URBANITE

Supporting the decision-making in urban transformation with the use of disruptive technologies

Deliverable D3.1

URBANITE Mobility Data Sources Analysis

Editor(s):	Fritz Meiners (FhG)
Responsible Partner:	Fraunhofer FOKUS
Status-Version:	Final (Privisional)
Date:	30.09.2020
Distribution level (CO, PU):	PU

Project Number:	GA 870338
Project Title:	URBANITE

Title of Deliverable:	URBANITE Mobility Data Sources Analysis
Due Date of Delivery to the EC:	30/09/2020

Workpackage responsible for the Deliverable:	WP3 - Data Management Platform
Editor(s):	JSI, Fraunhofer FOKUS
Contributor(s):	TECNALIA, C. Messina, ENG, BIL, MLC
Reviewer(s):	Alina Machidon (JSI)
Approved by:	All Partners
Recommended/mandatory readers:	WP4, WP5

Abstract:	Document that presents the results of the analysis of the data sources that are available to the project use cases. It will contain, among others, the available data structures, schematics and APIs for each data source.
Keyword List:	Data sources, overview
Licencing information:	This document is licensed under Creative Commons Attribution-ShareAlike 4.0 International (CC-BY-SA 4.0) https://creativecommons.org/licenses/by-sa/4.0/
Disclaimer	This document reflects only the author's views and neither Agency nor the Commission are responsible for any use that may be made of the information contained therein.

Document Description

Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
v0.1	28/04/2020	Draft ToC	FhG
v0.2	04/05/2020	Updated ToC	FhG
v0.3	21/07/2020	Added data sources	All
v0.4	28/08/2020	First draft	FhG
V0.5	11/09/2020	Final draft	JSI, FhG
V0.6	28/09/2020	Updates according to internal review	FhG
V1.0	30/09/2020	Final version	Tecnalia

Table of Contents

List of Figures

List of Tables

Terms and abbreviations

API	Application Programming Interface
BC	Black Carbon
CC	Creative Commons
CSV	Comma Separated Values
DB	Data Base
dBA	A-weighted Decibels
EC	European Commission
FTA	Finnish Transport Agency
GIS	Geographic Information System
GML	Geography Markup Language
GPS	Global Positioning System
HTML	HyperText Markup Language
HTTP	HyperText Transport Protocol
JSON	JavaScript Object Notation
MIF/MID	MapInfo Interchange Format
ODbL	Open Data Commons Open Database License
OSM	Open Street Map
PDF	Portable Document Format
REST	Representational State Transfer
RDW	Specific Open Data Portal of Amsterdam
SOAP	Simple Object Access Protocol
SPDP	Standard for Publishing Dynamic Parking Data
URL	Uniform Resource Locator
UTC	Coordinated Universal Time
WFS	Web Feature Service
WMS	Web Map Service
XLSX	Excel file Format
XML	eXtensible Markup Language
XSD	XML Schema Definition

Executive Summary

This deliverable contains a collection of data source descriptions for use in the URBANITE ecosystem. These data sources semantically cover the four pilots Amsterdam, Bilbao, Helsinki, and Messina. Additionally, a selection of data sources that are not associated with a pilot, but provide data from common domains, is included. Ultimately, the data harvested from these sources to be stored in the URBANITE data management platform will serve as input to the recommender engine developed in WP4. Therefore, use-case related input generated by WP6 is also included in this deliverable. Based on the insights gained from this deliverable the next step can commence. Due to the heterogeneous nature of the data sources it is vital to transform data from a common domain into a standardised structure/model. This is described in deliverable D3.4.

1 Introduction

The recommendation engine developed in the URBANITE context works on data. The kind of data, amount, and timeliness that is required varies between the various use cases. This deliverable shall provide an overview of the data sources available to the project.

1.1 About this deliverable

All pilots were asked to contribute data sources from their municipalities that they believed to be relevant. Additionally, FhG researched non-pilot data sources that are not limited to either of the pilots. Naturally, the data sources all vary greatly with regards to the quality of documentation, structure, and licencing, among others. Hence, further research may be required before production usage. It is also important to note that the data sources here can only be a snapshot of what was available at the time of writing. Also, the scope is limited by resources and the page limit that applies to deliverables.

1.2 Document structure

The data sources are structured by the pilots that contributed them. Hence, the data sources of Amsterdam, Bilbao, Helsinki, and Messina can be found in sections 2 to 5, respectively. However, some data sources are not attributed to a single pilot, but are available on the web and cover a varying number of cities and/or regions. These are categorized as “Other Data Sources” and can be found in section 6. Annex A contains a table that summarizes all data sources described in this document. For each data source, a short introduction is given, as well as information on the data structure. For a majority of cases, examples illustrate the content served. API endpoints are described where applicable.

2 Amsterdam Data Sources

This section contains descriptions of data sources from Amsterdam.

2.1 Amsterdam Maps

This data source[1] provides a collection of interactive maps about the City of Amsterdam. The maps cover different topics, for example, urban development & housing, traffic & infrastructure, and neighbourhood & amenities. Currently, about 100 maps and relative datasets are available.

2.1.1 Data Format and Structure

Datasets of these maps are provided in four different formats: CSV, GeoJSON, and MIF/MID. The two latter formats are for GIS (Geographic Information System) applications. The two files must be used together, since MIF files provide information about geometries; while MID files provide information about the attributes of those geometries. All data is served as static files[2].

The CSV files include different fields for the specific purposes of each dataset. Common fields are OBJECTNUMMER (an incremental ID), WKT_LNG_LAT and WKT_LAT_LNG (providing geographic information structured on the Well-known text - WKT specifications[3]), and LNG and LAT (coordinates in WGS84 format). An example of a CSV file is shown in Figure 1.

A	B	C	D	E	F	G	H
OBJECTNUM	Lijnnummer	Type_stadss	Type_hoofds	WKT_LNG_LAT	WKT_LAT_LNG	LNG	LAT
1	1	Stadsstraat	Hoofdstraat	LINESTRING(4.79098 52.356388,4.803471 52.358883)	LINESTRING(52.356388 4.79098,52.358883 4.803471)	47.972.305	523.576.285
2	2	Potentiele st	Hoofdstraat	LINESTRING(4.803471 52.358883,4.803849 52.358197,4.804017 52.357994,4.804287 52	LINESTRING(52.358883 4.803471,52.358197 4.803849,52.357994 4.804017,52.357828 4.80	48.058.455	523.583.145
3	3	Stadsstraat	Hoofdstraat	LINESTRING(4.793984 52.350854,4.801804 52.352422)	LINESTRING(52.350854 4.793984,52.352422 4.801804)	4.797.899	52.351.631
4	4	Potentiele st	Hoofdstraat	LINESTRING(4.827695 52.342603,4.827739 52.343231,4.827755 52.343725,4.827749 52	LINESTRING(52.342603 4.827695,52.343231 4.827739,52.343725 4.827755,52.344163 4.82	482.758	52.347.563
5	5	Stadsstraat	Hoofdstraat	LINESTRING(4.827399 52.352537,4.827306 52.356251,4.827224 52.358256)	LINESTRING(52.352537 4.827399,52.356251 4.827306,52.358256 4.827224)	48.273.165	523.553.895
6	6	Potentiele st	Hoofdstraat	LINESTRING(4.827224 52.358256,4.827109 52.363999,4.836487 52.364104,4.840442 52	LINESTRING(52.358256 4.827224,52.363999 4.827109,52.364104 4.836487,52.364131 4.84	4.835.642	52.361.204
7	7	Potentiele st	Hoofdstraat	LINESTRING(4.836487 52.364104,4.83641 52.365853,4.836385 52.366692,4.836476 52.	LINESTRING(52.364104 4.836487,52.365853 4.83641,52.366692 4.836385,52.367479 4.836	4.837.421	523.684
8	8	Potentiele st	Hoofdstraat	LINESTRING(4.830937 52.370625,4.837227 52.369709,4.838962 52.369513,4.840421 52	LINESTRING(52.370625 4.830937,52.369709 4.837227,52.369513 4.838962,52.369414 4.84	4.839.709	523.700.125
9	9	Potentiele st	Hoofdstraat	LINESTRING(4.846908 52.372457,4.845841 52.372267,4.843447 52.372203,4.842103 52	LINESTRING(52.372457 4.846908,52.372267 4.845841,52.372203 4.843447,52.372222 4.84	48.350.235	523.736.325
10	10	Stadsstraat	Hoofdstraat	LINESTRING(4.819021 52.375702,4.820113 52.378349,4.821402 52.381402,4.832222 52	LINESTRING(52.375702 4.819021,52.378349 4.820113,52.381402 4.821402,52.379749 4.83	48.373.465	523.790.745

Figure 1: Amsterdam Maps CSV example

2.1.2 Licence

Datasets of the maps are provided under the terms of a specific licence[4]. Dataset can be used both for non-commercial and commercial purposes.

