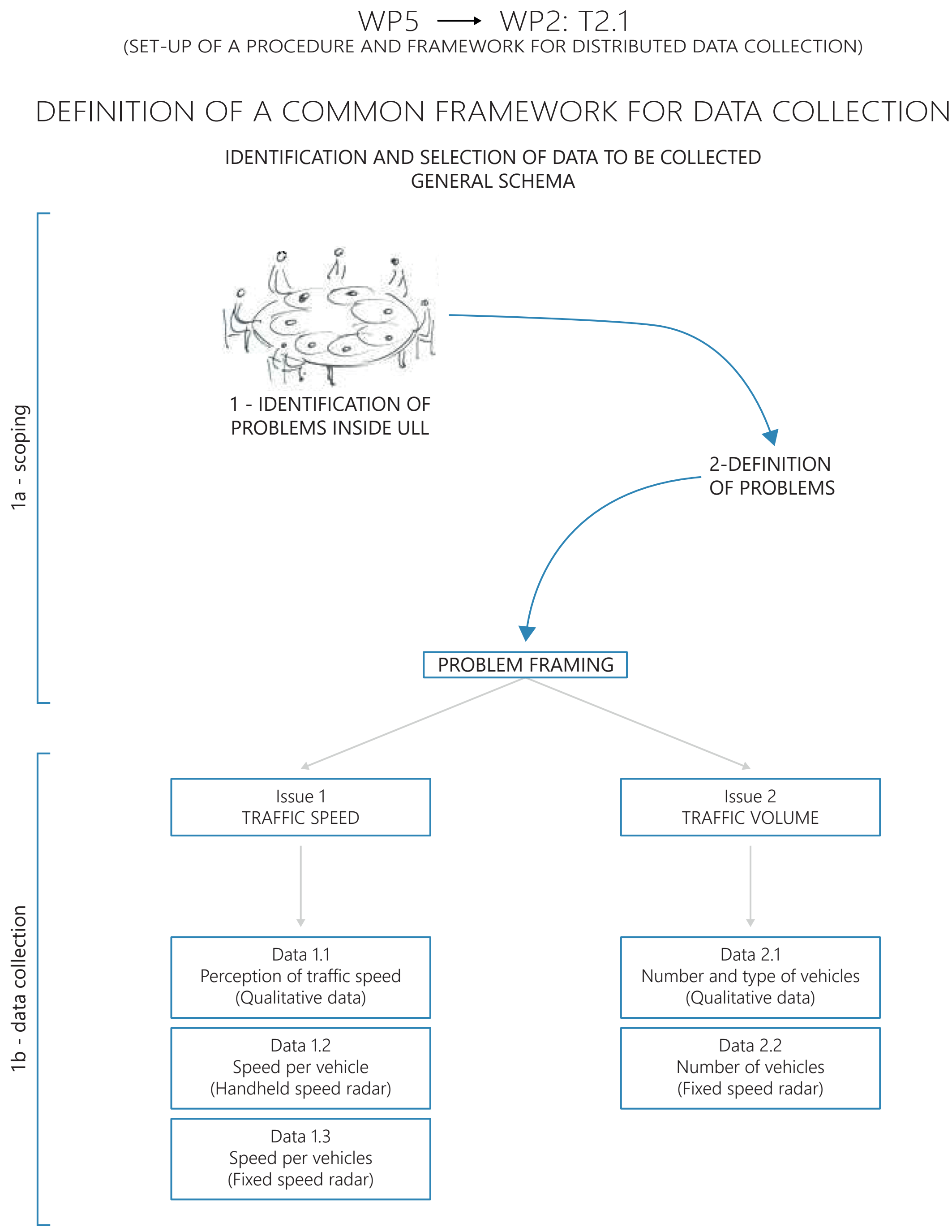


Figure 4: problem framing logical model of Brussels

2.2.2. Manchester

Four overlapping problems of concern to Brunswick residents have been identified:

Air Quality—Traffic—Safety—Community Spaces

Air quality is primarily linked to traffic, and particularly to the volume of traffic moving around and through the neighbourhood. Brunswick is bordered by an elevated expressway and two major roads. Residents frequently express concern about through traffic related to 'rat runs' (use by commuters to avoid traffic on other roads) and about parking by university and hospital employees. Volumes of traffic in the neighbourhood obviously also contribute to issues of **traffic safety** and therefore we hope to work closely with Transport for Greater Manchester (TfGM) which is interested in both issues and has considerable data. We will be able to take advantage of data from two government (DEFRA) air quality monitoring stations on either side of the neighbourhood and will collect our own data at specific locations—we expect to have some access to high quality handheld monitors to be used by university colleagues and to purchase low cost monitors that residents can use (possibly Plume *Flow* https://plumelabs.com). We will calibrate the less accurate handheld devices against the fixed monitors. We will also seek to obtain health information related to air quality from government and academic sources. There are particular concerns about children's health and safety and this relates to particular areas such as around the school and children's routes to school and so we will emphasise data collection in these areas.

Traffic safety (and air quality) is also linked to behaviours and practices with respect to how people drive and also about how pedestrians and cyclists move about the area and where children play or how people otherwise use community spaces. We will identify spaces and routes of concern and monitor behaviours in these areas. The movement and the activities in **public spaces** also relate to other sorts of **safety issues**. Interestingly, some residents perceive Brunswick, or at least the part of it in which they live, as very safe. (It might be interesting to see if the characteristics of specific areas could be identified and transferred to areas perceived as less safe.) The effect of changes in layout (particularly footpaths) has also been raised as a safety issue, particularly as layouts change constantly during construction. We plan to undertake walks with certain groups during which we will map routes and areas of concern.

Community spaces include **green spaces**, which may have an effect on **air quality** and other health and wellbeing concerns. Issues related to community spaces include the presence of necessary **amenities** (e.g. laundrette, shops) and quality spaces in the neighbourhood and also access to those a bit further afield. Brunswick is surrounded by amenities in the city centre and on the university campus but residents don't necessarily see these as part or their patch or take full advantage of them. We will facilitate participatory personal mapping exercises for residents to indicate their usual destinations and what they find to be satisfactory, unsatisfactory, inaccessible or lacking.

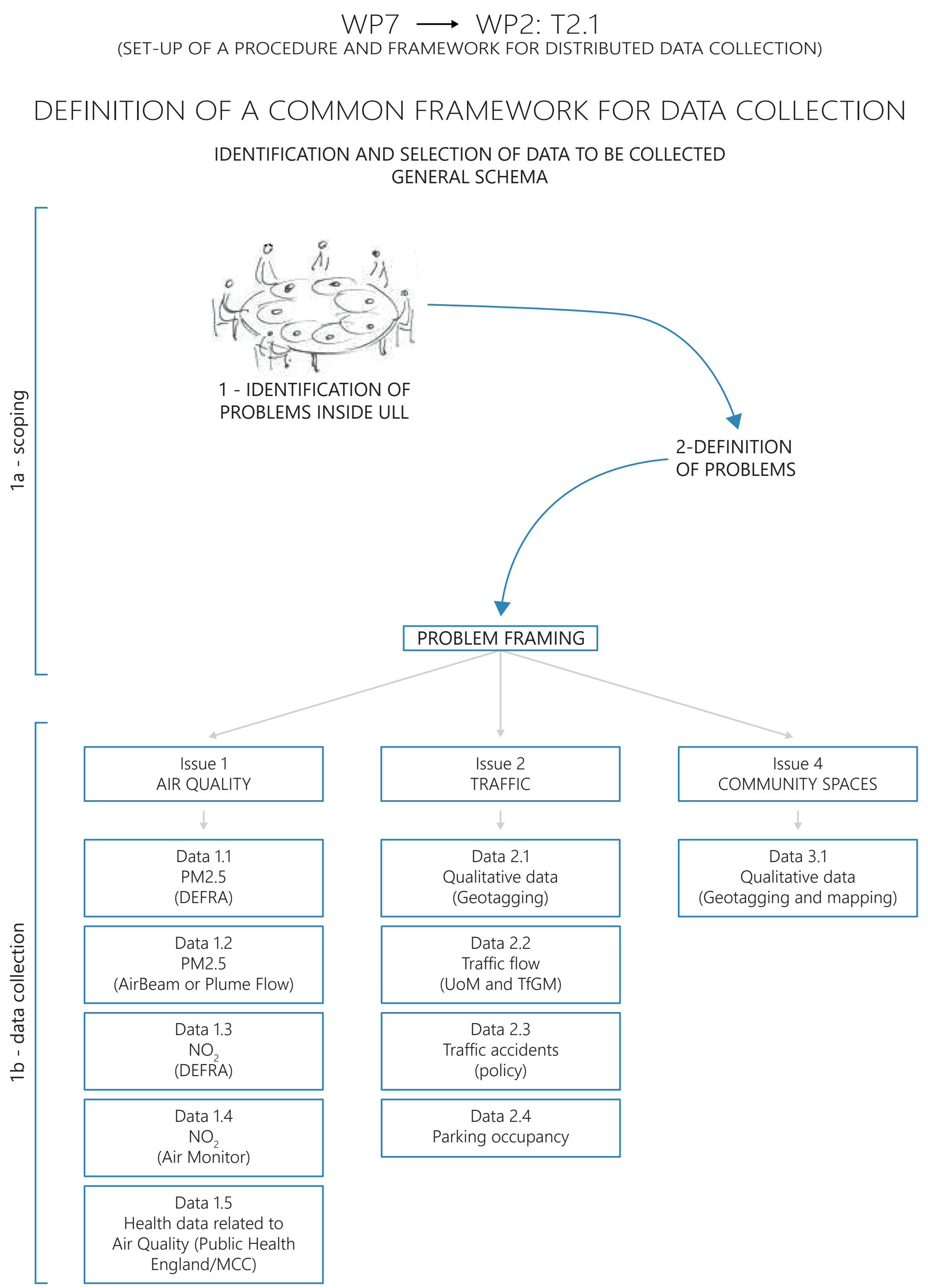


Figure 5: problem framing logical model of Manchester

2.2.3. Verona

The Urban Living Lab of Verona will focus on urban challenges related to environmental pollution caused by traffic and absence of natural mitigation measures (such as green areas).

To evaluate these issues the parameters that will be analyzed and measured will be air and noise pollution.

These parameters will be monitored using different methodologies and different tools.

Regarding methodologies, data will be acquired both by the national agency ARPA (Regional Agency for the environmental protection of Veneto) and by participatory data collection.

Tools used to collect this data are therefore different: both official instruments, such as, the mobile station of ARPAV, and portable devices operated by ULL members will be used.

The complete list of measured parameters and tools are listed in par. 5.2.3. while the following diagram shows the problem framing of Verona issues.

