



# URBAN CO-CREATION DATA LAB

## DATA SETS AND REQUIREMENTS FOR THREE CITIES

December 2021

<b>Milestone Title</b>	<b>MS 2 - Data sets and requirements defined for all 3 cities</b>
<b>Related Activity</b>	Activity 1: Data preparation and open data infrastructure assessment
<b>Related Task</b>	Task 1.1 Data definition and requirements definition
<b>Author's</b>	NOVA IMS
<b>Due Submission Date</b>	31.12.2020
<b>Actual Submission</b>	31.12.2021
<b>Action Number</b>	2018-EU-LA-0099
<b>Agreement Number</b>	INEA/CEF/ICT/A2018/1837945
<b>Instrument</b>	Connecting Europe Facility (CEF) – Telecommunication Sector
<b>Starting Date of Action</b>	01.10.2019
<b>Completion Date of Action</b>	28.02.2022
<b>Main objective</b>	Identify data sources e.g., from operational systems, mobile technologies, devices, open data infrastructures and platforms with special emphasis in Lisbon which can be replicated in the other two cities

### Table of Contents

1	Executive Summary .....	4
2	Objectives.....	4
2.1	Strategic Objective.....	4
2.2	Operational Objective .....	4
2.3	Tasks .....	4
2.4	Outputs:.....	5
2.5	Milestones and means of verification .....	5
3	Methodology.....	6
3.1	Understanding the municipality business needs.....	6
3.2	Identification and understanding of the datasets and their characteristics .....	6
4	Annex.....	7

# 1 Executive Summary

This report is the mean of confirmation of **Milestone 2 - Data sets and requirements defined for all 3 cities** of the Activity 1: Data Preparation and open data infrastructure assessment of the Action **Urban Co-creation Data Lab**, funded by the European Commission under the H2020ICT- 28-2017 Connecting Europe Facility (CEF) – Telecommunications Sector under the grant agreement n. INEA/CEF/ICT/A2018/1837945. In this report is presented the approach considered for the business understanding in the Lisbon City Council. Are also presented the datasets used for the development of the use cases in Lisbon and in the Provincial Council of Badajoz.

## 2 Objectives

### 2.1 Strategic Objective

The main objective of this activity was to prepare data and open data infrastructure for Smart Management Platform.

### 2.2 Operational Objective

Data preparation and open data infrastructure assessment.

### 2.3 Tasks

The development of this activity included the implementation of following tasks:

#### Task 1.1: Data Definition and requirements definition

This task intended to identify, characterise, evaluate, and increase the overall quality of the data that will be used in the first testing of the UCD Lab services in Lisbon. It involved a set of activities that will improve the overall quality and consistency, namely:

- Identification of data sources e.g., from operational systems, mobile technologies, devices, open data infrastructures and platforms with emphasis in Lisbon, but which can be replicated in the other two cities;
- Assessment of the quality of the data available in terms of accuracy, usefulness, reliability, latency, etc. and perform Data/Metadata Harmonization and Validation for preparing the UCD Lab services experiments;
- Identification of further stakeholders and end-users' needs and requirements to define the necessary features and functions of the platform particularly regarding the proposed services;
- Final definition and harmonization of the data sets, for the 5 services, with stakeholders and end-users' needs and requirements definition.

In the case of Lisbon, the set of selected open data from Lisboa Aberta and Lisboa Smart Management Platform (LSMP) was supposed to be made for the 5 proposed analytical services:

1. Assessment of human flows in highly crowded amusement areas;

2. Analytics to predict patterns in the production of solid urban waste;
3. Identify patterns and impact of illegal parking;
4. Predictive analytics for propagation of pollution in cities;
5. Predictive analytics for impact of events in mobility/transport.

Nevertheless, and after the global assessment of the data available quality and the identification of the stakeholders and end-users' needs and requirements, the proposed analytical services were reviewed as follows:

- MOBILITY - Evaluation and prediction of patterns and behaviours of micro mobility in the city of Lisbon, to support new planning and management approaches altogether with new tools to evaluate impact and prediction of micro mobility user behaviour;
- WASTE MANAGEMENT - Identification of patterns/profiles and solid waste production prediction in the city of Lisbon, to identify patterns to support the prediction of the production of urban waste associated with a variety of context information (e.g. events, climate situation, etc.);
- PARKING - Identification of patterns and prediction of irregular parking in the city of Lisbon to improve surveillance efficiency on irregular parking in Lisbon;
- POLLUTION - Elaboration of predictive models for the propagation of liquid and atmospheric pollutants in the city of Lisbon, to simulate their propagation after accidents with hazard substances in Lisbon;
- EMERGENCY – Identification of patterns and predictive modelling of traffic accidents in the city, to optimize the response of firefighters and civil protection services to emergency due to traffic accidents in Lisbon.

