



i-locate - Indoor/outdoor LOcation and Asset management Through open gEodata (GA 621040)

COVER PAGE



DELIVERABLE

Project Acronym: i-locate

Grant Agreement number: 621040

Project Title: Indoor/outdoor LOcation and Asset management Through open gEodata

D2.2 – Analysis of Open Data Repositories

Revision: 1.9

Authors: Lucian Brancovean (INDSOFT), Stefano Piffer (TRILOGIS), Mircea Greavu (INDSOFT), Catalin Popa (INDSOFT), Ramona Falamas (INDSOFT), Stefano Piffer (TRILOGIS), Giuseppe Conti (TRILOGIS)

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	X
C	Confidential, only for members of the consortium and the Commission Services	

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 1 of 39	Open Data Repositories

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision	Date	Author	Organisation	Description
v1.0	21/07/2014	Lucian Brancovean	INDSOFT	Structure of the document
v1.1	24/07/2014	Stefano Piffer	TRILOGIS	Geospatial database
v1.2	30/07/2014	Lucian Brancovean, Mircea Greavu	INDSOFT	Open Repositories Details
v1.8	20/07/2014	Lucian Brancovean, Catalin Popa, Ramona Falamas	INDSOFT	Open Repositories Details
v1.9	22/07/2014	Giuseppe Conti	TRILOGIS	Overall Review
V2.0	20/01/2015	Irene Facchin	TRILOGIS	Minor Typo corrections

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

1 List of references

Number	Full reference
1	i-locate deliverable D2.1 – Data Survey Report. Available online from: http://www.i-locate.eu/public-deliverables/
2	Creative Commons Public Domain License https://creativecommons.org/publicdomain/zero/1.0/legalcode
3	Creative Commons Attribution (CC BY v3.0) License https://creativecommons.org/licenses/by/3.0/legalcode
4	GTFS format https://developers.google.com/transit/gtfs/
5	Ogr2ogr tool http://www.gdal.org/ogr2ogr.html
6	GDAL License http://opensource.org/licenses/mit-license.php

2 Table of Acronyms

Acronym	Description
CSV	Comma-Separated Values
EPSG	European Petroleum Survey Group
GDAL	Geospatial Data Abstraction Library
GeoJSON	Open standard format for encoding collections of simple geographical features using JSON
GML	Geography Markup Language
GTFS	General Transit Feed Specification Reference
ID	Identifier
IODL	Italian Open Data License
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
KMZ	Zipped KML files with a .kmz extension
ODS	Operational Data Store
OGC	Open Geospatial Consortium

PDF	Portable Document Format
RDF	Resource Description Framework
SHP	ESRI Shapefile
SRID	ID of the reference system
TSV	Tab-Separated Values
URL	Uniform Resource Locator
WFS	Web Feature Service
WMS	Web Mapping Service
XLS	Microsoft Excel file format
XML	Extensible Markup Language
X/MIT	A specific type of open source license

3 Table of content

1	LIST OF REFERENCES	3
2	TABLE OF ACRONYMS	3
3	TABLE OF CONTENT	5
4	INTRODUCTION	7
5	GEOSPATIAL DATABASE	8
6	TOOLS	11
6.1	Conversion tool ogr2ogr	11
7	OPEN DATA REPOSITORIES	12
7.1	MH-OPN-02: Greece – Public Wi-Fi Hot-Spots	12
7.1.1	Technical capabilities	12
7.1.2	Integration solution	12
7.1.3	Regulatory and privacy concerns	13
7.2	MH-OPN-03: Greece – Public Buildings	14
7.2.1	Technical capabilities	14
7.2.2	Integration solution	14
7.2.3	Regulatory and privacy concerns	14
7.3	UMC-OPN-02 – Utrecht – Building DataBase	16
7.3.1	Technical capabilities	16
7.3.2	Integration solution	16
7.3.3	Regulatory and privacy concerns	16
7.4	RO-OPN-02 - Rovereto – roads	18
7.4.1	Technical capabilities	18
7.4.2	Integration solution	18
7.4.3	Regulatory and privacy concerns	18
7.5	RO-OPN-03 - Rovereto – Impaired people facilities	20
7.5.1	Technical capabilities	20
7.5.2	Integration solution	20
7.5.3	Regulatory and privacy concerns	20
7.6	RO-OPN-04 - Rovereto – Transportation facilities	22
7.6.1	Technical capabilities	22
7.6.2	Integration solution	22
7.6.3	Regulatory and privacy concerns	23
7.7	VE-OPN-02 - Province of Rome– Road Map	24
7.7.1	Technical capabilities	24
7.7.2	Integration solution	24
7.7.3	Regulatory and privacy concerns	24
7.8	VE-OPN-03 - Province of Rome– Viability	25
7.8.1	Technical capabilities	25
7.8.2	Integration solution	25
7.8.3	Regulatory and privacy concerns	26
7.9	VE-OPN-04 - Wi-Fi map of province of Rome	27