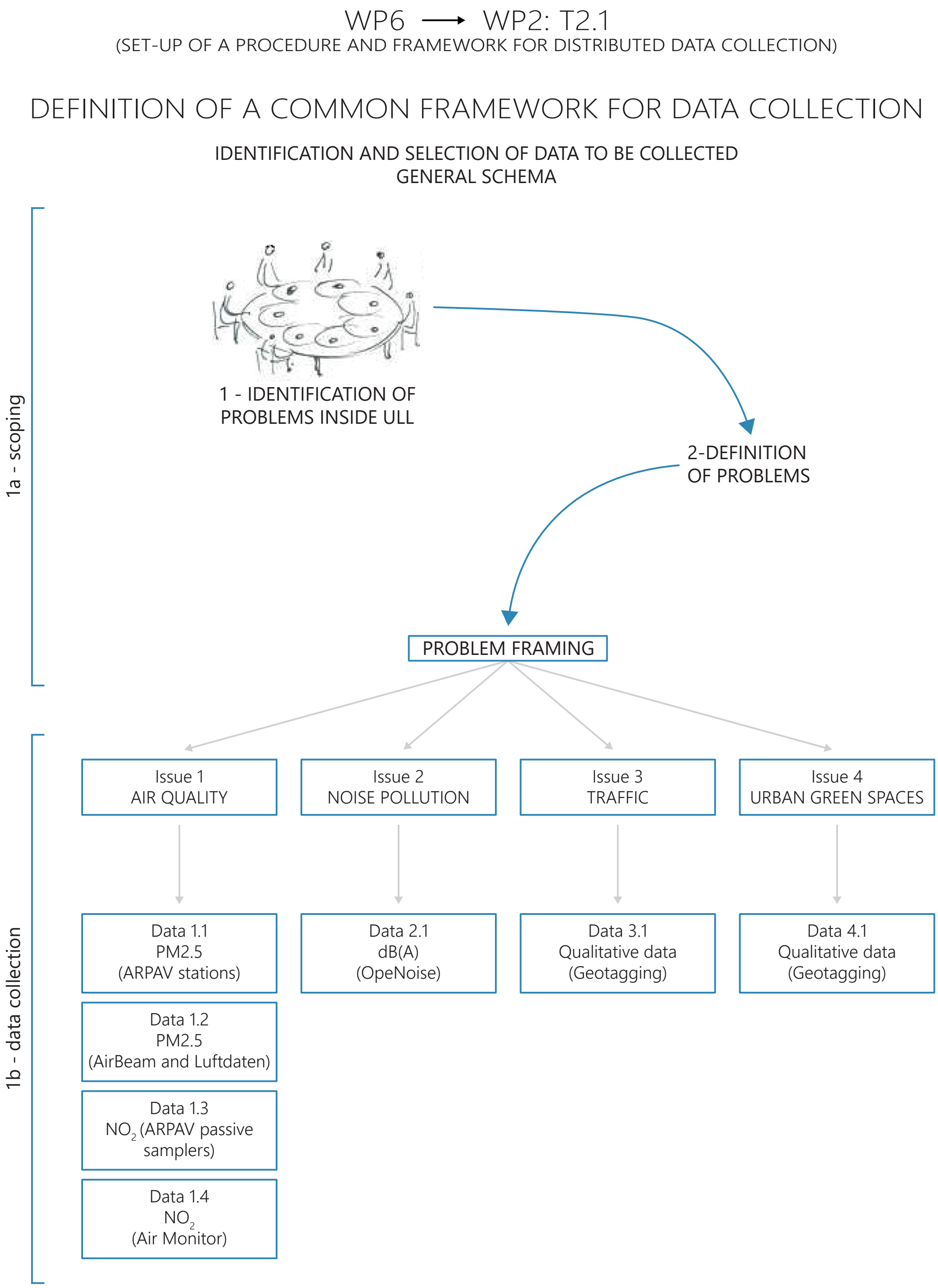


Figure 6: problem framing logical model of Verona

COMMON FRAMEWORK OF DATA TYPOLOGY

To setup a common framework for data collection and processing, we have to group the various types of samples that will be taken into specific categories in order to develop a valid conceptual data model for the LOOPER database and platform.

2.3.Data typology classification

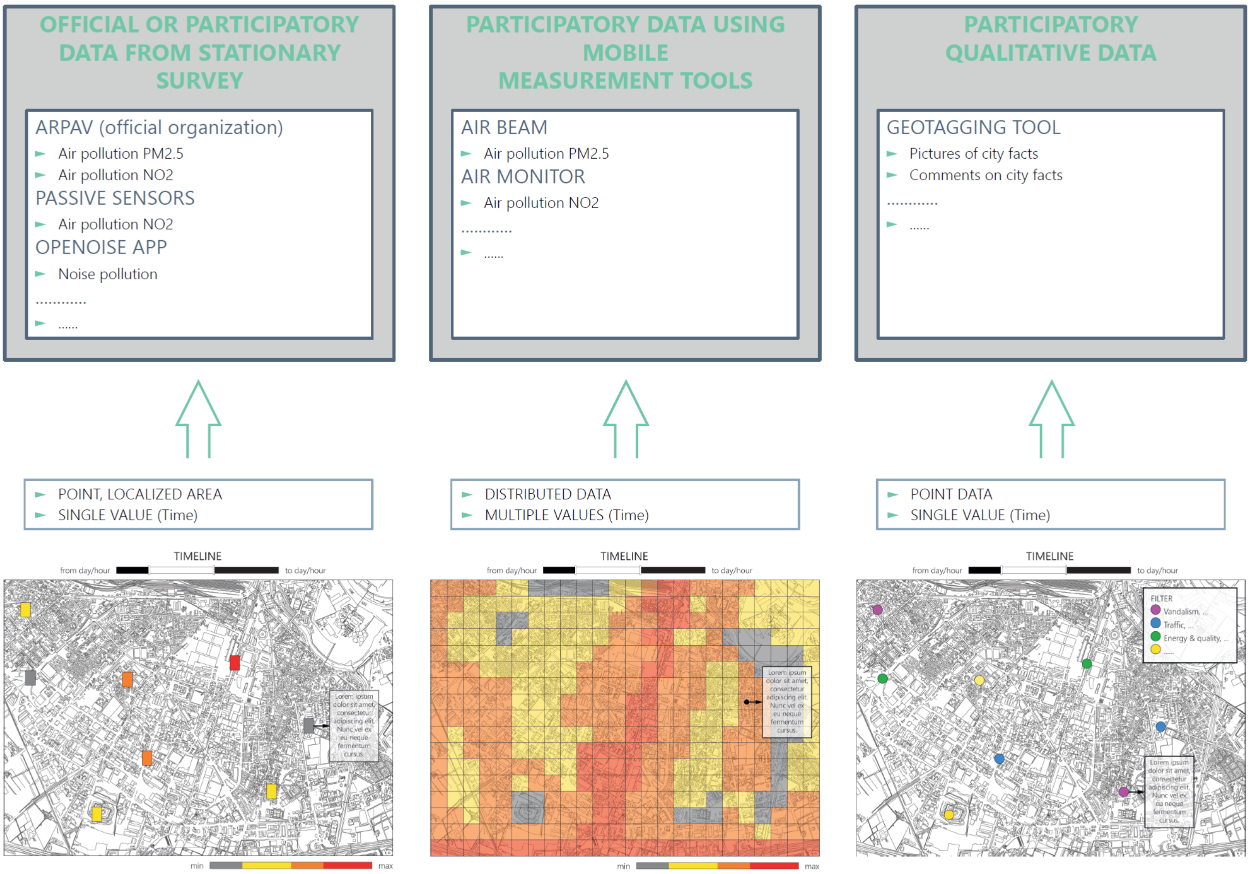


Figure 7: data typology classification scheme

2.4.Official data from stationary surveys

The first category of data typology classification regards surveys to be carried out using stationary sensors/stations made by both official agencies and participatory processes.

Such kind of surveys produce multi-temporal / multi-sensor series of data samples related to a single location. Data values will be visualized by placing a push-pin on the map which will be identifiable to have a pop-up info box showing some kind of graphs and summary information (see example in Figure 8).

For this kind of data, it is pointless to produce a layer with continuous coverage. This because there is no scientifically valid technique to interpolate values between the spots of the surveyed points.

In the case of multi-sensor survey station their info box will show graph/summary data about all sensors related to the identified station.

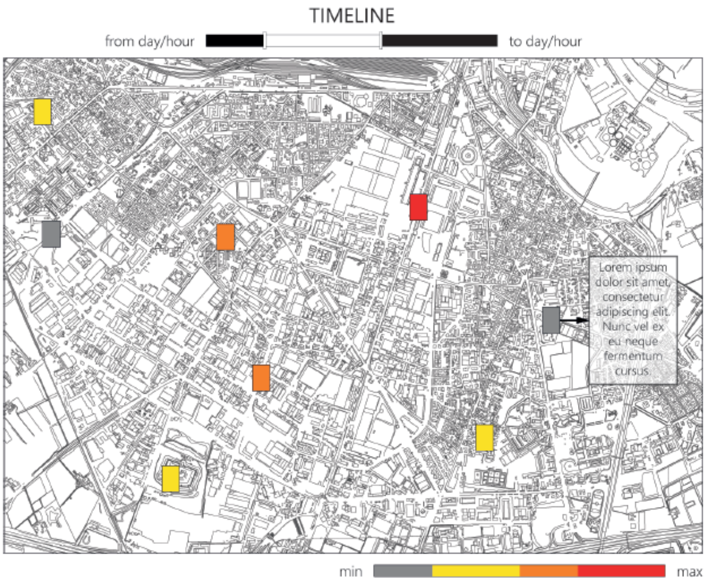


Figure 8: stationary survey data mapping example

As shown in Figure 8, push-pins might be thematic-colored to give a synoptic view of some quantitative parameter and put in evidence specific peak values.

2.5.Participatory data collected with portable measurements tools

Participatory data collected in a mobile manner, on the contrary of the previous case, must be mapped through a kind of "coverage layer" (see example in Figure 9) because survey points have no pre-defined location and samples distribution might be irregular.

In this case we could adopt two different approaches: the first one is based on a simple calculation on average values of concurrent measures in the same spatial unit (e.g. a grid cell), the second one is based on an interpolation function (e.g. heat-maps). The first approach is preferable in case of high irregularity of spatial distribution of measurements and temporal inhomogeneity of surveys with many cases of spatial samples overlapping, while the second is better when we have a limited number of samples and some homogeneous spatial sample distribution.

Since mobile surveys in this project will produce large samples dataset and irregular spatial distribution of surveyed areas, the simple average values based approach seems to be preferable.

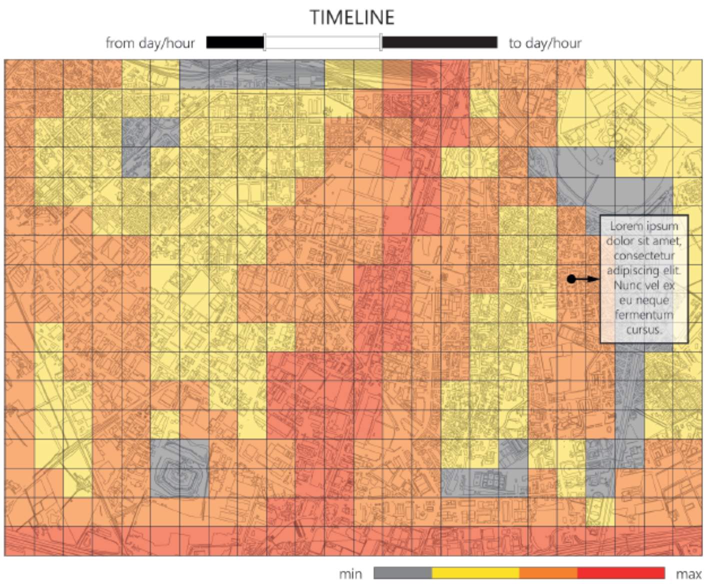


Figure 9: mobile survey grid-based data mapping example

One of the initial assumptions will be about the choice of the spatial unit and resolution to be used to build the visualization layer.