Due to the pandemic the selection of the two additional cities for testing and validation of local use cases during the Action, had to be delayed. Nevertheless, the data selection and definition will be performed also for the Provincial Council of Badajoz, namely for the development of use case #2 Waste management.

### Task 1.2: Services and Use Cases refinement

Under this task, the services and use cases full definition was produced after:

- Refining and analysing further the initial set of proposed services utilization scenarios (use cases) and identify those with highest priority that could serve as reference implementations in the co-creation labs and city validation;
- Providing more detailed specifications for the city services validations, clearly identifying the similarities and differences between them.

## 2.4 Outputs:

Data sets definition and requirements definition for each use case.

## 2.5 Milestones and means of verification

Milestone number	Milestone description	Indicative completion date	Completion date	Means of verification
2	Data sets and requirements defined for all 3 cities	31/12/2020	31/12/2021	Data sets definition and requirements definition completed and available for Lisbon city

### 3 Methodology

The methodology developed in this phase of the Action was implemented in two stages:

- I. understating the municipality business needs for each dimension of the city addressed in this Action (i.e., micro mobility, waste management, parking, pollution and emergency);
- II. identification and understanding of the datasets and their characteristics available in the Lisbon Smart Management Platform (LSMP) considering Lisbon municipality business needs, and datasets for the development of use case #2 Waste management for the Provincial Council of Badajoz.

#### 3.1 Understanding the municipality business needs

In this first stage of the Action, several meetings were made between the research team that is developing the technical work, the Lisbon Urban Intelligence Management Centre (CGIUL) (responsible for the Lisbon municipality data management), and departments of Lisbon municipality, namely the Mobility Department, Urban Hygiene Department, Civil Protection Service, Sanitation Department, the Municipality Police and external entities (EMEL – a parking and mobility company that supports Lisbon Municipality). The main goal of these meetings was to understand several problems that these departments face in operational terms regarding their daily activity that are not being addressed and understand what are their mainly activities in the municipality operation. The intention was to provide a first view of possible use cases that these departments could have interest in developing. To notice that this was an iterative process along this phase of the Action until the final definition and refinement of the use cases that will be developed in this Action.

Also, meetings with the Provincial Council of Badajoz were carried out, namely with Promedio, a municipal company that is responsible for the waste management in the Provincial Council of Badajoz.

#### 3.2 Identification and understanding of the datasets and their characteristics

In this phase of the Action a literature review was carried out by the research team to identify the relevant datasets that are used as explanatory variables to address use cases regarding each one of the city dimensions addressed in the Action (i.e., micro mobility, waste management, parking, pollution, and emergency). Several meetings and requests were made between the research team and the CGIUL to assess the available datasets in the LSMP that have a potential interest for the development of the case studies. The identification of the available datasets and their characteristics allowed the beginning of the use cases refinement. Several characteristics about the identified datasets (see Annex 1) were collected to increase the understating about the available data. The data characteristics and their description are presented in Table 1.

Data characteristics	Description
GROUP	Thematic group that the data set belongs
DATASET	Identification of the dataset
DESCRIPTION	Description of the dataset
SOURCE	Source (organization/portal/platform) that holds the dataset
FORMAT	File format of the dataset
INITIAL DATE	Date of the first record in the dataset

<b>Data characteristics</b>	<b>Description</b>
END DATE	Date of the last record in the dataset
SPATIAL REPRESENTATION	Spatial representation of the dataset (data model of the dataset)
ATTRIBUTES	Attributes of the dataset
TYPE	Type of the attributes in the dataset
UPDATING FREQUENCY	Time frequency or temporal resolution of the dataset
COMMENTS	Field to assign any necessary comment to the dataset
STATUS	Status of availability to use in the development of the analytical models
#1 MICROMOBILITY	The dataset is checked if it will be used in the development of the case study #1 - Micromobility
#2 WASTE MANAGEMENT	The dataset is checked if it will be used in the development of the case study #2 - Waste Management
#3 PARKING	The dataset is checked if it will be used in the development of the case study #3 - Parking
#4 POLLUTION	The dataset is checked if it will be used in the development of the case study #4 - Pollution
#5 EMERGENCY	The dataset is checked if it will be used in the development of the case study #5 - Emergency