[1] https://maps.amsterdam.nl
[2] https://maps.amsterdam.nl/open_geodata
[3] https://www.ogc.org/standards/wkt-crs
[4] https://maps.amsterdam.nl/open_geodata/terms.php

2.2 Netherlands RDW

The RDW[5] is an open data portal focusing on traffic and mobility data from the entire country of the Netherlands. Also included are parking related data. There are currently roughly 1000 datasets available. The portal is powered by Socrata[6], a data platform developed for governments. Despite the ability to toggle between Dutch and English, most of the site stays in Dutch regardless. The available datasets are grouped into categories. The most relevant are listed below:

- Car and Lorry registration
- Info on parking lots
- Dynamic parking data

2.2.1 Data Format and Structure

All datasets contain tabular data. The dynamic parking data follows the SPDP[7].

2.2.2 API

Since the portal is based on Socrata their APIs apply[8]. An example is shown in Figure 2.


```
    "resource": {
      "name": "Total Percentage of Available Funds Paid Out by the Commission 2007-2013",
      "id": "w8x7-cqjd",
      "parent_fxf": Array[1][↔],
      "description": "This data is refreshed every 24 hours. EU Cohesion Policy (European Regional Development Fund + Cohesion Fund + European Social Fund) including European Territoria
Cooperation.  Percentage of funds paid (including interim payments and pre-financing) compared to total available budget.  Source: Infoview > Financial Management >
Financial_execution_by_period_fund_country",
      "attribution": "European Commission DG REGIO",
      "attribution_link": "http://ec.europa.eu/regional_policy/index_en.cfm",
      "contact_email": "",
      "type": "chart",
      "updatedAt": "2020-07-16T05:47:48.000Z",
      "createdAt": "2015-12-04T14:57:26.000Z",
      "metadata_updated_at": "2019-12-18T09:55:19.000Z",
      "data_updated_at": "2020-07-16T05:47:48.000Z",
      "page_views": {↔},
      "columns_name": Array[7][↔],
      "columns_field_name": Array[7][↔],
      "columns_datatype": Array[7][↔],
      "columns_description": Array[7][↔],
      "columns_format": Array[7][↔],
      "download_count": 1460,
      "provenance": "official",
      "lens_view_type": "tabular",
      "blob_mime_type": null,
      "hide_from_data_json": false,
      "publication_date": "2017-06-07T10:16:00.000Z"
    },
    "classification": {↔},
    "metadata": {↔},
    "permalink": "https://cohesiondata.ec.europa.eu/d/w8x7-cqjd",
    "link": "https://cohesiondata.ec.europa.eu/2007-2013/Total-Percentage-of-Available-Funds-Paid-Out-by-th/w8x7-cqjd",
    "owner": {↔}
  },
```


Figure 2: Amsterdam RDW example

2.2.3 Licence

All datasets are published under the Creative Commons Zero licence. RDW requests that "it is not permitted to state that the data comes from the RDW upon reuse, and it is not permitted to use the RDW logo or corporate identity in the developed applications"[9].

[5] https://www.rdw.nl/
[6] https://www.tylertech.com/products/socrata
[7] https://data.openparking.nl/downloads/Standard_for_the_Publication_of_Dynamic_Parking_Data_v2.0.pdf
[8] https://dev.socrata.com
[9] https://www.rdw.nl/over-rdw/dienstverlening/open-data/bijsluiter-open-data (translated by Google)

2.3 Data in Amsterdam

This open data portal[10] provides a diverse range of data sources, mainly in the field of geography, but also including information on education, health and politics. The various visualization features are especially notable. However, it is important that the data from different topics can be accessed via an API. This is provided with a number of RESTful interfaces. In addition to the data from the topics mentioned above, separate data sources with a spatial focus exist as well. These are provided in the same way.

2.3.1 Data Format and Structure

The data APIs (not the spatial data) are using JSON by default. The geo web services are using the WMS and WFS standards, which are based on XML and cohere to a predefined vocabulary. Map tiles for a map visualization are provided too.

2.3.2 API

The APIs[11] are grouped sorted according to the following topics:

- Waste management
- Addresses and building information
- Land register
- Cadastre information
- Monuments
- Commercial register
- Parking spaces
- Care services
- Election results
- Milieu topics
- Tellus
- Tram and Metro lines
- Tourism accommodation
- Geographic airport information (Schipol)

2.3.3 Licence

Most data are under a public domain licence available, i.e. CC0. There are some exceptions, which comes with a CC BY 4.0 licence. In this case, a name/credit is required; otherwise, the data can also be used freely.

[10] https://data.amsterdam.nl/
[11] https://api.data.amsterdam.nl/api/

3 Bilbao Data Sources

Bilbao's Council's Website provides the majority of Bilbao's open data. For example, there is data available for electric charging stations, taxi stops, bicycle collection points, traffic status, public transport, and parking.

3.1 Licence

All data is licenced under CC BY 3.0 ES[12].

3.2 Electric Charging Stations

This data source provides the list of electric charging stations available in Bilbao. The information is updated daily.

3.2.1 Data Format and Structure

The information about electric charging stations and the current occupation is provided in two formats: GML and GeoJSON. An example is shown in Figure 3. "NumeroMangeras" indicates the total number of charging points ("Hoses"), whereas "NumeroPlazas" indicates the number of free charging points.


```
"type": "FeatureCollection",
"features": [{
    "type": "Feature",
    "properties": {
        "IdVrtPPElectrolinera": "23",
        "IdPerVertPPElectrolinera": "23",
        "NombreDireccion": "Plaza Ugarteko",
        "TV_NOMBRE": "UGARTEKO (PL)",
        "TDIS_DES_NOMBRE": "1 DEUSTU",
        "TBAR_DES_NOMBRE": "103 IBARREKOLANDA",
        "NumeroPlazas": "2",
        "NumeroMangeras": "3",
        "TipoCarga": "R&#225;pida (43-50 kW)",
        "Origen": "Convenio Ayto-IBERDROLA"
    },
    "geometry": {
        "type": "Point",
        "coordinates": [-2.9547684, 43.2735633]
    }
}, {
```


Figure 3: Bilbao Electric Charging Stations example

3.2.2 API

There are dedicated APIs for the two formats mentioned in the previous section; GML[13] and GeoJSON[14].

[12] https://creativecommons.org/licences/by/3.0/es/
[13] https://www.bilbao.eus/aytoonline/srvDatasetParadas?tipo=electrolineras&formato=gml
[14] https://www.bilbao.eus/aytoonline/srvDatasetParadas?tipo=electrolineras&formato=geojson

3.3 Taxi Stops

This data source provides the list of taxi stops available in Bilbao. The information is updated daily.

3.3.1 Data Format and Structure

The information about taxi stops is provided in two formats, GML and GeoJSON. An example is shown in Figure 4.


```
}, {
    "type": "Feature",
    "properties": {
        "IdParada": "6",
        "CodigoParada": "60202",
        "NombreParada": "Urrutia 1 / Plaza Indautxu",
        "Vigente": "1",
        "FechaAlta": "2019-05-02 00:00:00.0",
        "FechaBaja": "2019-05-02 00:00:00.0",
        "Visible": "1",
        "PlazasTeoricas": "8",
        "Calle": "URRUTIA (CL)",
        "Barrio": "602 INDAUTXU",
        "Distrito": "6 ABANDO"
    },
    "geometry": {
        "type": "Polygon",
        "coordinates": [
            [
                [-2.939938, 43.2603595],
                [-2.9399595, 43.2603732],
                [-2.9400346, 43.2601955],
                [-2.940107, 43.2600177],
                [-2.9400869, 43.2600109],
                [-2.939938, 43.2603595]
            ]
        ]
    }
}, {
```


Figure 4: Bilbao Taxi Stops example

3.3.2 API

There are dedicated APIs for the two formats mentioned in the previous section; GML[15] and GeoJSON[16].