In most cases, users will be able to click on a cell of the grid covered with a layer and will get a pop-up info box showing summary data such as average values, number of samples, time ranges/graphs etc. We can also decide to design info-boxes showing data from other sensors and/or different sensor value in comparisons or with relations.

2.6.Participatory qualitative data

The third category of data typology classification regards surveys to be carried out using smartphone geotagging applications. As in the case of stationary surveys, qualitative data collection is made of information related to a fixed geographical location.

When placing a tag, user is requested to fill out a form with some place-related information. In general, the form may contain a basic free text input box and some other more structured input controls; in most cases there may be a choice selection from a predefined list and the possibility to attach multimedia files such as pictures or videos. Also, some ranking value (e.g. five-star selection tool) might be provided. Basically, tags might not contain any quantitative value to make any calculation or thematic mapping rather than a simple and user-friendly ranking tool since submitting data is mostly a kind of observed situation, not a measured value.

Figure 10: qualitative data mapping example

For this type of data, we can assume to adopt the push-pin representation technique again (see example in Figure 10). Color code may be useful to visualize both a ranking value (if provided) or to differentiate pins according to the category list. The identify info box will allow users to view/download attached multimedia files.

3. LOOPER ICT SYSTEM AND LOOPER PLATFORM DATA FLOW

The LOOPER project ICT system is going to be a sort of "suite" of interconnected web tools that will provide several communication and collaborative functions gathered inside a common frame.

This data-flow schema is intended to show the system architecture of the LOOPER ICT system and the connections between each part of it. The diagram also shows where data are to be generated, processed, transferred and visualized within the ULL collaborative processes.

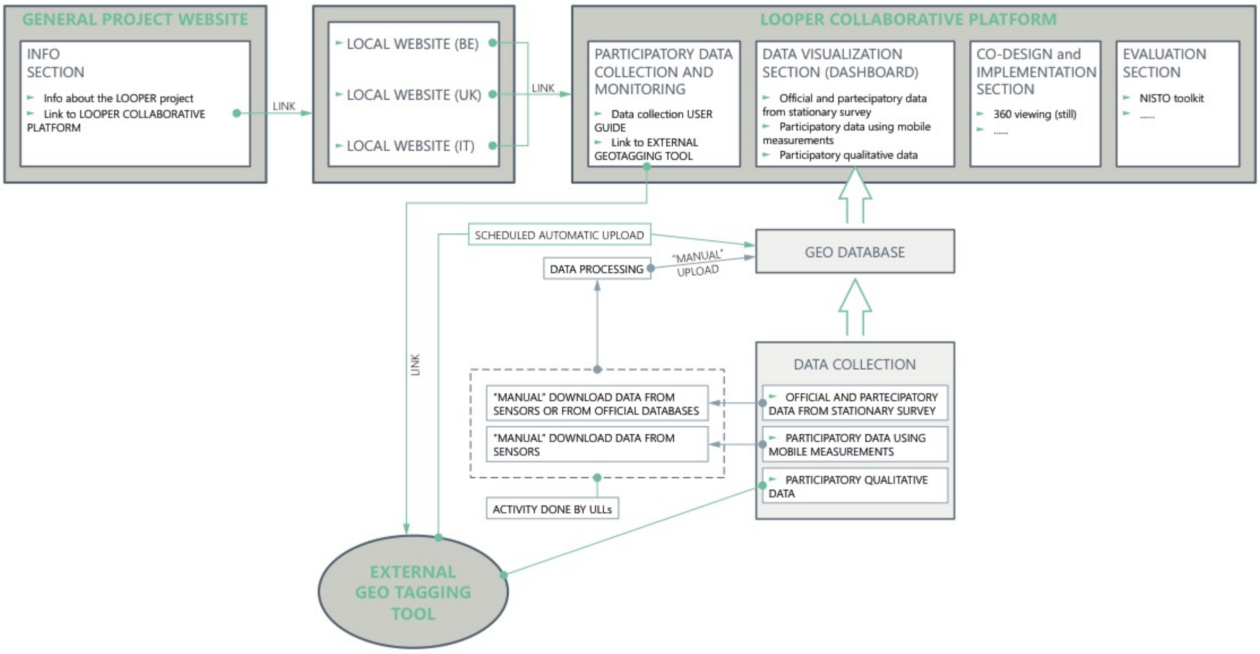


Figure 11: LOOPER ICT system and LOOPER platform data flow overall schema

In the upper part of the diagram there are two boxes depicting the main sections of the LOOPER system: the general project website and the LOOPER collaborative platform (the infrastructure of the platform will be divided into two parts, and one will be the frontend section with the three local websites). While the general project website is the section where all information about the development of LOOPER project are published, the LOOPER collaborative platform is a collection of tools specially developed for the Learning Loop processes in the living labs.

The structure of the websites is outlined in D8.2 Dissemination to the public.

The general project website and the LOOPER collaborative platform are going to be separated; simple hyperlinks will allow visitors to jump from one section to the other.

The LOOPER collaborative platform is made up of the following sub-sections:

- General introduction to the ULL (including location, problem description, news, events, local contact details)
- The participatory data collection and monitoring section;
- The data visualization section;
- The co-design and implementation section;
- The evaluation section.

While the co-design and co-evaluation sections definition stage is still to be approached, the data collection and visualization sections are about to be detailed and developed.

For each ULL a local website (collaborative web platform) will be developed which will follow the above sections in the local language (Brussels: FR/NL; Verona: IT; Manchester: ENG).

With regard to the participatory data collection section, there will be a direct link to the external application for the geotagging suggested in section 4.3. As shown in the diagram, data collected with this tool will be stored in the main LOOPER platform geo-database via a scheduled automatic upload procedure.

The data visualization section is a section specially designed to display to explore and better know the issues about their cities. In this section, there will be several interactive maps and reports with which users will be able to perform custom searches, to access data details and statistics, and to navigate through thematic maps that will show the results of all data collection campaign.

The data visualization section is powered by a geo-database that will store and process data coming from stationary surveys, mobile measurements as well as data from the mentioned geotagging qualitative data. In particular, data from participatory surveys and official agencies data gathering will be collected and pre-processed by ULLs participants in order to upload them into the main geo-database with a blended system of manual/semi-automatic procedures depending on the various used sensors and tools.

4. TOOLS FOR DATA COLLECTION

This chapter reports basic useful information and suggestion about tools that can be used by all Urban Living Labs to carry out stationary, mobile and qualitative surveys during the project.

4.1.Tools for stationary surveys

4.1.1. ARPAV fixed and mobile station

ARPAV checks the air quality in the regional territory through a network of fixed control units integrated by surveys carried out by mobile vehicles that are periodically moved to carry out air quality monitoring campaigns in areas of the territory not completely covered by fixed stations.

Fixed control units work as background stations, and are used to have a complete view of the situation of air pollution during the year. Mobile stations are often used to cover more specific areas, depending on the needs of the Municipality and of citizens.

It is important to understand that all the instruments, samples, component, and the analysis made on the samples collected with these two types of stations, are entirely the same.

Campaigns carried out by mobile stations last an average of 8 to 10 weeks. To ensure representativeness in terms of weather and climate conditions, the campaigns are divided into two periods: one in spring-summer, the other in autumn-winter.

At the end of each monitoring campaign, ARPAV technicians prepare a report describing the data collected and the conclusions on the air quality situation in the monitored area. As soon as they are available, the reports are published on the ARPAV portal in the dedicated section of the Provincial Department of competence.

Figure 12 ARPAV fixed station

Figure 13 ARPAV mobile station

4.1.2. Passive sensors

The diffusive sampler is a closed box, usually cylindrical. Of its two opposite sides, one is “transparent” to gaseous molecules which cross it, and are adsorbed onto the second side. The former side is named diffusive surface, the latter is the adsorbing surface (marked with S and A in the figure). Q is the sampling rate and has the dimensions of a gaseous flow. Therefore, if Q is constant and measured, to calculate the ambient air concentration you need only to quantify the mass of analyte trapped by the adsorbing material and to keep note of the time of exposure of the diffusive sampler.

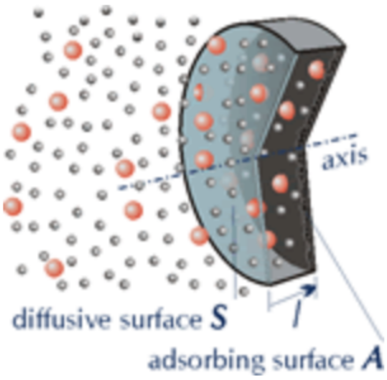


Figure 14 In the diffusive sampler, the adsorbing and the diffusive surfaces are two opposing planes of a closed box. Driven by the concentration gradient, the gaseous molecules (coloured in the figure) pass through the diffusive surface and are trapped from the adsorbing surface.

How is it possible to improve Q then? The answer is to improve the sampler geometry to a radial design. From this idea, the Radiello[1] sampler has been developed, its cylindrical outer surface acting as diffusive membrane: the gaseous molecules move axially parallel towards an adsorbent bed which is cylindrical too and coaxial to the diffusive surface.

When compared to the axial sampler, Radiello shows a much higher diffusive surface without increase of the adsorbing material amount. Even if the adsorbing surface is quite smaller than the diffusive one, each point of the diffusive layer faces the diffusion barrier at the same distance.

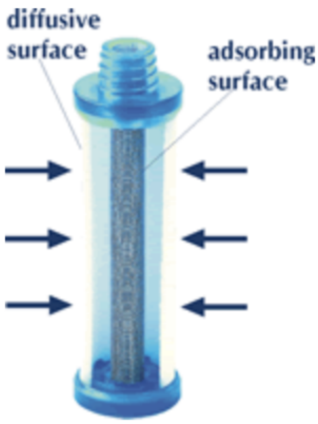


Figure 15 Section of Radiello. Diffusive and adsorbing surfaces are cylindrical and coaxial: a large diffusive surface face, at a fixed distance, the small surface of a little concentric cartridge.

[1] Read how to use it from the following URL: http://www.radiello.com/english/usoen.htm

4.1.3. GPS PM2.5 and PM10 logger

Luftdaten is a tool used to get instantaneous data of the exposure to PM2.5 and PM10 (Figure 16).

Figure 16 Luftdaten components

This GPS PM2.5 and PM10 logger measure PM levels by assembling a SDS011 fine dust sensor with a NodeMCU firmware installed on a ESP8266 module that can be connected with a Wi-Fi router, to which a DHT22 temperature and humidity sensor can be added. The whole sensor is based on an Arduino system that is used to upload the firmware to allow the functioning of the module. Once the firmware is uploaded and the other components are assembled it is possible to install the sensor to update on the crowdsourcing platform the data collected continuously. The sensor here uses the light scattering method to analyse the amount of PM particles to be found in the air sucked in the sensing chamber.