*Table 1. Datasets characteristics collected and respective description in the assessment of the datasets needed for the development of the case studies.*

## 4 Annex

Annex 1 – Data sets used for the development of the use cases in Lisbon

Annex 2 – Data sets used for the development of the use case in the Provincial Council of Badajoz

### Annex 1 – Data sets used for the development of the use cases in Lisbon

GROUP	DATASET	DESCRIPTION	#1 Micromobility	#2 Waste management	#3 Parking	#4 Pollution	#5 Emergency
Education	Public schools - 1st cycle	General information and location of 1st cycle schools (public network) in the city of Lisbon			X		
Education	Private schools - 1st cycle	General information and location of private 1st cycle schools in Lisbon			X		
Education	Public schools - 2nd cycle	Generic information and location of school's 2nd cycle (public network) in the city of Lisbon			X		
Education	Private schools - 2nd and 3rd cycle	General information and location of 2nd and 3rd cycle private schools in the city of Lisbon			X		
Education	Public schools - 3rd cycle	General information and location of 3rd cycle schools (public network) in the city of Lisbon			X		
Education	Public schools - pre-scholar	General information and location of pre-schools (public schools) in the city of Lisbon			X		
Education	Private schools - pre-scholar	General information and location of private kindergartens in Lisbon			X		
Education	Public schools - Secondary	Generic information and location of public secondary schools in the city of Lisbon			X		
Education	Private schools - Secondary	General information and location of private secondary schools in the city of Lisbon			X		
Education	Higher education institutions	General information on "Higher Education Institutions" strategic actors located in the city and region according to clusters, selected by the Direção Municipal de Economia e Inovação.			X		
Environment	Weather data	Weather data recorded in the three weather stations in Lisbon	X		X	X	X
Environment	Weather station's location	Location of the three weather stations in Lisbon	X		X	X	X
Environment	Digital terrain model	Digital terrain model of Lisbon				X	
Health	Health centres	General information and location of health centres in the city of Lisbon			X		
Health	Public hospitals	General information and location of public hospitals in the city of Lisbon			X		
Health	Private hospitals	General information and location of private hospitals in the city of Lisbon			X		
Infrastructures	Traffic lights	Intersections with traffic lights					X
Infrastructures	Road network	Road network of Lisbon			X		X
Infrastructures	City buildings 2.5 D model	This dataset provides information about buildings occupation area and their height				X	
Mobility	GIRA docks occupation	This data set provides information about the dock's occupation of GIRA service	X				
Mobility	GIRA station's location	Location of GIRA stations	X				
Mobility	Train stations	General information and location of train stations in the city of Lisbon			X		
Mobility	Metro stations	General information and location of metro stations in Lisbon			X		



## MS 2 - Data sets and requirements defined for all 3 cities

GROUP	DATASET	DESCRIPTION	#1 Micromobility	#2 Waste management	#3 Parking	#4 Pollution	#5 Emergency
Mobility	Bus stations	Location of BUS stops			X		
Parking	Illegal parking occurrences	Parking illegalities reported by the municipality police regarding illegal parking			X		
Safety	Traffic accidents	Traffic accidents occurrences recorded by the firefighters					X
Urban hygiene	Garbage collection circuits	This dataset provides information about the location of garbage collection circuits		X			
Urban hygiene	Mixed waste collection loads	This dataset provides information about the total amount of garbage collected for each collecting circuit in a specific day		X			

### Annex 2 – Data sets used for the development of the use case in the Provincial Council of Badajoz

GROUP	DATASET	DESCRIPTION	#1 - Micromobility	#2 - Waste management	#3 - Parking	#4 - Pollution	#5 - Emergency
Urban hygiene	Mixed waste collection loads	This dataset provides information about the total amount of garbage collected for each collecting circuit in a specific day		X			
Social economic	Employment rate	Variation coefficient of the annual employment rate		X			