[15] https://www.bilbao.eus/aytoonline/srvDatasetParadas?tipo=taxi&formato=gml
[16] https://www.bilbao.eus/aytoonline/srvDatasetParadas?tipo=taxi&formato=geojson

3.4 Bicycle Collection Points

This data source provides the list of bicycle collection points available in Bilbao and offered by the bicycle loan service, with an indication of occupation. The publisher of the information is Bilbao City Council and the information is updated monthly.

3.4.1 Data Format and Structure

The information about bicycle collection points is provided in WMS[17] and dedicated schema[18]. An example is shown in Figure 5.

```
−<WMS_Capabilities version="1.3.0" updateSequence="35850"
 xsi:schemaLocation="http://www.opengis.net/wms https://www.geobat.eus/geoserver/schemas/wms/1.3.0
 /capabilities_1_3_0.xsd">
  +<Service></Service>
  −<Capability>
    −<Request>
      +<GetCapabilities></GetCapabilities>
      +<GetMap></GetMap>
      +<GetFeatureInfo></GetFeatureInfo>
     </Request>
    +<Exception></Exception>
    +<Layer></Layer>
   </Capability>
 </WMS_Capabilities>
```

Figure 5: Bilbao Bicycle Collection Points example

3.5 Traffic Status

This data source provides information about Bilbao's traffic status in real time. The publisher of the information is Bilbao City Council and the information is updated daily. The following data typology for traffic is provided in this platform:

- bnw:BioRoadNode
- bnw:BioRoadLink
- eti:Activities
- eti:ConstructionWorks
- eti:GeneralObstruction
- eti:MaintenanceWorks

3.5.1 Data Format and Structure

The information about traffic status points is provided in WMS[19], GML[20] and GeoJSON[21] formats. The URL returns the sections of the map with data about intensity, speed, occupation. An example is shown in Figure 6.

[17] https://www.geobilbao.eus/geobilbao/Main?gsservice=Ls&gsrequest=WMS&idservice=MyS_Bicicletas
[18] http://schemas.opengis.net/wms/1.3.0/capabilities_1_3_0.xsd
[19] https://www.geobilbao.eus/geobilbao/Main?gsservice=Ls&gsrequest=WMS&idservice=MyS_GestionTrafico
[20] https://www.bilbao.eus/aytoonline/srvDatasetTrafico?formato=gml
[21] https://www.bilbao.eus/aytoonline/srvDatasetTrafico?formato=geojson

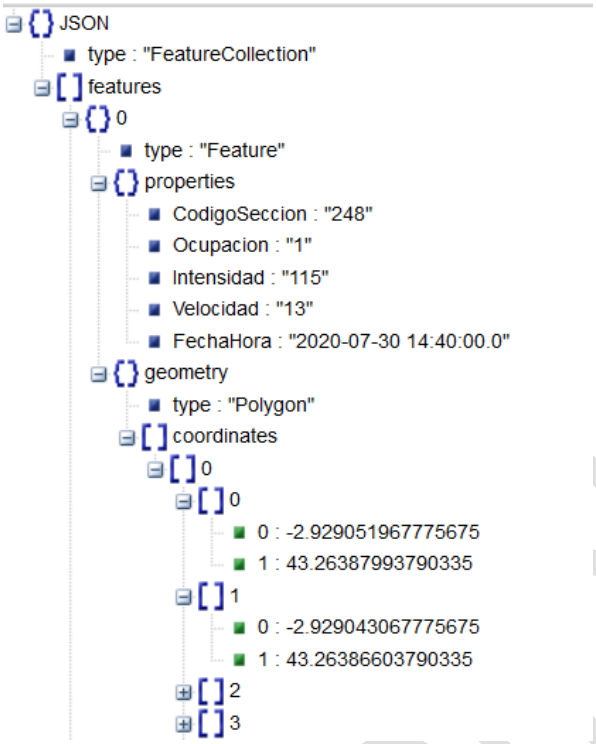

Figure 6: Bilbao Traffic Status example

3.6 Bilbo Bus

This data source provides static and dynamic information about the Bilbobus service, i.e. the public bus transport service in Bilbao.

3.6.1 Data Format and Structure

This data source provides the list of operating Bilbobus lines, divided by day[22]- and night[23]-time. Data typology related to the bus service available in the platform is:

- ept:DatedTimetableVersionFrame:
- ept:EstimatedCall:
- ept:JourneyPattern:
- ept:Line:
- ept:PointInJourneyPattern:
- ept:ProductionTimetable:
- ept:Route:
- ept:Service:
- ept:StopPoint:

The information on bus lines is provided in a CSV file with two columns. The first field represents the code of the bus line and the second field is the description of the line. An example is shown in Figure 7.

[22] http://www.bilbao.eus/aytoonline/jsp/opendata/bilbobus/od_lineas.jsp?idioma=c&formato=csv&tipo=diurnas
[23] http://www.bilbao.eus/aytoonline/jsp/opendata/bilbobus/od_lineas.jsp?idioma=c&formato=csv&tipo=nocturnas

	A	B	C	D
1	Codigo linea	Descripcion linea		
2	1	ARANGOITI - PLAZA BIRIBILA		
3	3	PLAZA BIRIBILA - OTXARKOAGA		
4	10	ELORRIETA - PLAZA BIRIBILA		
5	11	DEUSTU - ATXURI		
6	13	SAN IGNAZIO - TXURDINAGA		
7	18	SAN IGNAZIO - ZORROTZA		
8	22	SARRIKUE - ATXURI		
9	27	ARABELLA - BETOLATZA		
10	28	URIBARRI - ALTAMIRA		
11	30	TXURDINAGA - MIRIBILLA		
12	34	OTXARKOAGA - SANTUTXU		
13	38	OTXARKOAGA - INTERMODAL		
14	40	LA PEÑA - PLAZA BIRIBILA		
15	43	GARAIZAR - SANTUTXU		
16	48	SANTUTXU - LEZEAGA		
17	50	BUIA - LA PEÑA		
18	56	LA PEÑA - JESUSEN BIHOTZA		

Figure 7: Bilbao Bilbobus example

In addition to information on bus lines, this data source also provides the GPS positions of currently operating buses. These positions are approximately refreshed every 30 seconds. The data is provided as WMS[24] and coheres to a fixed XSD schema[25].

3.7 Parking

This data source provides a map of the situation of the public parking lots in Bilbao, in which any citizen or visitor can park, in accordance with current rates. The information is updated daily. The data typology supported by the platform is:

- edi:ParkingPoint:
- eti:CarParkDynamic:

3.7.1 Data Format and Structure

The information is provided in WMS[26] format, and a data schema is available[27].

3.8 Air Quality

This data source[28] provides information about the air quality in Euskadi, from which five measurement stations are located in the city of Bilbao. The information is updated hourly. It provides information about location of the stations and daily and hourly rates for:

- NO (µg/m3)
- NO2 (µg/m3)
- NOX (µg/m3)

[24] https://www.geobilbao.eus/geobilbao/Main?gsservice=Ls&gsrequest=WMS&idservice=MyS_Bilbobus
[25] http://schemas.opengis.net/wms/1.3.0/capabilities_1_3_0.xsd
[26] https://www.geobilbao.eus/geobilbao/Main?gsservice=Ls&gsrequest=WMS&idservice=MyS_AparcamientosRotacion
[27] http://schemas.opengis.net/wms/1.3.0/capabilities_1_3_0.xsd
[28] https://opendata.euskadi.eus/catalogo/-/calidad-aire-en-euskadi-2020/

- O3 (µg/m3)
- O3 8h (µg/m3)
- PM10 (µg/m3)
- PM2,5 (µg/m3)
- SO2 (µg/m3)

3.8.1 Data Format and Structure

Data is provided as XLSX, CSV and JSON. The respective endpoints are listed in a text file[29].