4.1.4. Android Sound Level Meter

The sound level meter apparatus is provided by an android smartphone, a lavalier microphone, a power bank and a waterproof enclosure.

Figure 17 Android Sound Level Meter components - Smartphone

Figure 18 Android Sound Level Meter components - Microphone

Figure 19 Android Sound Level Meter components – Power bank

Figure 20 Android Sound Level Meter components - Enclosure

The price of the apparatus depends on the type of hardware chosen but it can be less than 150 euro.

Example:

Android Smartphone Wiko Lenny:	82.49	€
Microphone Boya BY-LM10	19.95	€
Puridea 20000mAh Power Bank	19.54	€
Waterproof enclosure (190x140 mm):	7.99	€
Total	129.97	€

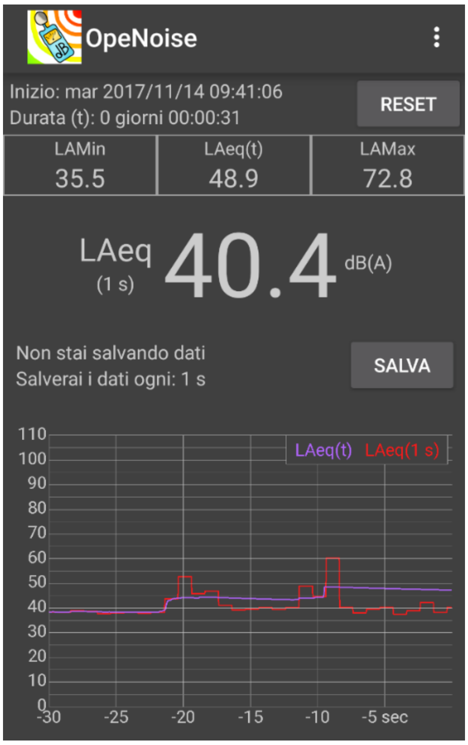
The sound level meter software used for the Sound Level Meter apparatus is OpeNoise by Arpa Piemonte[2], it is an open source software downloadable for free from the play Store[3].

Unfortunately, it is only in Italian, but it is still easy to use.

In the following figure, the main screen is depicted. The App allows the user to store the A weighted equivalent level in a .txt file.

In the .txt file is it also possible to store the third octave band log.

In the settings, it is possible to calibrate the microphone and to change the time step of the logging from one second to one hour.



Main screen:

Date, hour of the measurement and total length

Maximum, minimum and equivalent levels of the whole measurement.

Real-time LAeq (every second)

Logging (on – off). The measure is stored in a .txt file. The time step can be chosen from 1s to 60 minutes

The graphic shows the logging data of LAeq integrated in the whole measure (t) and in one second.

The Linear and A weighted spectrums can be also shown and stored in the log file.

[2] http://www.arpa.piemonte.gov.it/approfondimenti/temi-ambientali/rumore/rumore/openoise-2
[3] https://play.google.com/store/apps/details?id=it.piemonte.arpa.openoise

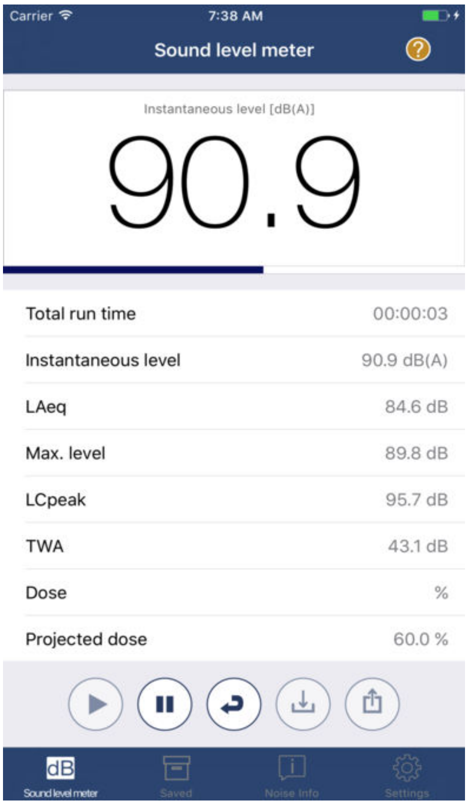
4.1.5. iOS Sound Meter APP

Because the OpeNoise app works only on Android smartphones, one option for iOS users is the *Niosh SLM APP.*

The *NIOSH Sound Level Meter* (SLM) app automatically provides a readout of the instantaneous sound level in A, C, or Z weighted decibels using the smartphone's built-in microphone (or preferably using an external, calibrated, microphone).

The APP was tested and verified (± 2 dB) with a reference type 1 Sound Level Meter. The APP allows to save and share data measurements using the smartphone media features. If location services are enabled, it is possible to provide geospatial references for the noise measurement. It is also possible to calibrate the APP accessing the Setting screen, and an external calibrated microphone can be added.

In conclusion, like the *OpeNoise* APP the *Niosh SLM* APP is for free, it is possible to add an external microphone that can be calibrated and it can register the noise level choosing the period of time manually.



Main screen:

Real time dB(A)

Various data among which total run time, real time LAeq, maximum level (dB)

Start and stop of logging, clear and start again, store, export

Figure 21 Niosh SLM APP homepage interface

Other APPs that can be used are:

- deciBel (Android)
- Sound Meter PRO (Android)
- dB Volume Meter (iOS)
- Decibel X (iOS)

4.1.6. Black CAT Radar

The Black CAT Radar unit (Figure 22) allows to collect traffic data without the need of an in-road traffic sensors. This sensor allows to detect the lane position of vehicles, and with one sensor it is possible to collect data of two lanes at time. The algorithm that controls the sensors is also able to detect the length of the vehicles and differentiates the vehicles by groups - i.e. bikes and motorbikes, cars, vans, short lorries and lorries.

Figure 22 Black CAT Radar on a light pole

The device can been powered by using batteries that need to be changed on average every six days. The device is also assembled to be powered by using solar panels or by plugging it directly to an electric plug.

This device allows both real-time data download, meaning that the device automatically sends the data to the in-station, and what they call 'historical mode' so that the sensor waits for the user to manually - and physically - download the data from the device itself.

4.1.7. Telraam

This traffic counting sensor is based on a Raspberry Pi microcomputer, sensors and a low-resolution camera and it collects data as it is able to recognize vehicles or pedestrian when analysing the captured images (Figure 23). It can send data about traffic counting to its central database if continuously connected to a Wi-Fi source - the data upload takes place every hour. As other stationary devices it needs a power outlet within reach.

Figure 23 Telraam sensor installed on a window inside a private house

It is important to correctly position the sensor to allow a reliable data collection, it needs to be mounted on the inside of an upper-floor window with a view on the street. The idea behind this type of sensors is to have as many of them as possible around the city.

4.2.Tools for mobile measurements

4.2.1. GPS NO_2 CO logger

The instrument is based on an *Arduino UNO* board. It provides a GPS shield *Adafruit* that include also an SD card for data logging. It is also possible to connect NO_2 and CO sensor in the free digital pins of the shield and store the data in the SD card within GPS coordinates date and time.

Figure 24 Components of the GPS NO_2 CO logger - Arduino UNO

Figure 25 Components of the GPS NO_2 CO logger - GPS shield

Figure 26 Components of the GPS NO_2 CO logger - NO_2 CO sensor

Figure 27 Components of the GPS NO_2 CO logger - Power bank

Figure 28 Components of the GPS NO_2 CO logger - Enclosure

Figure 29 Components of the GPS NO_2 CO logger - Micro SD card

Arduino is a flexible board that allows the user to connect many different sensors so the apparatus can be object of further developments.

The software is made in the Arduino IDLE and it is flexible and continuously improved.

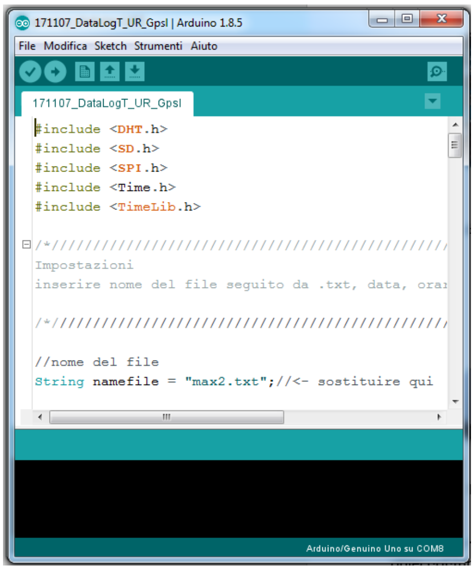


Figure 30 Arduino board interface

4.2.2. GPS PM2.5 logger

AirBeam is a tool used to get instantaneous data of the exposure of the person to PM2.5.

The AirBeam measure PM2.5 using a light scattering method; air is sucked in a sensing chamber wherein a LED bulb scatters off particles in the airstream, is registered by a detector and converted into a measurement that estimates the number of the particles in the air.

The device works via Bluetooth and sends data to the AirCasting Android APP every second, the APP then maps and graphs the data in real time on the smartphone.

At the end of each measuring session all of the collected data are sent to the AirCasting server and there gets crowdsourced with data collected from all the other devices to create a map showing where PM2.5 concentrations are highest and lowest.

Because it is an open-source platform it is possible to modify the components to take other measurements and or to transmit the data to other websites or APPs. Because of the presence of an expansion port on the device it is possible to add sensors.

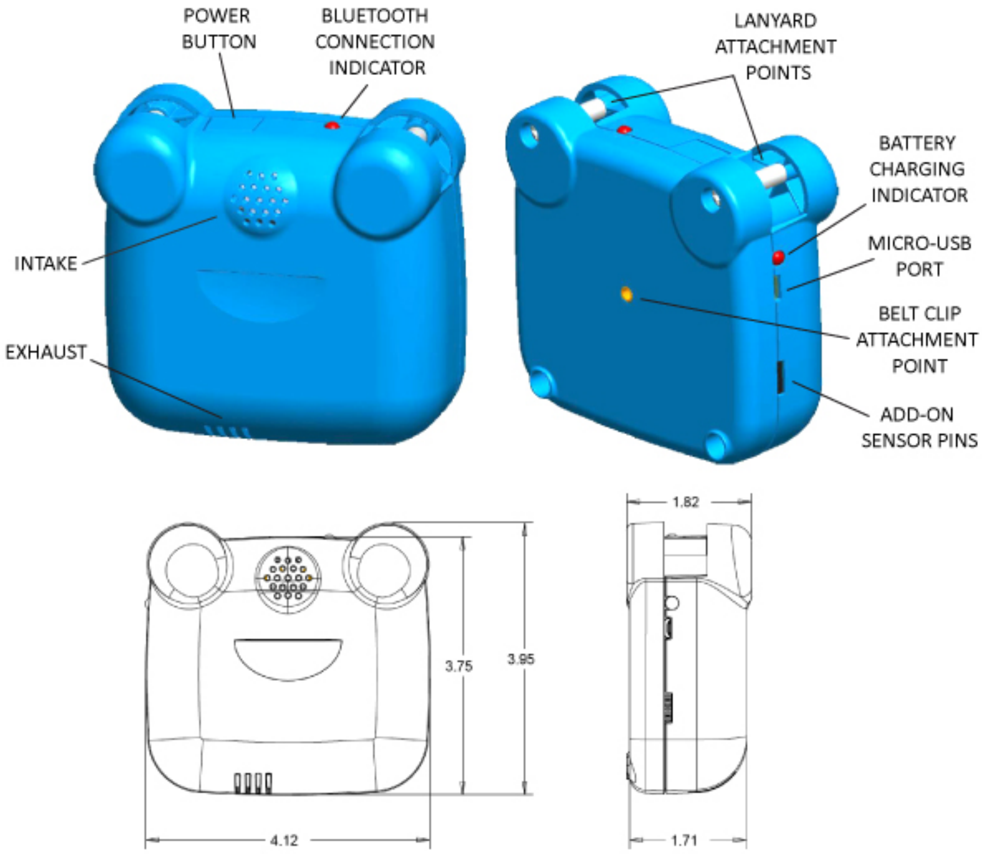


Figure 31 AirBeam design

4.3. Tools for qualitative data collection (feedback and report)

For collecting qualitative data in mobile mode using mobile phones or tablets, an application able to get positional data as well as multimedia and textual information is needed. It is also required that the application must be multi-platform and the most user-friendly as possible. There are several native-developed mobile APPs for smartphones but a web-based (HTML 5) customizable application accessible via standard web browser seems to be the more effective solution for the LOOPER project.