3.8.2 Licence

Licencing information is available online[30], albeit not in English.

[29] https://opendata.euskadi.eus/contenidos/ds_informes_estudios/calidad_aire_2020/es_def/adjuntos/url.txt
[30] https://www.euskadi.eus/informacion/-/informacion-legal/

4 Helsinki Data Sources

This section contains descriptions of data sources from Helsinki.

4.1 Count of Car Traffic and Speeds

This data source provides information about the number of cars and relative speed detected at specific monitor points of roads in Helsinki.

4.1.1 Data Format and Structure

Data is provided in JSON format; the structure depends on the specific APIs that is invoked. Detection points are represented by JSON objects with the following fields:

- id: unique ID of the detection point
- name: title of the detection point
- latitude: latitude of the detection point
- longitude: longitude of the detection point

An example is shown in Figure 8.


```
[{
    "direction": "Itään",
    "data": [{
        "count": 18237,
        "speed": 57,
        "time": "2020-06-08"
    }, {
        "count": 19004,
        "speed": 58,
        "time": "2020-06-09"
    }, {
        "count": 19612,
        "speed": 57,
        "time": "2020-06-10"
    }, {
        "count": 19752,
        "speed": 58,
        "time": "2020-06-11"
    }, {
```


Figure 8: Helsinki Count of Car Traffic and Speeds example

4.1.2 API[31]

Nine different REST APIs are available via HTTP GET requests. The APIs are specified via OpenAPI[32].

- /api/Public/getAllPoints
- /api/Public/getAllDigitrafficPoints
- /api/Public/getDigitrafficPointData/{measuringPeriodId}
- /api/Public/getPointData/{measuringPeriodId}
- /api/Public/getDayById/{measuringPeriodId}
- /api/Public/getHourById/{measuringPeriodId}
- /api/Public/getMonthById/{measuringPeriodId}
- /api/Public/getTotalCount
- /api/Public/getTotalCountDigitraffic

4.1.3 Licence

A licence for the data is not specified.

4.2 Ringroad Vehicle Counts

This data source provides data about the volume of traffic in Helsinki. This data was published for the first time in 2012 and contains information starting from 2010. Types of vehicles are cars, vans, trucks, trucks, buses, motorcycles and trams. The traffic volume is reported hourly, except during peak periods (6.00-9.00 and 15.00-18.00) for which the volume is reported at intervals of 15 minutes.

4.2.1 Data Format and Structure

Data is available in CSV format. The CSV file comprises of the following fields:

- piste: identifier of the calculation point
- nimi: title of the calculation point
- x_gk25: x coordinate of the calculation point; ETRS-GK25 (EPSG 3879) coordinate system.
- y_gk25: y coordinate of the calculation point; ETRS-GK25 (EPSG 3879) coordinate system.
- suunta: direction
- aika: calculation time
- vuosi: year
- ha: number of cars
- pa: number of vans
- ka: number of trucks
- ra: number of lorries
- la: number of buses
- mp: number of motorcycles
- rv: number of trams
- autot: total number of vehicles

An example is shown in Figure 9.

[31] Base URL: https://lamapi.azurewebsites.net
[32] https://lamapi.azurewebsites.net/swagger/index.html

piste	nimi	x_gk25	y_gk25	suunta	aika	vuosi	ha	pa	ka	ra	la	mp	rv	autot
A01	LAUTTASAAREN SILTA	25494426	6672169	1	0	2010	60	4	1	0	4	0	0	69
A01	LAUTTASAAREN SILTA	25494426	6672169	1	100	2010	35	3	1	0	2	0	0	41
A01	LAUTTASAAREN SILTA	25494426	6672169	1	200	2010	17	1	0	0	1	0	0	19
A01	LAUTTASAAREN SILTA	25494426	6672169	1	300	2010	17	1	0	0	1	0	0	19
A01	LAUTTASAAREN SILTA	25494426	6672169	1	400	2010	36	2	0	0	2	0	0	40
A01	LAUTTASAAREN SILTA	25494426	6672169	1	500	2010	71	5	1	0	5	0	0	82
A01	LAUTTASAAREN SILTA	25494426	6672169	1	600	2010	28	3	1	0	5	0	0	37
A01	LAUTTASAAREN SILTA	25494426	6672169	1	615	2010	44	4	3	0	8	0	0	59

Figure 9: Helsinki Ringroad Vehicle Count example

4.2.2 API

The data is provided as a static file; it's available on Helsinki Region Infoshare website[33] and can also be downloaded directly[34].

4.2.3 Licence

The data is licenced under Creative Commons Attribution 4.0.

4.3 Parking Lots

This data source consists of a real-time REST APIs providing information about the parking lots in Helsinki, such as their locations and the number of parking spaces. This information is collected from parking ticket machines and private mobile payment operators operating in Helsinki. The REST interface is documented on the Helsinki Region Infoshare website[35]

4.3.1 Data Format and Structure

Data is provided in JSON format; the structure depends on the specific API that is invoked. For instance, information about parking areas are provided as a GeoJSON collection, and includes the following fields:

- Count: Total number of returned objects
- Next: Next page URL (for pagination of results)
- Previous: Previous page URL (for pagination of results)
- Features: Array of objects (ParkingArea)

An example is shown in Figure 10.

[33] https://hri.fi/data/en_GB/dataset/liikennemaarat-helsingissa
[34] http://www.hel.fi/hel2/tietokeskus/data/helsinki/Liikenne/hki_liikennemaarat.csv
[35] https://hri.fi/data/fi/dataset/rajapinta-helsingin-pysakointipaikkojen-kaytosta

Figure 10: Helsinki Parking Lots example

4.3.2 API[36]

Nine different REST APIs are available via HTTP GET requests. The APIs are specified via OpenAPI[37].

- /parking_area/
- /parking_area/{parking_area_id}/
- /parking_area_statistics/
- /parking_area_statistics/{parking_area_id}/

4.3.3 Licence

A licence for the data is not specified.

4.4 Western Harbour Traffic

This data source consists of a REST API for querying outbound traffic data related to two harbour exits for heavy vehicles and other types of vehicles, respectively. Number of vehicles for the two harbour exits are provided by the API as follow:

- vehicles per minute from the last 15 minutes
- vehicles per hour from the last 24 hours
- vehicles per day from the last 2 weeks

4.4.1 Data Format and Structure

Data is provided in JSON format and includes the following fields:

- timestamp: response timestamp in ISO 8601 format with time offsets from UTC
- results: an array of JSON objects with the following structure:
  - startTime: start time in ISO 8601 format with time offsets from UTC.
  - endTime: end time in ISO 8601 format with time offsets from UTC.,
  - count: number of vehicles
- error: possible error-text

An example is shown in Figure 11.

[36] Base URL: https://pubapi.parkkiopas.fi/public/v1/
[37] https://api.parkkiopas.fi/docs/public/

```
{
  "timestamp": "2020-07-08T17:52:46+03:00",
  "results": [
    {
      "startTime": "2020-06-25T00:00:00+03:00",
      "endTime": "2020-06-26T00:00:00+03:00",
      "count": 1078
    },
    {
      "startTime": "2020-06-26T00:00:00+03:00",
      "endTime": "2020-06-27T00:00:00+03:00",
      "count": 701
    },
    {
      "startTime": "2020-06-27T00:00:00+03:00",
      "endTime": "2020-06-28T00:00:00+03:00",
      "count": 676
    },
    {
      "startTime": "2020-06-28T00:00:00+03:00",
```


Figure 11: Helsinki Western Harbour Transport Data example

4.4.2 API[38]

One REST API is available via HTTP GET requests. The API is specified via OpenAPI[39].

- /v1/calculated_data_results

4.4.3 Licence

A licence for the data is not specified.

[38] Base URL: https://it102.infotripla.fi/ItSensorServiceApi/rest/
[39] https://it102.infotripla.fi/ItSensorServiceApi/swagger/#/

4.5 Bike Traffic

This data source provides information about the number of bikers at a given location and at a given time. This information allows calculating the distribution of bikers within the city. A visualization is also available.

4.5.1 Data Format and Structure

The data itself is divided into two parts. The API provides information about the measuring station, containing longitude and latitude as well as the update interval and further information. The second part provides the counts at each measuring station. JSON is used as exchange format in both cases. An example of both (measuring stations left, count right) is shown in Figure 12.

Figure 12: Helsinki Bike Traffic example

4.5.2 API

The use of the API has to be requested at the provider first. OAuth2 authentication is planned for the future. Documentation[40] and a demo[41] are available for the API.

4.5.3 Licence

The licence terms are unknown at this time. They will be communicated after a request for use.

[40] http://eco-test2.com/apidoc/wso2/apidoc.html
[41] https://apiadmin.eco-counter-tools.com/store/apis/info

4.6 Bicycle Calculations

This data source provides information about the calculated number of bikes in specific road segments of the City of Helsinki from 2009–2019 (where the calculation is made regularly, the results of the most recent calculation year are reported). This data is shown on a map[42].The results of the calculations can be viewed point by point from a pop-up window that opens by clicking on the highlighted road segments.

4.6.1 Data Format and Structure

Calculated data is made available by GeoJSON embedded in a JSON file. The most relevant information is enclosed in an array of JSON objects, named “features”. Each object contains a "properties" field, which provided information like location, calculated bike number and year. The "geometry" field provided information about road segments in the GeoJSON format. The data is provided as a static file[43]. An example is shown in Figure 13.


```
varjson_PP2009_2019_polyline_0={
  "type": "FeatureCollection",
  "crs": {
    "type": "name",
    "properties": {
      "name": "urn:ogc:def:crs:OGC:1.3:CRS84"
    }
  },
  "features": [
    {
      "type": "Feature",
      "properties": {
        "PAIKKA": "Vaskisalmentie",
        "LASKENTA": 663.0,
        "VRK": "818",
        "KESÃ„KUUN_A": "707",
        "HUIPPUVRK": "911",
        "VUOSI": "2019",
        "HUOM": null,
        "NimiÃ¶": "818 \/ 707 \/ 911 \/\n2019"
      },
      "geometry": {
        "type": "LineString",
        "coordinates": [
          [
            24.845921609368002,
            60.165042490322058
          ],
          [
            24.848036937420616,
            60.164704313743101
          ]
        ]
      }
    },
```


Figure 13: Helsinki Bicycle Calculations example

[42] https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/pyoraliikenne/webmap/index.html

[43] https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/pyoraliikenne/webmap/layers/PP2009_2019_polyline_0.js

4.6.2 Licence

The licence terms are unknown at this time. They will be communicated after a request for use.

4.7 Pedestrian Counts

This data source provides information about the calculated number of pedestrians in specific areas of the City of Helsinki from 2000 to 2018. This data is reported on a map[44]. The results of the calculations can be viewed point by point from a pop-up window that opens by clicking on the highlighted road segments.

4.7.1 Data Format and Structure

Calculated data is provided through six different JSON files. Each file follows the same structure and contains a unique JSON object, based on GeoJSON format. The most relevant information is enclosed in the field "features", consisting of an array of JSON objects. Each object contains the field "properties", which provides different information about the location, the calculated number of pedestrians and the year. The field "geometry" includes information about the road segments in the GeoJSON format.

An example is shown in Figure 14.


```
varjson_Alueellisetjalankulkulaskennat2016_0={
  "type": "FeatureCollection",
  "name": "Alueellisetjalankulkulaskennat2016_0",
  "crs": {
    "type": "name",
    "properties": {
      "name": "urn:ogc:def:crs:OGC:1.3:CRS84"
    }
  },
  "features": [
    {
      "type": "Feature",
      "properties": {
        "paikka": "PukinmÃ¤ki",
        "linkki": "<A HREF=\"https:\/\/www.hel.fi\/static\/liitteet\/kaupunk
        "Huom_": "Laskenta-aika klo 6-20"
      },
      "geometry": {
        "type": "Polygon",
        "coordinates": [
          [
            [
              24.993525371159038,
              60.243873269059456
            ],
            [
              24.994335392742517,
              60.244031121825465
            ],
            [
              24.995605278021422,
              60.243486977741817
            ],
            [
              24.99622923596441,
              60.24341081002008
            ],
            [
              24.997586549124421,
              60.244047546504241
            ],
            [
              24.998144834527686,
              60.2439496008883
            ],
```


Figure 14: Helsinki Pedestrian Counts example

[44] https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/Jalankulku/webmap/index.html#12/60.1776/24.9647

4.7.2 API

The data is provided as static files via the following URLs:

- Alueellisetjalankulkulaskennat2016_0: https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/Jalankulku/webmap/data/Alueellisetjalankulkulaskennat2016_0.js
- 2018_5: https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/Jalankulku/webmap/data/2018_5.js
- 2017_4: https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/Jalankulku/webmap/data/2017_4.js
- 2016_3: https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/Jalankulku/webmap/data/2016_3.js
- 20102015_2: https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/Jalankulku/webmap/data/20102015_2.js
- 20002009_1: https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/Jalankulku/webmap/data/20002009_1.js

4.7.3 Licence

The licence terms are unknown at this time. They will be communicated after a request for use.

4.8 Car Traffic and Speeds

This data source provides information about vehicle traffic calculated in specific zones of the City of Helsinki. The traffic is reported for specific intersections and roads. Data is accessible via a map[45] that shows the zones for which traffic calculation is performed.

4.8.1 Data Format and Structure

By clicking points reported on the map, it is possible to access traffic data calculated for the same point. The data is provided as static web pages, which are accessible via the map. The data is represented by the following properties:

- Calculated numbers of cars, vans, trucks, buses, motorcycles, bicycles and trams.
- The direction of the calculated traffic
- The total number of vehicles
- The date of the calculation

An example is shown in Figure 15.

4.8.2 Licence

The licence terms are unknown at this time. They will be communicated after a request for use.

[45] https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/autoliikenne/Liikennelaskennat/Laskennat.html

Kehälaskenta Keskustaan
SUUNTA 1 Tiistaina, 10. syyskuuta 2019
KESKIARKIPÄIVÄLIIKENNE, KÄSILASKENTA SUORITETTU KLO 7 - 19.
Kartta

PISTE=B09 SUUNTA=1 suunta_=Keskustaan PVM=Tiistaina, 10. syyskuuta 2019

PISTE	SUUNTA	Aika2	HA	PA	KA	RA	LA	MP	RV	Autot	M_ajon	Yht	Hay
B09	1	0000	53	6	0	0	6	0	0	65	65	65	71
B09	1	0100	40	3	0	0	3	0	0	46	46	46	49
B09	1	0200	34	4	0	0	4	0	0	42	42	42	46
B09	1	0300	25	3	0	0	3	0	0	31	31	31	34
B09	1	0400	58	6	0	0	8	0	0	72	72	72	80
B09	1	0500	232	9	3	0	28	0	0	272	272	272	303
B09	1	0600	120	11	5	0	10	0	1	146	146	147	164
B09	1	0615	155	21	9	0	13	0	2	198	198	200	226
B09	1	0630	196	26	11	0	17	0	2	250	250	252	284
B09	1	0645	200	26	11	0	18	0	2	255	255	257	290
B09	1	06-07	671	84	36	0	58	0	7	849	849	856	964
B09	1	0700	171	47	9	0	15	0	1	242	242	243	269
B09	1	0715	193	46	10	0	23	3	5	272	275	280	322
B09	1	0730	204	42	13	0	23	7	8	282	289	297	346

Figure 15: Helsinki Car Traffic and Speeds example

4.9 Helsinki Ferry Traffic

This data source provides information about ferry traffic. Departure and arrival times are available as well as the name of the ferry and the name of the agency since 2005. Some keys can contain values for information about passengers or the crew, but these are not always present. The data source is maintained by the Finnish Transport Agency (FTA).

4.9.1 Data format and structure

The data format is JSON and can be requested via a REST[46] interface. The API provides data almost in real-time, but no standard is used to structure the data. An example is shown in Figure 16.