The web-based (HTML 5) application solution is the most suitable for the LOOPER project because:

- it is multiplatform (works with iOS, Android, Windows, etc.);
- it can be used with every device (laptop, tablet, etc.);
- it can be used during the ULLs workshops.

One of these kinds of web applications is actually hosted on www.geotagging.it. Geotagging.it is an experimental web based application that allows you to collect geo-data in a collaborative manner on the Internet. It can host an unlimited number of thematic sections (www.geotagging.it/something) each with its own configuration of content and graphics features as well as its own account database.

The geotagging.it front-end tool is optimized for mobile devices whose GPS allows immediate pointing of the area where the user is located. Geotagging.it user interface is actually in Italian but it is easily translatable in other languages.

4.3.1. Back-end section

The back-end section is a standard website providing with administrative functions.

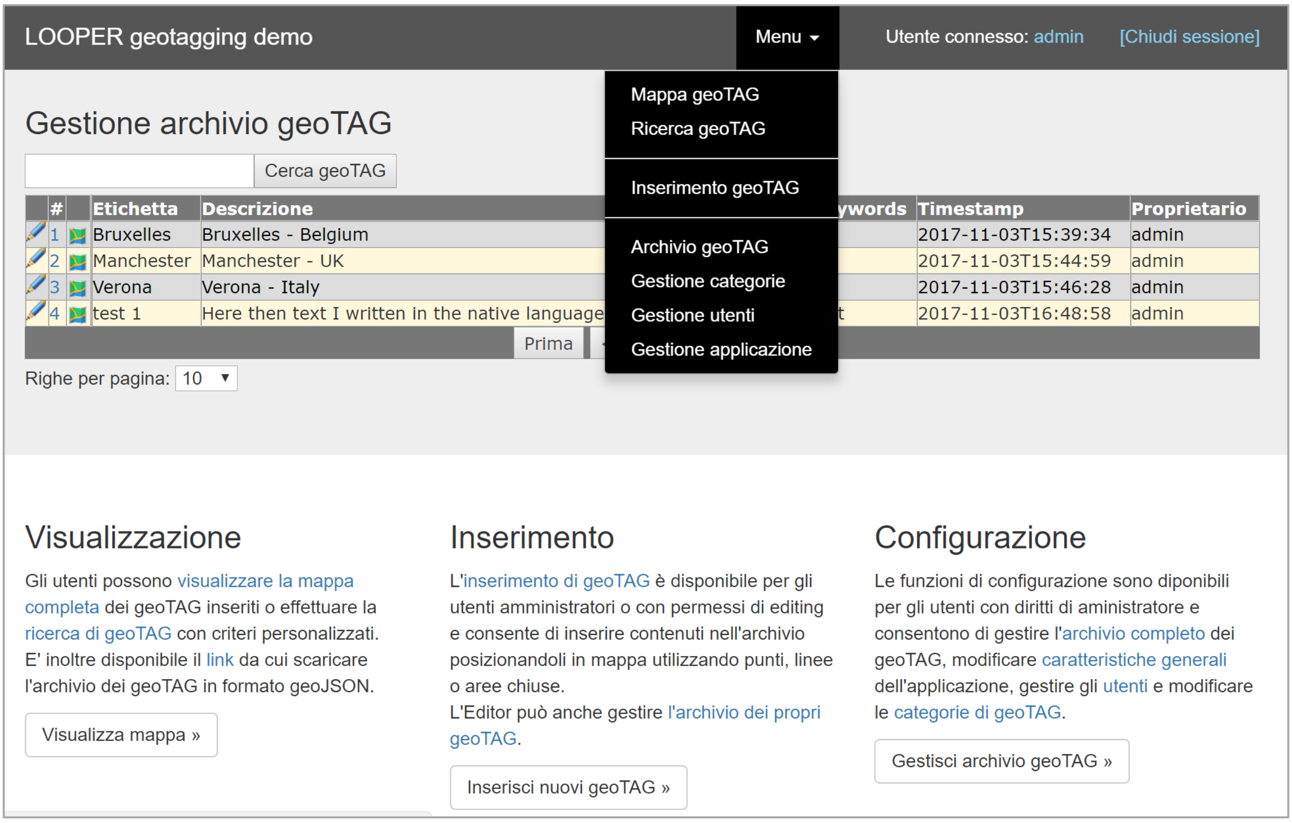


Figure 32: geotagging.it web application back-end section

Geotagging.it back end provides the below listed basic functions:

- Geo-tag archive explorer / manager;
- User accounts / groups manager with a 3-level permission system;
- Basic geo-tag map visualization;
- Basic geo-tag search engine;
- Geo-tag classification manager (to define custom classes for the front-end combo list);
- Overall application parameters configuration (e.g. section title, description, UI colours, etc. …).

Geotagging.it is a lightweight web-app based on Microsoft ASP.NET framework and MS Access portable database.

Geotagging.it provides direct access to the geo-tag database in geojson format through the address: www.geotagging.it/something/geoJSON.aspx link.

4.3.2. Front-end tool

Geotagging.it front-end tool is a frameless web-map page with a floating toolbox with which you can draw a geometry on a base map (point, line or polygon) and attach to it additional information as well as multimedia file.

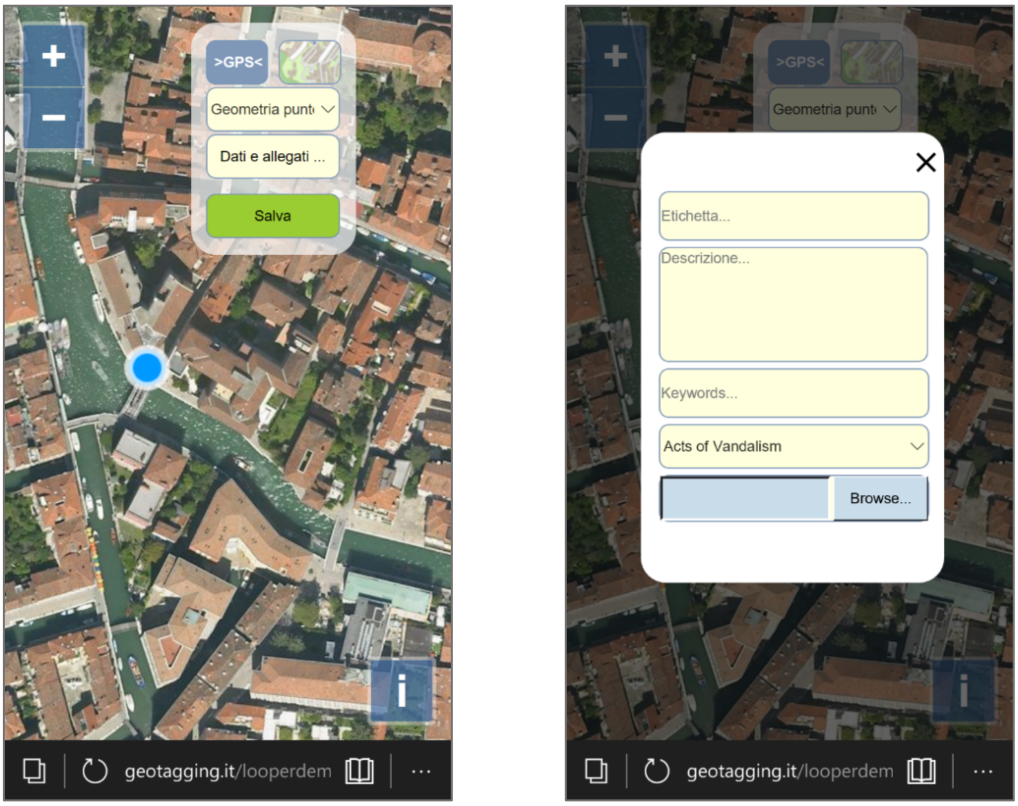


Figure 33: geotagging tool front-end

The geo-tag input form actually provides three text field, a combo box to choose a value from a preset list and a browse-button to attach a document, a picture or a video.

4.3.3. UI re-design proposal

To use the application for the LOOPER project, some minor changes to the user interface may be needed. Below some mock-up images of the re-design proposal.

Figure 34: geotagging tool mobile version mock-up for the home page

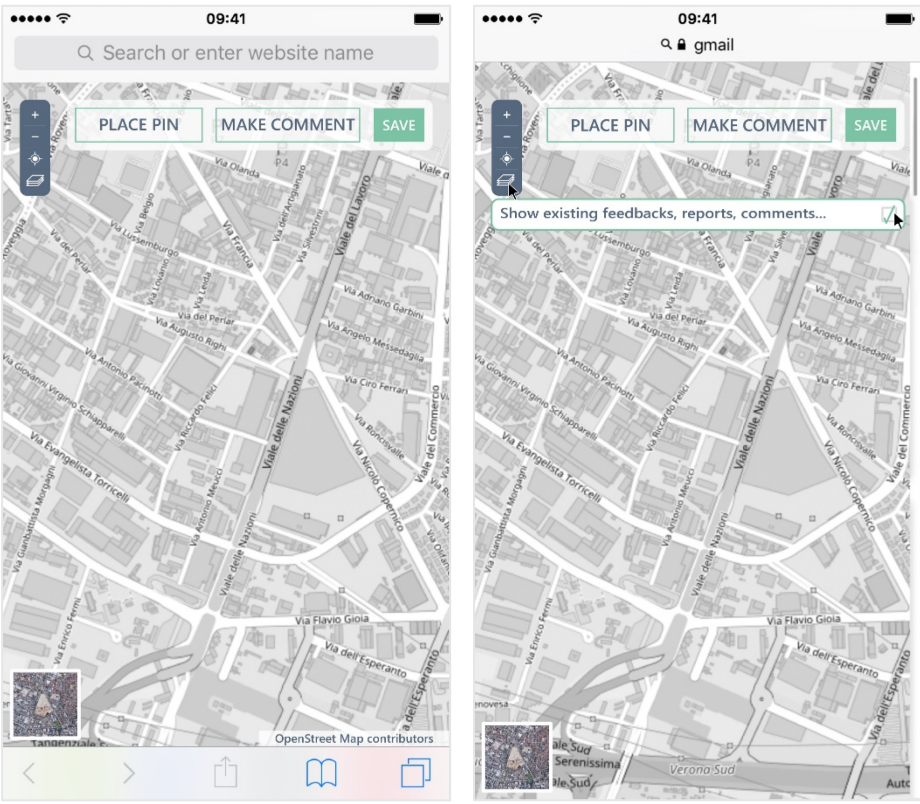


Figure 35: geotagging tool mobile version mock-up for place-pin mode

Figure 36: geotagging tool mobile version mock-up for data-entry mode

Summary of re-design requirements:

- Main theme re-design;
- Toolbar buttons re-design and responsive layout optimization;
- Add of a ranking value field;
- Minor changes for the back-end section.

5. URBAN LIVING LABS DATA COLLECTION PRELIMINARY SURVEY

5.1. Dataset characterization sheet

To setup the common framework data model we need to know which data will be collected and how they will be collected. For this scope, we need to make a preliminary survey using a simple form to gather those basic information from the LOOPER partners.

The form, basically a sort of "dataset characterization sheet", is detailed in the next paragraph.

5.1.1. Dataset characterization sheet template

Below the template of the form which is to be filled for each single type of expected survey data (e.g. air pollution; noise exposure; traffic …). It provides details about 5 different attribute classes about collection mode, subject, time, location and operator.

NOTES
…
1. WHAT
Object of the survey
a) What kind of data will be collected? b) What is the measurement unit or the detection unit (or classification for qualitative data) of the collected data?
2. HOW
Collection mode, tools and devices
a) Brief general description on how the survey campaign will take place. b) Which tools or devices will be used to collect data? c) Where will surveys datasets be firstly stored (before final submission to the platform database)?
3. WHEN
Time information and accuracy
a) Will the measurements / data sampling be made on a regular pace/frequency? What will approximatively be the pace/frequency? b) Will a timestamp be recorded? Will it be an instant timestamp or a time range/period?
4. WHERE
Location information and accuracy

a) Will a location data be recorded? If yes, will it be a strict latitude/longitude geographical position or an indirect reference (e.g. street address, neighbourhood name …)
b) How will the location be determined?

5. WHO
User/operator information

a) Will an operator userID/role be recorded for each measure/survey item? If yes, what user/operator role classification will be adopted?

Table 1: dataset characterization sheet template

5.2.Data collection preliminary survey results

5.2.1. Brussels

5.2.1.1. Table of data to be collected

ISSUE 1	TRAFFIC SPEED
	Data 1.1: Perception of traffic speed (qualitative data)
	Data 1.2: Speed per vehicle (handheld speed radar)
	Data 1.3: Speed per vehicle (fixed speed radar)

ISSUE 2	TRAFFIC VOLUME
	Data 2.1: Number and type of vehicles (pen and paper)
	Data 2.2: Number of vehicles (fixed speed radar)

Table 2: Brussels issue data tables

5.2.1.2. Dataset characterization sheets

Issue 1: Traffic speed
Data 1.1: Perception of traffic speed (qualitative data)

NOTES
The perception of traffic speed will be collected using an online geotagging tool which can be accessed at https://www.loopertagging.eu/brussels/
1. WHAT
Objective of the survey
a) Find spots/streets where residents have the perception that traffic generally drives too fast.
2. HOW
Collection mode, tools and devices
a) Qualitative geo-located observation will be collected by citizens using a web application on their smartphones. b) Using a geotagging tool, citizen can give a description of the situation, attach multimedia files, and gi e a 5-value ranking (from "very critical" to "optimal"). c) Geotagging tool will store data in its internal database then it will provide a geojson output as interchange data format.
3. WHEN
Time information and accuracy
a) Observation will be collected (without a regular frequency) during a one-month campaign. b) A timestamp will be recorded automatically for each geo-tag.
4. WHERE
Location information and accuracy
a) Point location data (latitude-longitude) will be recorded for each geo-tag. WGS84 geographical coordinate system (EPSG: 4326) will be adopted. b) Location will be determined by geotagging tool from on-board mobile device GPS.
5. WHO
User/operator information
a) An operator user ID will be recorded for each geo-tag. Operator user ID will be referred to a user list with role definition (e.g. ULL member, project team member).

5.2.1.3. Dataset characterization sheet

Issue 1: Traffic speed
Data 1.2: Speed per vehicle (handheld speed radar)

NOTES
The speed of vehicles will be collected by residents using a Bushnell speed radar gun.
1. WHAT
Objective of the survey
a) Measure the speed of vehicles in different locations, on different days and on different times of the day. b) Inform citizens on methods to collect traffic data.
2. HOW
Collection mode, tools and devices
a) Bushnell handheld speed radar. b) Pen and paper to note down registered speeds. c) Pen and paper to note down number of vehicles.
3. WHEN
Time information and accuracy
a) Data collection will take place during three ‘pop-up’ data collection moments with residents. b) The Bushnell mobile speed radar has an accuracy of +/- 1.6 km/h.
4. WHERE
Location information and accuracy
a) The data collection locations will be registered using pen and paper.
5. WHO
User/operator information
a) No operator user ID will be provided.

5.2.1.4. Dataset characterization sheet

Issue 1: Traffic speed
Data 1.3: Speed per vehicle (fixed speed radar)

NOTES
The speed per vehicles will be collected using a fixed speed radar (ICOMS TMS-SA) from the Brussels police department.
1. WHAT
Objective of the survey
a) Measure speed of vehicles before and after the implementation of the co-designed idea. b) Analyse whether the implemented idea has an effect on the speed of vehicles.
2. HOW
Collection mode, tools and devices
a) Fixed street radar from the Brussels police.
3. WHEN
Time information and accuracy
a) Speed data will be collected one week before and one week after the implementation of the co-designed idea. b) Data will be exported to an Excel-file with one line per measurement. Collected data include day, time, recorded speed, and vehicle category. c) The ICOMS TMS-SA has an accuracy of +/- 3 km/h.
4. WHERE
Location information and accuracy
a) The data collection location is based on the location of the co-designed idea.
5. WHO
User/operator information
a) No operator user ID will be provided.

5.2.1.5. Dataset characterization sheet

Issue 2: Traffic volume
Data 2.1: Number and type of vehicles (pen and paper)

NOTES
The number and type of vehicles will be collected by residents.
1. WHAT
Objective of the survey
a) Count the number of vehicles on different location, on different days and on different times of the day. b) Inform citizens on methods to collect traffic data.
2. HOW
Collection mode, tools and devices
a) Pen and paper. b) Participants count per 15 minutes the number of small vehicles, large vehicles, cyclists, and pedestrians.
3. WHEN
Time information and accuracy
a) Participants will write down the date and time.
4. WHERE
Location information and accuracy
a) Participants will note down the location and the direction of count, a well as the weather conditions.
5. WHO
User/operator information
a) No operator user ID will be provided.

5.2.1.6. Dataset characterization sheet

Issue 2: Traffic volume
Data 2.2: Number of vehicles (fixed speed radar)

NOTES
The number of vehicles will be collected using a fixed speed radar (ICOMS TMS-SA) from the Brussels police department.
1. WHAT
Objective of the survey
a) Count the number of vehicles driving through the street where the co-designed idea will be implemented.
2. HOW
Collection mode, tools and devices
a) Fixed street radar from the Brussels police.
3. WHEN
Time information and accuracy
a) Traffic volume data will be collected one week before and one week after the implementation of the co-designed idea. b) Data will be exported to an Excel-file with one line per measurement. Collected data include day, time, recorded speed, and vehicle category. c) The ICOMS TMS-SA has an accuracy of +/- 3 km/h.
4. WHERE
Location information and accuracy
a) The data collection location is based on the location of the co-designed idea.
5. WHO
User/operator information
a) No operator user ID will be provided.

5.2.2. Manchester

5.2.2.1. Table of data to be collected

ISSUE 1	**AIR QUALITY**
	Data 1.1: PM2.5 (DEFRA)
	Data 1.2: PM2.5 (AirBeam or Plume Flow)
	Data 1.3: NO_2 (DEFRA)
	Data 1.4: NO_2 (Air Monitor)
	Data 1.5: Health data related to Air Quality (Public Health England/MCC)

ISSUE 2	**TRAFFIC**
	Data 2.1: Qualitative data (Geotagging)
	Data 2.2: Traffic flow (UoM and TfGM)
	Data 2.3: Traffic accidents (policy)
	Data 2.4: Parking occupancy

ISSUE 3	**COMMUNITY SPACES**
	Data 3.1: Qualitative data (Geotagging and mapping)

Table 3: Manchester issue data tables

5.2.2.2. Dataset characterization sheets

Due to the difficulties found by organisers in finding participants with the needed abilities, it is not expected to use digital tools and sensors in Manchester. Therefore, there is no need to fill in the dataset characterisation sheets.

5.2.3. Verona

The area of South Verona has a complex situation both from an urban planning point of view and from the environmental pollution point of view. The area is closed in-between two major highways (south and west), the train station (north) and a canal (east) which makes it difficult to connect it with the services and structures of the rest of the city which lies beyond the train station. The number of infrastructures and the presence of the industrial area make it possible to find higher levels of noise and air pollution. Also, the geographical position of the area makes the accumulation of air pollution easier.

The data that will be collected for the project are therefore: PM2.5, PM10, NO_2 for air pollution; dB for noise pollution; citizen feedbacks and reports via geotagging tools.

5.2.3.1. Table of data to be collected

ISSUE 1	AIR QUALITY
	Data 1.1: PM 2,5 (ARPAV)
	Data 1.2: PM 2,5 (AirBeam)
	Data 1.3: NO_2 (Passive sensors ARPAV)
	Data 1.4: NO_2 (Air Monitor)

ISSUE 2	NOISE POLLUTION
	Data 2.1: dB(A) (OpeNoise)

ISSUE 3	TRAFFIC
	Data 3.1: Qualitative data

ISSUE 4	URBAN GREEN SPACES
	Data 4.1: Qualitative data

Table 4: Verona issue data tables

5.2.3.2. Dataset characterization sheets

Issue 1: Air quality.
Data 1.1: Official ARPAV air monitoring – Particulate Matter (PM 2.5)

NOTES
None
1. WHAT
Objective of the survey
b) Particular Matter 2.5 (PM2.5) air sampling c) Measurement unit: µg/m^3
2. HOW
Collection mode, tools and devices
d) Monitoring will be carried out by an official ARPAV base station. e) ARPAV air quality monitoring network is the official institutional regional network. Data will be stored in ARPAV central database. f) Surveys datasets will be provided in a CSV text file or XLS spreadsheet file.
3. WHEN
Time information and accuracy
c) Data sampling will be made in continuous for about 15 days. d) Time resolution is 1hour (Each CSV row contains 1hour sampling value).
4. WHERE
Location information and accuracy
c) Location data will be recorded using latitude and longitude and referred to a base station geographical position. WGS84 geographical coordinate system (EPSG: 4326) will be adopted. d) Location will be determined by querying Google Maps.
5. WHO
User/operator information
b) No operator user ID will be provided.