4.9.2 Licence

The data is licenced under Creative Commons 4.0.

[46] https://meri.digitraffic.fi/api/v1/port-calls

```
"portCallId": 2587077,
"portCallTimestamp": "2020-08-03T06:47:56.000+0000",
"customsReference": "14/80024148",
"portToVisit": "FIHEL",
"prevPort": "RUULU",
"nextPort": "FIHEL",
"domesticTrafficArrival": false,
"domesticTrafficDeparture": false,
"arrivalWithCargo": true,
"notLoading": true,
"discharge": 1,
"shipMasterArrival": "X",
"shipMasterDeparture": "X",
"managementNameArrival": " ",
"managementNameDeparture": " ",
"forwarderNameArrival": ".",
"forwarderNameDeparture": ".",
"freeTextArrival": " ",
"freeTextDeparture": "Seuraava satama ei vielä tiedossa.",
"vesselName": "Kallio",
"vesselNamePrefix": "ms",
"radioCallSign": "OJQJ",
"radioCallSignType": "real",
"imoLloyds": 9319064,
"mmsi": 230645000,
"nationality": "FI",
"vesselTypeCode": 70,
"certificateIssuer": "Traficom",
"certificateStartDate": "2019-05-05T21:00:00.000+0000",
"certificateEndDate": "2024-05-11T21:00:00.000+0000",
"currentSecurityLevel": 1,
"agentInfo": [
  {
    "role": 1,
    "portCallDirection": "Arrival or whole PortCall",
    "ediNumber": "003714714063",
    "name": "GAC Finland Oy"
  },
  {
    "role": 2,
    "portCallDirection": "Arrival or whole PortCall",
    "ediNumber": null,
    "name": "ESL Shipping Oy"
  }
],
```


Figure 16: Helsinki Ferry Traffic example

5 Messina Data Sources

This section contains descriptions of data sources from Messina.

5.1 Maps

The Map server[47] is a service that offers map layers that show toponomy, communal boundaries, streets, green areas, rivers and streams, stations, and galleries, among others.

5.1.1 Data Format and Structure

The data is presented in two different ways, JSON[48] and SOAP[49]. The latter provides the data as ARCGIS, a geographic information system (GIS) produced by Esri [1]. It is used for creating and using maps, compiling geographic data; map analysis, sharing geographic information and managing geographic information in a database. The data can be viewed in a web interface.

An example is shown in Figure 17.

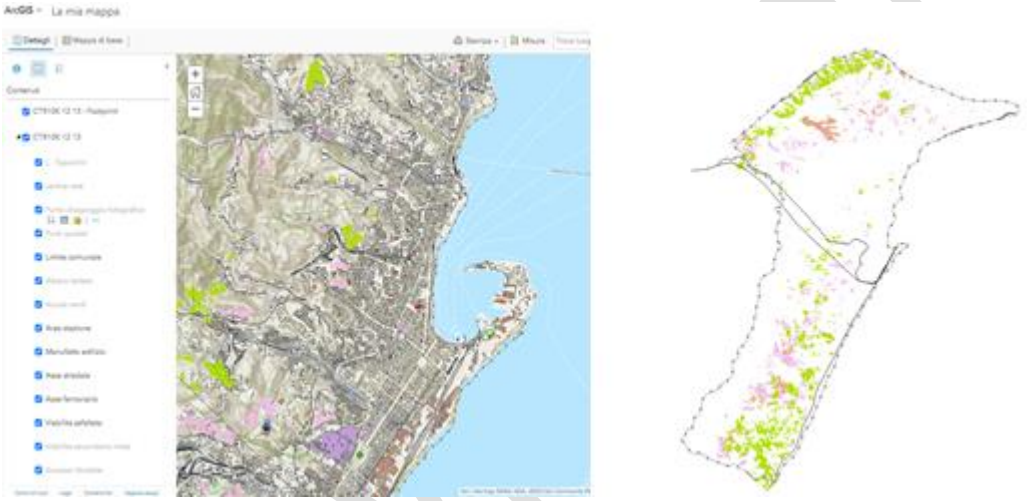

Figure 17: Messina Maps example

5.1.2 API

The supported operations are listed below:

- Export Map
- Identify
- Find
- Return Updates
- Generate KML

5.1.3 Licence

The data are defined as open, but lack a specific licence.

[47] http://2.237.156.214:6080/arcgis/rest/services/CTR10K_12_13/MapServer
[48] http://2.237.156.214:6080/arcgis/rest/services/CTR10K_12_13/MapServer?f=pjson
[49] http://2.237.156.214:6080/arcgis/services/CTR10K_12_13/MapServer?wsdl

5.2 Zafferia Weather

This data source[50] provides data from a weather station located in Zafferia, hamlet of Messina Municipality.

5.2.1 Data Format and Structure

Data are presented in two ways, table or visual graphics. The following values are measured:

- Temperature
- moisture
- wind
- daily wind rose/weekly wind rose/annual wind rose
- daily rainfall/monthly rainfall/annual rainfall
- hourly rainfall -> real time update
- barometer -> real time update

5.2.2 API

The API provides two different types of reports:

- Climatological Summaries: monthly and annual report
- Browse Archive Records: day records

5.2.3 Licence

The data are defined as open, but lack a specific licence.

5.3 Noise Monitoring

This data source[51] provides data about the following two types of noise:

- Electromagnetic pollution monitoring network
- Vehicle traffic noise monitoring network

5.3.1 Data Format and Structure

Data are presented as monthly reports in PDF format. The data are updated in the following intervals:

- Electromagnetic:
  - daytime (08:00 / 23:00) [V/m]
  - nighttime (23:00 / 08:00) [V/m]
  - hourly trend of a general day of the month
- Vehicular noise:
  - daily [dBA]
  - weekly [dBA]
  - monthly [dBA]

Data regarding electromagnetic pollution are collected in the following seven locations:

[50] http://www.messinameteo.it/index.php?option=com_wrapper&view=wrapper&Itemid=274
[51] https://onlinepa.info/index.php?page=moduli&mod=6&ente=88&node=265

- Scuola media G. Mazzini
- Scuola elementare Montepiselli
- Scuola elementare G. Mazzini
- Scuola elementare F. Crispi
- Scuola elementare Beata Eustochia
- Casa di cura San Camillo
- Istituto Don Orione

Data regarding vehicular flow are collected in the following six locations:

- Viale Giostra
- Viale Europa
- Via Celi
- Via Taormina
- Via La Farina
- Viale della LIbertà
- Viale Boccetta

5.3.2 Licence

The data are defined as open, but lack a specific licence.

6 Other Data Sources

This section contains descriptions of data sources that do not stem from a particular pilot city, but are relevant regardless of location.

6.1 OpenWeatherMap

OpenWeatherMap[52] provides numerous types of weather data, like temperature, precipitation, and humidity. The weather data can, among others, be limited by city, geographical coordinates or specification of area. In addition to current forecasts, historical weather data is available.

6.1.1 Data Format and Structure

Most API methods serve JSON. When accessing bulk data, single files are compressed to a gzip archive. For requests involving multiple files tarballs are used in conjunction with gzip. Depending on the type of data requested formats like XML, CSV, and HTML are available. Additionally, OpenWeatherMap offers access to layered maps via API, where each layer refers to a certain weather metric. Units and language can be toggled via the query parameter. An example[53] of the data returned by the current weather API is shown in Figure 18. Note that only actually measured values are returned by the API. For instance, if no precipitation was measured, the according “rain” and “snow” attributes will be missing from the payload.

- Location coordinates
- Weather description
- Wind (speed, direction)
- Temperature (actual, felt, min, max)
- Humidity
- Atmospheric pressure
- Cloudiness
- Visibility
- Rain and snow volume
- Time stamp

An example is shown in Figure 18.