Issue 1: Air quality.
Data 1.2: Airbeam – Particulate Matter (PM2.5)

NOTES
None
1. WHAT
Objective of the survey
a) Particulate Matter (PM2.5) air sampling b) Measurement unit: µg/m3
2. HOW
Collection mode, tools and devices
a) Monitoring sensor box will be distributed to Urban Living Lab members that will be in charge of their use, care and get back. b) AirBeam wearable air monitor will be used. Data will be stored in the internal memory card. (see www.takingspace.org/aircasting/airbeam). c) Surveys datasets will be stored in an TXT delimited text file for each sensor box containing all recorded values.
3. WHEN
Time information and accuracy
a) Data sampling will be made in mobile mode by different users for periods of about 15 days. b) Time resolution is 1second (Each CSV row contains 1second sampling value).
4. WHERE
Location information and accuracy
a) GPS location data will be recorded in latitude/longitude format and written in the CSV file. WGS84 geographical coordinate system (EPSG: 4326) will be adopted. b) Location will be automatically recorded by device.
5. WHO
User/operator information
a) An operator user ID will be recorded for each used device. Operator user ID will be referenced to a user list with role definition (e.g. ULL member, project team member, ARPAV member …).

Issue 1: Air quality.
Data 1.3: Passive sensors – Nitrogen Dioxide (NO_2)

NOTES
None
1. WHAT
Objective of the survey
a) Nitrogen Dioxide (NO_2) air sampling b) Measurement unit: µg/m^3
2. HOW
Collection mode, tools and devices
a) Air pollution passive sensors will be distributed to Urban Living Lab members that will be in charge of their installations, care and get back. Samplers then will be sent to ARPAV labs that will analyse the cartridge and measure pollution values. b) Radiello™ Cartridge Adsorbents passive sensor will be used for sampling (http://www.sigmaaldrich.com/catalog/product/supelco/rad166?lang=en&region=GB) c) ARPAV laboratory will produce a CSV delimited text file containing all recorded values.
3. WHEN
Time information and accuracy
a) Data sampling will be made 1 time per loop in about 20 different location. Exposure time is about 2 weeks. b) It will be recorded in the CSV file start date and end date for each cartridge.
4. WHERE
Location information and accuracy
a) GPS location data will be recorded in format latitude/longitude and written in the CSV file. WGS84 geographical coordinate system (EPSG: 4326) will be adopted. b) Location will be determined by querying Google Maps.
5. WHO
User/operator information
a) An operator user ID will be recorded for each cartridge exposure. Operator user ID will be referenced to a user list with role definition (e.g. ULL member, project team member, ARPAV member …).

Issue 1: Air quality.
Data 1.4: Air Monitor – Nitrogen Dioxide ($N0_2$)

NOTES
None
1. WHAT
Objective of the survey
a) Nitrogen Dioxide (N02) air sampling b) Measurement unit: µg/m3
2. HOW
Collection mode, tools and devices
a) Monitoring sensor box will be distributed to Urban Living Lab members that will be in charge of their use, care and get back. b) AirCasting Air Monitor DIY low cost sensor will be used. Data will be stored in the internal memory card (see http://habitatmap.org/habitatmapdocs/HowToBuildAnAirCastingAirMonitor.pdf). c) Surveys datasets will be stored in an TXT delimited text file for each sensor box containing all recorded values.
3. WHEN
Time information and accuracy
a) Data sampling will be made in mobile mode by different users for periods of about 15 days. b) Time resolution is 1second (Each CSV row contains 1second sampling value).
4. WHERE
Location information and accuracy
a) GPS location data will be recorded in format latitude/longitude and written in the CSV file. WGS84 geographical coordinate system (EPSG: 4326) will be adopted. b) Location will be automatically recorded by device.
5. WHO
User/operator information
a) An operator user ID will be recorded for each used device. Operator user ID will be referenced to a user list with role definition (e.g. ULL member, project team member, ARPAV member …).

Issue 2: Noise pollution.
Data 2.1. Low cost noise data sampler/logger – Urban noise

NOTES
None
1. WHAT
Objective of the survey
a) Noise sampling b) Measurement unit: dB(A)
2. HOW
Collection mode, tools and devices
a) Low cost noise sensor box will be distributed to Urban Living Lab members that will be in charge of their installations, care and get back. b) Low cost noise sensor box made up by an Android smartphone connected to an external high-quality microphone. Data are stored in the internal mobile phone memory. c) Surveys datasets will be stored in an TXT delimited text file for each sensor box containing all recorded values.
3. WHEN
Time information and accuracy
a) Data sampling will be made 1 time per loop in about 20 different location. Exposure time is about 1 week. b) Time resolution is 10 minutes (Each TXT row contains 10 minutes average sampling value).
4. WHERE
Location information and accuracy
a) Location data will be recorded using latitude and longitude column in XLS sheet. WGS84 geographical coordinate system (EPSG: 4326) will be adopted. b) Location will be determined by querying Google Maps.
5. WHO
User/operator information
a) An operator user ID will be recorded for each sensor box exposure. Operator user ID will be referred to a user list with role definition (e.g. ULL member, project team member, ARPAV member …).

Issue 3: Traffic.
Data 3.1. Qualitative data

NOTES
None
1. WHAT
Objective of the survey
a) Road traffic condition b) Free description of the situation, optional multimedia file, 5-value ranking (from “very critical” to “optimal”)
2. HOW
Collection mode, tools and devices
a) Qualitative geo-located observation will be collected by citizens using a web application through their smartphones. b) Geotagging tool (see par. 4.3) c) Geotagging tool will store data in its internal database then it will provide a geojson output as interchange data format.
3. WHEN
Time information and accuracy
a) Observation will be collected (without a regular frequency) during campaign of about 2-3 months each learning loop concurrently with other surveys. b) A timestamp will be recorded automatically for each geo-tag.
4. WHERE
Location information and accuracy
a) Point location data (latitude-longitude) will be recorded for each geo-tag. WGS84 geographical coordinate system (EPSG: 4326) will be adopted. b) Location will be determined by geotagging tool from on-board mobile device GPS.
5. WHO
User/operator information
a) An operator user ID will be recorded for each geo-tag. Operator user ID will be referred to a user list with role definition (e.g. ULL member, project team member, ARPAV member …).

Issue 4: Urban green spaces.
Data 4.1. Qualitative data

NOTES
None
1. WHAT
Objective of the survey
a) Urban green space presence and condition b) Free description of the situation, optional multimedia file, 5-value ranking (from "very critical" to "optimal")
2. HOW
Collection mode, tools and devices
a) Qualitative geo-located observation will be collected by citizens using a web application through their smartphones. b) Geotagging tool (see par. 4.3) c) Geotagging tool will store data in its internal database then it will provide a geojson output as interchange data format.
3. WHEN
Time information and accuracy
a) Observation will be collected (without a regular frequency) during campaign of about 2-3 months each learning loop concurrently with other surveys. b) A timestamp will be recorded automatically for each geo-tag.
4. WHERE
Location information and accuracy
a) Point location data (latitude-longitude) will be recorded for each geo-tag. WGS84 geographical coordinate system (EPSG: 4326) will be adopted. b) Location will be determined by geotagging tool from on-board mobile device GPS.
5. WHO
User/operator information
a) An operator user ID will be recorded for each geo-tag. Operator user ID will be referred to a user list with role definition (e.g. ULL member, project team member, ARPAV member …).

6. COMMON FRAMEWORK FOR DATA MODELLING AND PROCESSING (ALPHA VERSION)

The common framework for data modelling and processing is a set of rules and protocols needed to manage in a unified way the data collected by all the three Urban Living Labs. The common framework basically consists of the database conceptual data model, which defines how the data will be stored and mutually related, and the set of procedures that will be used to process data according to the output needs.

In this phase, we describe the “alpha version” of the framework since problems and data needs by the three Urban Living Labs is not yet completely defined.

6.1.Conceptual data model

As conceptual model, this schema is a preliminary simplified version of the Entity-Relationship (ER) diagram of the LOOPER platform database. As in any ER diagram, main boxes stands for entities (e.g. database tables) and each connecting segment stands for the relationship between two entities which will be done by linking key attribute fields (Primary Key of the master entity and Foreign Key of the child entity).

LOOPER platform data model contains 4 main entities to manage data collections and several minor entities and lookup tables. Main entities are the 3 on the left of the diagram, “DATAStatSurvey”, “DATADistSurvey” and “DATAPointQual” (see chapter 0), that will record the three type of data collection and the fourth one, “Sensor”, which will record the devices/users that will be used to carry out the surveys. The other 7 (or more) entities will record auxiliary information and lookup values as detailed in the next paragraph.

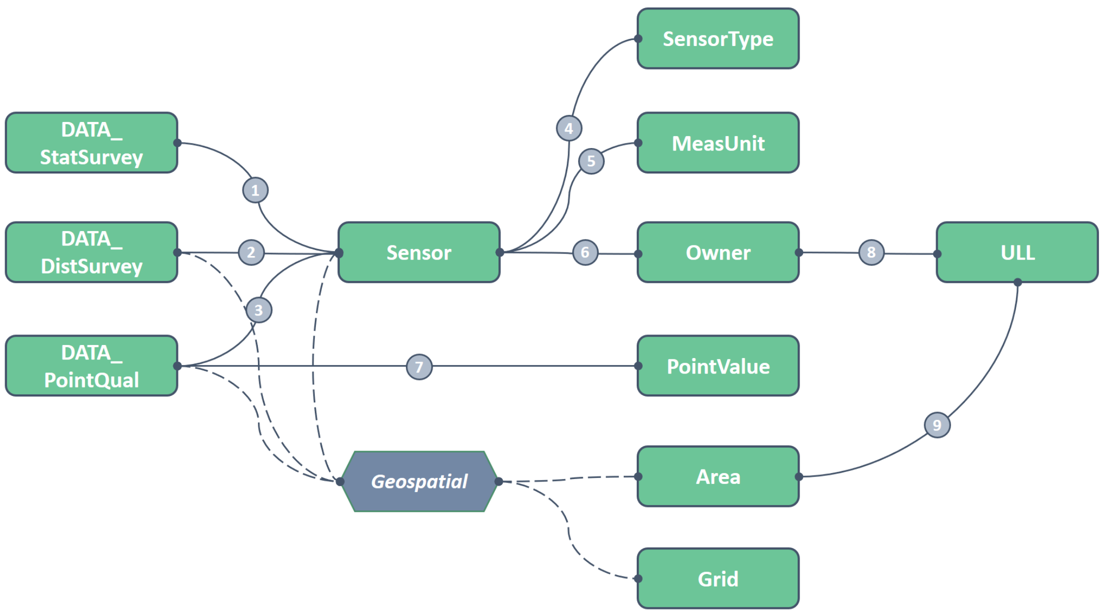


Figure 37: LOOPER platform database preliminary diagram

Since entities are detailed descripted in the next paragraph, with the table below we briefly explain each relationship shown in the diagram and that are numbered from 1 to 9.