[52] https://openweathermap.org/
[53] https://openweathermap.org/current

```
{
  "coord": {
    "lon": -0.13,
    "lat": 51.51
  },
  "weather": [{
      "id": 300,
      "main": "Drizzle",
      "description": "light intensity drizzle",
      "icon": "09d"
    }],
  "base": "stations",
  "main": {
    "temp": 280.32,
    "pressure": 1012,
    "humidity": 81,
    "temp_min": 279.15,
    "temp_max": 281.15
  },
  "visibility": 10000,
  "wind": {
    "speed": 4.1,
    "deg": 80
  },
  "clouds": {
    "all": 90
  },
  "dt": 1485789600,
  "sys": {
    "type": 1,
    "id": 5091,
    "message": 0.0103,
    "country": "GB",
    "sunrise": 1485762037,
    "sunset": 1485794875
  },
  "id": 2643743,
  "name": "London",
  "cod": 200
}
```


Figure 18: OpenWeatherMap example

6.1.2 Licence

Basic weather data is available for free, with an allowed 60 calls per hour up to 1M calls per month. The available APIs include current weather, minute/hourly/daily forecasts, 5-day historical data, basic weather maps as well as weather triggers and widgets. Paid plans increase API quota as well as the granularity of the weather data. Certain types of data, like historical forecasts further in the past, are only available with the paid plans. All data, regardless of pricing, is licenced under ODbL[54].

[54] https://opendatacommons.org/licences/odbl/

6.2 OpenAQ

Motivated by preventing premature death caused by air pollution OpenAQ[55] offers open data from the air quality domain. The provided datasets are harvested from official authorities portals and can either be viewed directly on the web or accessed via API.

6.2.1 Data Format and Structure

The data can be accessed in multiple ways, namely manual download from the website or accessed via API. The former serves the data as CSV files, the latter as JSON. The API has multiple endpoints for retrieving different kinds of data, for example the available countries and cities, or a list of the primary data sources. However, the API only provides data for the last 90 days. Older data can be obtained from an Amazon AWS endpoint. Instead of using the API, snapshots from the underlying PostgreSQL database can also be downloaded.

Measurements of the following pollutants are provided:

- Nitrogen Dioxide (NO_2)
- Sulfur Dioxide (SO_2)
- Carbon Monoxide (CO)
- Ozone (O_3)
- Black Carbon (BC)
- Particles (>2.5µm and >10µm)

An example is shown in Figure 19.


```
{
  "parameter": "Ammonia",
  "date": {
      "utc": "2015-07-16T20:30:00.000Z",
      "local": "2015-07-16T18:30:00.000-02:00"
  },
  "value": "72.9",
  "unit": "µg/m3",
  "location": "Anand Vihar",
  "country": "IN",
  "city": "Delhi",
  "sourceName": "Anand Vihar",
  "averagingPeriod": {
     "value": 1,
     "unit": "hours"
  },
  "coordinates": {
     "latitude": 43.34,
     "longitude": 23.04
  }
}
```


Figure 19: OpenAQ example

[55] https://openaq.org/

6.2.2 Licence

All data is licenced under CC-BY 4.0[56].

6.3 Open Street Map

Created in 2004, OpenStreetMap[57] (OSM) is a collaborative service offering free editable maps of the world. Crowdsourced data is available under the terms of Open Database Licence and mainly consists of geodata. Offered data can be used for different purposes.

6.3.1 Data Format and Structure

APIs directly provided by OSM allow read and write operations on the raw map data contained in the OpenStreetMap database. These APIs are mainly used for editing applications, or to create new tools. OSM also provides a dedicated "Overpass" API[58] for querying by specific search criteria (e.g. location, type, tag, etc.) or combinations of them. The available formats are CSV, JSON, and GeoJSON.

The main elements of OSM data formats are:

- Tags: This element provides semantic information; tags consist of a key and a value. OSM does not specify any tag; tags are chosen spontaneously by contributors.
- Nodes, Ways, Relations: These are the three types of objects of OSM. Objects can be associated with an arbitrary number of tags and must have an ID.
    - Nodes are coordinates associated with an ID and tags; they can represent a point, or a small object. Only nodes have coordinates.
    - Ways are sequences of nodes. Also, ways are associated with an ID and tags. Thereby ways get a geometry by using the coordinates of the referenced nodes. This also creates a topology: two ways are connected if both points are positioned to the same node.
    - Relations are groups of elements (e.g. nodes, ways) and define logical or geographic relationships between other elements. Relations are used to represent different concepts, such as borders of countries, counties, routes, etc. Also, relations are associated with an ID and tags.
- Areas: These elements consist of enclosing ways or relations.
- Meta Data: These elements consist of tags used for meta-data purposes (e.g. to display information in map editors)

An example is shown in Figure 20.

[56] https://creativecommons.org/licences/by/4.0/
[57] https://www.openstreetmap.org/
[58] https://wiki.openstreetmap.org/wiki/Overpass_API

```
<?xml version="1.0" encoding="UTF-8"?>
<osm version="0.6" generator="Overpass API">
  <node id="691566183" lat="49.7982193" lon="13.4686623">
    <tag k="ele" v="295"/>
    <tag k="information" v="guidepost"/>
    <tag k="name" v="Dolanský most (pravý břeh)"/>
    <tag k="tourism" v="information"/>
  </node>
  <area id="3600435511"/>
  <area id="3600442466"/>
  <node id="691566191" lat="49.8003120" lon="13.4679726">
    <tag k="ele" v="295"/>
    <tag k="information" v="guidepost"/>
    <tag k="name" v="Dolanský most (levý břeh)"/>
    <tag k="tourism" v="information"/>
  </node>
  <area id="3600435511"/>
  <area id="3600442466"/>
[...]
  <area id="3600435511">
    <tag k="admin_level" v="4"/>
    <tag k="boundary" v="administrative"/>
    <tag k="name" v="Jihozápad"/>
    <tag k="name:cs" v="Jihozápad"/>
    <tag k="name:de" v="Südwesten"/>
    <tag k="population" v="1209298"/>
    <tag k="ref" v="CZ03"/>
    <tag k="ref:NUTS" v="CZ03"/>
    <tag k="source" v="cuzk:ruian"/>
    <tag k="source:population" v="csu:uir-zsj"/>
    <tag k="type" v="boundary"/>
    <tag k="wikipedia" v="cs:NUTS Jihozápad"/>
  </area>
  <area id="3600442466">
    <tag k="ISO3166-2" v="CZ-PL"/>
    <tag k="admin_level" v="6"/>
    <tag k="boundary" v="administrative"/>
    <tag k="name" v="Plzeňský kraj"/>
    <tag k="name:cs" v="Plzeňský kraj"/>
    <tag k="name:de" v="Region Pilsen"/>
    <tag k="name:ru" v="Пльзенский край"/>
    <tag k="population" v="572687"/>
    <tag k="ref" v="PL"/>
    <tag k="ref:NUTS" v="CZ032"/>
    <tag k="source" v="cuzk:ruian"/>
    <tag k="source:population" v="csu:uir-zsj"/>
    <tag k="type" v="boundary"/>
    <tag k="wikipedia" v="cs:Plzeňský kraj"/>
  </area>
```

Figure 20: Open Street Map example

6.3.2 Licence

All data is licenced under Open Database Licence (ODbL) v1.0

6.4 CityBikes

This project provides bike sharing data. Currently, more than 400 cities are included. Out of the four URBANITE pilots, data is available for Bilbao and Helsinki.

6.4.1 Data Format and Structure

Data is returned in the JSON format. The payload contains relevant information like available stations, their geographical coordinates, and the number of currently free bikes, among others.

An example is shown in Figure 21.


```
{
  "networks": [
    {
        "company": "JCDecaux",
        "href": "/v2/networks/velib",
        "location": {
          "latitude": 48.856612,
          "city": "Paris",
          "longitude": 2.352233,
          "country": "FRA"
        },
        "name": "Vélib'",
        "id": "velib"
    },
    {...}
  ]
}
```


Figure 21: CityBikes example

6.4.2 API[59]

The API provides two endpoints. One allows retrieval of a list of networks that share bikes. For each network, a second endpoint then provides a list of stations for a given network. To reduce traffic, the AI provides whitelist-based filtering, i.e. an array of the desired fields can be specified using the "fields" query parameter. The endpoints are listed below:

- `/networks`
- `/networks/{network_id}`

6.4.3 Licence

CityBikes is free to use. However, developers should indicate that CityBikes data has been used.