Relationships		
#1 #2 #3	Relationships between the data values entities ("DATAStatSurvey", "DATADistSurvey", "DATAPointQual") and "Sensor" entity.	*We intend "Sensors" in a broad sense as "what is used to take samples", so sensors may be both physical devices or human observers.* *Relationships 1,2,3 are all one-to-many related from Sensor (one) to values data (many).*
#4	Relationship between each individual sensor and its specific type.	*One-to-many relationship from type to sensor since we can deploy n sensors of the same type.*
#5	Relationship between each sensor and its measure unit.	*One-to-many relationship from unit to sensor since each sensor provide only one type of data.*
#6	Relationship between sensors and its "owner" that is the responsible of the survey carried out with it.	*Theoretically one-to-many relationship since each sensor will be related to one user but virtually many-to-many if there will be groups of users owning a sensor*
#7	Relationship between point data value table and qualitative classes used to express ratings.	*Many-to-many relationship since same classes groups may be used by several (human) sensors.*
#8	Relationship between owners and Urban Living Labs. It basically expresses owner affiliation.	*Basic one-to-many relationship.*
#9	Relationship between geographically defined study area(s) and related Urban Living Lab.	*Basic one-to-many relationship.*
#Geospatial: (pseudo-relationships)		*Segments connecting "Geospatial" box indicate the "virtual joins" between all the entities that have positional/shape attributes since each of these tables can be visualized and processed as a map layer using a GIS software.*

Table 5: LOOPER platform database diagram relationships details

6.2.Data model details tables

The details of each entity are reported using the below table template:

ENTITY NAME			
Entity description			
Keys	Field name	Field data type	*Field description and notes.*

Table 6: data model detail table template

6.2.1. Main entities

According to the conceptual data model, there are 4 main entities intended to store samples data and sensors descriptive information.

DATAStatSurvey			
Data from stationary sensors (see par. 2.4). Each record is a single measure/observation; to rebuild value series filter by sensorid and order by starttime. join with sensor positional data in order to map values.			
PK	dataid	Numeric integer	*Sequential unique identifier.*
FK	sensorid	Numeric integer	*Foreign Key referring to Sensor entity Primary Key (see Sensor entity details table).*
	value	Numeric float	*Measured data value without unit.*
	starttime	Timestamp	*Timestamp (date + hour-min-sec) of start survey time.*
	endtime	Timestamp	*Timestamp (date + hour-min-sec) of end survey time. In case of instant sample, endtime = starttime.*
	note	Text	*Auxiliary note field.*

DATADistSurvey			
Distributed data collected by participatory surveys with portable measurements tools (see par. 2.5). Each record is a single measure/observation with relative positional data; to rebuild value series filter by sensorid and order by starttime. Related sensors will not have positional data.			
PK	dataid	Numeric integer	*Sequential unique identifier.*
FK	sensorid	Numeric integer	*Foreign Key referring to Sensor entity Primary Key (see Sensor entity details table).*
	value	Numeric float	*Measured data value without unit.*
	time	Timestamp	*Timestamp (date + hour-min-sec) of sample.*
	poslat	Numeric float	*Latitude in decimal degrees of positioning data*
	poslon	Numeric float	*Longitude in decimal degrees of positioning data*
	note	Text	*Auxiliary note field.*

(see whether could be useful to populate a cellid value with a scheduled/triggered procedure)

DATAPointQual			
Point qualitative data collected by participatory surveys with geotagging tools (see par. 2.6). Each record means a single observation. In this case we consider a specific topic (e.g. traffic, urban green, etc..) as a "pseudo-sensor" so to get value series by topic filter by sensorid and order by starttime.			
PK	dataid	Numeric integer	*Sequential unique identifier.*
FK	sensorid	Numeric integer / Text	*Composite Foreign Key referring to PointValue entity Primary Key (see PointValue entity details table).*
	value	Numeric integer	
	time	Timestamp	*Timestamp (date + hour-min-sec) of sample*
	attachment	Text	*Name/Code of attachment file (image, audio, video ...)*
	note	Text	*Auxiliary note field.*

(see whether could be better naming attachment files using a string with dataid value)

Sensor			
Sensors list. Each record of this table means both a hardware device used to acquire data and a qualitative issue observed by a human operator like traffic, urban green etc. which can also be intended as a sort of "observation class".			
PK	sensorid	Numeric integer / Text	*Sequential or alphanumeric-mnemonic unique identifier.*
	name	Text	*Mnemonic label of the sensor.*
FK	type	Numeric integer	*Foreign Key referring to SensorType entity Primary Key (see SensorType entity details table).*
FK	measunitid	Numeric integer	*Foreign Key referring to MeasUnit entity Primary Key (see MeasUnit entity details table).*
	poslat	Numeric float	*Latitude in decimal degrees of positioning data*
	poslon	Numeric float	*Longitude in decimal degrees of positioning data*
	stationary	Boolean	*Flag to mark stationary sensors.*
FK	ownerid	Numeric integer	*Foreign Key referring to Owner entity Primary Key (see Owner entity details table).*
	note	Text	*Auxiliary note field.*

6.2.2. Auxiliary/Lookup entities

The entities described below are basically "auxiliary" tables used to store data like series of key/value (known as *lookup tables*), users lists, classifications or other minor information.

SensorType			
Lookup/detail table of different measure units of sensors. Sensors will be both hardware device and a human-perceived qualitative issue; this entity has a two-level classification (type/category).			
PK	sensortypeid	Numeric integer / Text	*Sequential or alphanumeric-mnemonic unique identifier.*
	type	Text	*Description of sensor typology (e.g. "Mobile air sampling station", "Geotagging qualitative sampling" ...).*
	category	Text	*Category (sub-type; e.g. "Nitrogen Dioxide – NO2", "Traffic" ...). To be referred to "issues" (see par. 2.1)*
	note	Text	*Auxiliary note field.*

MeasUnit			
Lookup table of different measure units of sensors. Since sensors will be both hardware device and a human-perceived qualitative issue, measure units may be physical dimensions or a generically named "qualitative" measure unit.			
PK	measunitid	Numeric integer / Text	*Sequential or alphanumeric-mnemonic unique identifier.*
	measunit	Text	*Description of measure unit (e.g. "NO2", "PM2.5" ... "Qualitative")*
	note	Text	*Auxiliary note field.*

Owner			
Auxiliary account table of owner/manager of sensors.			
PK	ownerid	Text	*Account unique identifier. Might be a nickname or an email address.*
	fullname	Text	*User full name.*
	...	...	*(other details to be defined as needed)*
FK	ullid	Text	*Foreign Key referring to ULL entity Primary Key (see ULL entity details table).*
	note	Text	*Auxiliary note field.*

PointValue			
Lookup table of point qualitative sample values.			
PK	sensorid	Numeric integer / Text	*Qualitative value composite unique identifier. Value field might be a “ranking” value (e.g. “VeryHigh-High-Medium-Low-VeryLow”) or another additional quantitative component of the observation.*
PK	value	Numeric integer	
	valuedesc	Text	*Qualitative value extended description*
	note	Text	*Auxiliary note field.*

ULL			
Lookup table of Urban Living Labs.			
PK	ullid	Text	*Mnemonic unique identifier.*
	ullname	Text	*Name of the Urban Living Lab*
	…	…	*(other details to be defined as needed)*
	note	Text	*Auxiliary note field.*

Area			
Auxiliary geographic table of study areas. May be useful to build final maps or to carry out some geoprocessing operations.			
PK	areaid	Numeric integer	*Sequential unique identifier.*
	description	Text	*Area description.*
	geom	Geometry	*Geometric shape of the area.*
	…	…	*(other details to be defined as needed)*
FK	ullid	Text	*Foreign Key referring to ULL entity Primary Key (see ULL entity details table).*
	note	Text	*Auxiliary note field.*

Grid			
Auxiliary geographic table of mapping grid for distributed data. This is a polygon GIS layer of square/hexagonal cells covering all study area; grid cells will be used to aggregate quantitative data coming from participatory surveys and produce thematic maps visualizations.			
PK	cellid	Numeric integer	*Sequential unique identifier.*
	geom	Geometry	*Geometric shape of the cell.*
	…	…	*(other details to be defined as needed)*

(grid optimal size to be defined)

6.3.Protocols and procedures

Protocols and procedures will be defined to ensure effective data flows and processing from data collection storages and inside LOOPER platform. Basically, protocols will be implemented as guidelines and specification, while procedure will be implemented as DBMS *views* or *stored procedures*. It will be considered whether to setup one or more web services in order to ensure data transfer between multi-platform applications.

There are two phases of data processing: -pre- and -post- storage into the DBMS. The following paragraph briefly introduce the key issues that will be developed once the framework model is completed.

6.3.1. Pre-processing protocol

Pre-processing protocols will define how data collected with many different tools will be finally stored into the LOOPER platform database.

Devices used to carry out surveys will store data in various ways. We need specific procedures to ensure easy and efficient data migration into LOOPER platform database, taking into account the different tools that will be used and different users that will be involved in data collection.

Basically, we can assume to have (at least) two types of scenarios that will make necessary to follow two different protocols:

1. Tools based on a “closed” data management system that provides some standard interchange format;
2. Tools based on a connectable (“open”) data management system.

For the first scenario, the most probably way to ensure data migration may be to follow some case-specific guidelines and specifications to put collected data into a spreadsheet or delimited-text file according to a strict attribute schema. We can consider to setup a basic web upload system to make the process more fluid and user-friendly.

For the second scenario, for which the needs have to be verified case-by-case, it may be a good solution to establish a remote database connection through LOOPER server and other application servers, or even to develop a specific web-service to schedule regular updating of the main database from remote repository.

6.3.2. Post-processing

Post processing procedures are needed to make users able to perform data visualization, reporting, searching and identifying.

There are two fundamental type of procedures to process data in a DBMS: *views* and *stored procedures*. Mostly we will create views directly to output results (on the fly) from storing tables and stored procedures to perform complex processing that needs to be saved as static dataset due to long computing time.

We must use views every time we need real-time update of the outputs, e.g. maps showing survey data changing every hour (or even one time per day). Probably in LOOPER project there will not be this need and we could do many of the computations using stored procedures and temporary output tables.

Below a simple schematic list of probably needed procedures for the LOOPER platform database related to functions and data typology:

data typology functions	Stationary survey data	Mobile survey data	Qualitative data
Visualization and reporting	*Output of map pin layer with simple thematic symbol*	*Output of map grid layer with thematic symbol of average values*	*Output of map pin layer with thematic symbol*
Map objects identifying	*Single pin with info-box summary data about all issues*	*Single cell summary of data values about a time range (e.g. last week, weekends, peak hours etc. …)*	*Single pin info-box with attachment URLs (e.g. picture, video etc.)*
Data searching	*Output lists / reports with summary data (e.g. number of records found, average values etc.)*	*Filtered map grid layer according to search criteria.*	*Output lists / reports with summary data (e.g. number of records found, average values etc.)*

7. ACKNOWLEDGEMENTS

The support of 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 is gratefully acknowledged.