[59] https://api.citybik.es/v2/

6.5 Kaggle

The *Kaggle* platform focuses on the analysis of data. It offers open data as well as tools and applications for analysis directly in the provided web application (like Python notebooks). In addition, there are competitions that anyone can take part in. The goal is to discover new knowledge from the provided data. This way, the community is always involved and motivated. The data stems from different domains, with the bulk belonging to the traffic, mobility, and urbanity domains.

6.5.1 Data Format and Structure

The data[60] is mostly open and always available in an open format (no proprietary formats are desired here). Tabular data is often available as CSV or SQLite DB. Document-oriented content is usually provided in JSON format. Image data for computer vision is also available.

An example of CSV data is shown in Figure 22.

Country	Year	No. of cases	No. of deaths	No. of cases_me...
Country	Year	No. of cases	No. of deaths	No. of cases median
107 unique values	856 total values	0 16% 3 1% Other (714) 83%	0 30% 1 2% Other (584) 68%	0 62.0m
Afghanistan	2017	630308[495000-801000]	298[110-510]	630308
Algeria	2017	0	0	0
Angola	2017	4615605[3106000-6661000]	13316[9970-16600]	4615605
Argentina	2017	0	0	0
Armenia	2017	0	0	0
Azerbaijan	2017	0	0	0
Bangladesh	2017	32924[30000-36000]	76[3-130]	32924

Figure 22: Kaggle example

6.5.2 API

There is a public API[61] available. However, it requires authentication via key, which is available after free registration.

6.5.3 Licence

Different licences are used depending on the dataset. The description of the provider should be read completely. There are some datasets without a licence, so the use of this kind of datasets should be taken with caution. However, most of the data are under a public domain licence and can be used for free.

[60] https://www.kaggle.com/docs/datasets
[61] https://github.com/Kaggle/kaggle-api

7 Conclusion

This document describes a number of data sources, both pilot-specific and globally applicable. Out of the many data sources that were initially gathered, only a select few could be included in this document. This is due to various reasons, such as lack of documentation or inaccessible links at time of writing. However, the deliverable provides a first overview of the domains from which data is available, as well as the most commonly used formats. Especially traffic related data seems to be readily available in comparably high quality and granularity. It will have to be seen which of these data sources are ultimately fit for use with the recommendation engine. Of course, data may also be transformed to align their structure to be compatible with the recommendation engine.

8 References

[1] ESRI, “ESRI” [Online]. Available: https://www.esri.com/

9 APPENDIX A: Data Sources Overview

Table 1: Overview of Data Sources

Pilot	Title	Formats	URLs	Licence
Amsterdam	Amsterdam Maps	CSV, GeoJSON, GIS	https://maps.amsterdam.nl https://maps.amsterdam.nl/open_geodata	https://www.ogc.org/standards/wkt-crs
Amsterdam	Netherlands RDW	CSV, SPDP	https://www.rdw.nl/	CC0
Amsterdam	Data in Amsterdam	JSON, WMS, WFS	https://data.amsterdam.nl/	CC0, CC-BY-4.0
Bilbao	Electric Charging Stations	GML, GeoJSON	https://www.bilbao.eus/aytoonline/srvDatasetParadas?tipo=electrolineras&formato=gml https://www.bilbao.eus/aytoonline/srvDatasetParadas?tipo=electrolineras&formato=geojson	CC BY 3.0 ES
Bilbao	Taxi Stops	GML, GeoJSON	https://www.bilbao.eus/aytoonline/srvDatasetParadas?tipo=taxi&formato=gml https://www.bilbao.eus/aytoonline/srvDatasetParadas?tipo=taxi&formato=geojson	CC BY 3.0 ES
Bilbao	Bicycle Collection Points	WMS	https://www.geobilbao.eus/geobilbao/Main?gsservice=Ls&gsrequest=WMS&idservice=MyS_Bicicletas	CC BY 3.0 ES
Bilbao	Traffic Status	GML, GeoJSON, WMS	https://www.geobilbao.eus/geobilbao/Main?gsservice=Ls&gsrequest=WMS&idservice=MyS_GestionTrafico https://www.bilbao.eus/aytoonline/srvDatasetTrafico?formato=gml https://www.bilbao.eus/aytoonline/srvDatasetTrafico?formato=geojson	CC BY 3.0 ES

Pilot	Title	Formats	URLs	Licence
Bilbao	Bilbobus	CSV, WMS	http://www.bilbao.eus/aytoonline/jsp/opendata/bilbobus/od_lineas.jsp?idioma=c&formato=csv&tipo=diurnas http://www.bilbao.eus/aytoonline/jsp/opendata/bilbobus/od_lineas.jsp?idioma=c&formato=csv&tipo=nocturnas	CC BY 3.0 ES
Bilbao	Parking	WMS	https://www.geobilbao.eus/geobilbao/Main?gsservice=Ls&gsrequest=WMS&idservice=MyS_AparcamientosRotacion	CC BY 3.0 ES
Bilbao	Air Quality	CSV, JSON, XLSX	https://opendata.euskadi.eus/catalogo/-/calidad-aire-en-euskadi-2020/	CC BY 3.0 ES
Helsinki	Count of Car Traffic and Speeds	JSON	https://lamapi.azurewebsites.net/swagger/index.html	*Not specified*
Helsinki	Ringroad Vehicle Counts	CSV	https://hri.fi/data/en_GB/dataset/liikennemaarat-helsingissa http://www.hel.fi/hel2/tietokeskus/data/helsinki/Liikenne/hki_liikennemaarat.csv	Creative Commons Attribution 4.0
Helsinki	Parking Lots	GeoJSON	https://pubapi.parkkiopas.fi/public/v1/ https://api.parkkiopas.fi/docs/public/	*Not specified*
Helsinki	Western Harbour Transport	JSON	https://it102.infotripla.fi/ItSensorServiceApi/rest/ https://it102.infotripla.fi/ItSensorServiceApi/swagger/#/	*Not specified*
Helsinki	Bike Traffic	JSON	http://eco-test2.com/apidoc/wso2/apidoc.html https://apiadmin.eco-counter-tools.com/store/apis/info	*Not specified*
Helsinki	Bicycle Calculations	JSON, GeoJSON	https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/pyoraliikenne/webmap/index.html https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/pyoraliikenne/webmap/layers/PP2009_2019_polyline_0.js	*Not specified*
Helsinki	Pedestrian Counts	JSON, GeoJSON	https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/Jalankulku/webmap/index.html #12/60.1776/24.9647	*Not specified*
Helsinki	Car Traffic and Speeds	Online Map	https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/kadut/liikennetilastot/autoliikenne/Liikennelaskennat/Laskennat.html	*Not specified*
Helsinki	Ferry Traffic	JSON	https://meri.digitraffic.fi/api/v1/port-calls	Creative Commons 4.0

Pilot	Title	Formats	URLs	Licence
Messina	Maps	JSON, GIS	http://2.237.156.214:6080/arcgis/rest/services/CTR10K_12_13/MapServer http://2.237.156.214:6080/arcgis/rest/services/CTR10K_12_13/MapServer?f=pjson http://2.237.156.214:6080/arcgis/services/CTR10K_12_13/MapServer?wsdl	Open, but not specified
Messina	Zafferia Weather	Online Table	http://www.messinameteo.it/index.php?option=com_wrapper&view=wrapper&Itemid=274	Open, but not specified
Messina	Noise Monitoring	PDF	https://onlinepa.info/index.php?page=moduli&mod=6&ente=88&node=265	Open, but not specified
Other	OpenWeatherMap	JSON, XML, CSV, HTML	https://openweathermap.org/	ODbL
Other	OpenAQ	CSV, JSON	https://openaq.org/	CC-BY-4.0
Other	Open Street Map	CSV, JSON, GeoJSON	https://www.openstreetmap.org https://wiki.openstreetmap.org/wiki/Overpass_API	ODbL
Other	CityBikes	JSON	https://api.citybik.es/v2/	Open, but not specified
Other	Kaggle	CSV, SQLite, JSON	https://www.kaggle.com/docs/datasets https://github.com/Kaggle/kaggle-api	Varying, mostly public domain