7.9.1	Technical capabilities	27
7.9.2	Integration solution	27
7.9.3	Regulatory and privacy concerns	28
7.10	RJK-OPN-02 - Rijeka – City Streets	29
7.10.1	Technical capabilities	29
7.10.2	Integration solution	29
7.10.3	Regulatory and privacy concerns	29
7.11	ITA-OPN-01 - Italy – Public Health Agencies	30
7.11.1	Technical capabilities	30
7.11.2	Integration solution	30
7.11.3	Regulatory and privacy concerns	30
7.12	ITA-OPN-02 - Italy – Automatic pharmaceutical vending machines	32
7.12.1	Technical capabilities	32
7.12.2	Integration solution	32
7.12.3	Regulatory and privacy concerns	33
7.13	ITA-OPN-03 - Italy – Public para-pharmacies	34
7.13.1	Technical capabilities	34
7.13.2	Integration solution	34
7.13.3	Regulatory and privacy concerns	35
7.14	ITA-OPN-04 - Italy – Public pharmacies	36
7.14.1	Technical capabilities	36
7.14.2	Integration solution	36
7.14.3	Regulatory and privacy concerns	37
7.15	SJH-OPN-02 - Malta – Facilities	38
7.15.1	Technical capabilities	38
7.15.2	Integration solution	38
7.15.3	Regulatory and privacy concerns	39



4 Introduction

This document lists the Open Data Repositories identified within D.2.1 (Data Survey Report) as a good fit for the i-locate project and for each of them it provides a technical description of the necessary procedures (in terms of data transformation, format translation etc.) required to make them usable by i-locate, together with any “connector” (i.e. software module) required to integrate it with the project.

It should be noted that all the entries identified in the previous data survey (D.2.1) that refer to Open Street Map data are not listed here, since integration with OSM is already detailed in the System Architecture (D.1.4), and will be an integral part of the system and portal. This will ensure full integration with the repositories made available by Open Street Map and therefore no further translation or transformation will be required for those data (hence they have not been introduced within this document).

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 7 of 39	Open Data Repositories

5 Geospatial database

In general terms, i-locate will follow a federated approach. In other words, wherever technically possible, data from (external) open repositories will be directly read from those online repositories (to avoid data duplication) and automatically converted through ad-hoc software “connectors” as described below, and then finally stored in a geospatial database. The structure of this database is described below.

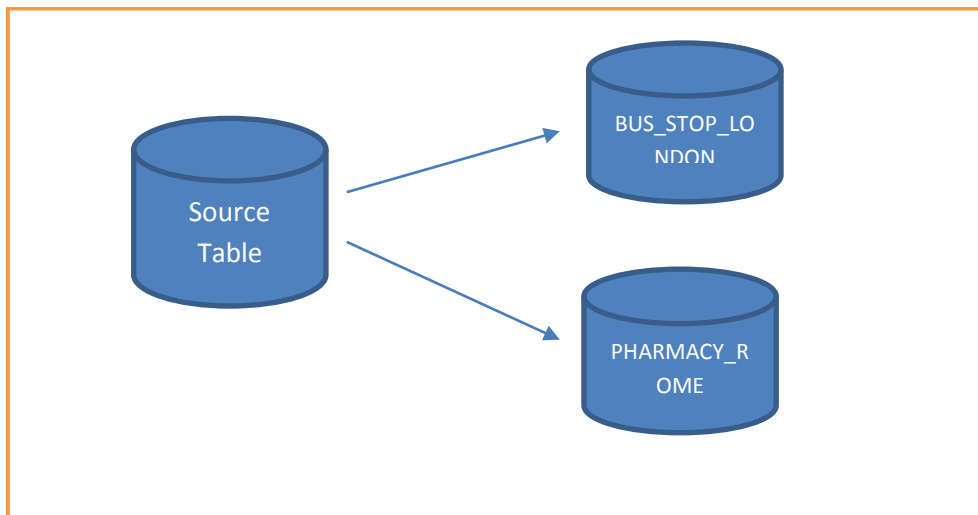


Figure 1: the overall approach followed for the structure of the database.

More specifically, the source table, meant as a catalogue of sources, has been structured with the following attributes:

- ID: a unique identifier.
- Short Name, the name used to refer to the resource.
- Description, a brief textual explanation of the resource.
- URL, the address the resources was retrieved from.
- BBox, the bounding box (the box surrounding the full extent) of the dataset.
- SRID, the reference system used by the dataset (using the EPSG code).
- DataType, the type of resource (e.g. POINT, MULTIPOINT, etc.).
- RefDataTable, the table referred to.

For each open data repository, an entry in this table will exist, pointing to the table that contains the data imported from the given external resource, as in the following example in Table 1.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 8 of 39	Open Data Repositories

Table 1: an example of data describing two datasets.

ID	Short Name	Description	URL	BBox	SRID	Data Type	RefDataTable
1	BUS STOP	Bus stop position in London	http://...	(min – max) Lat (min – max) Lon	EPSG: 4326	POINT	BUS_STOP_LONDON
2	PHARMACY	Pharmacy position in Rome	http://...	(min – max) Lat (min – max) Lon	EPSG: 3857	MULTIPOINT	PHARMACY_ROME

For each data table the following mandatory fields have been identified:

- ID, the unique identifier.
- SourceID, the reference to the source table.
- Geom, the geospatial feature.
- MainAttribute, a “blob” of attributes in JSON/XML format
- MinorAttributes, a “blob” of attributes in JSON/XML format.

Furthermore, an additional set of optional fields can be provided for each data table. It is worth noting that these fields can be different for any different table:

- License
- CreationDate,
- OnlineInformation

Continuing with the example in Table 1, the following table shows an example for a table called for is provided below for the two tables referred to within it, respectively called “BUS_STOP_LONDON” and “PHARMACY_ROME”.

Table 2: possible example for the “BUS_STOP_LONDON” data table.

ID	SourceID	Geom	MainAttribute	MinorAttributes
101	1	POINT(2.0,4.5)	Bus Stop A2	Lines: A2,A5,L6; Days:Monday, Friday; Frequency: 2h
102	1	POINT(2.2,5.1)	Bus Stop F34	Lines: F34,G4; Days:All days; Frequency: 0.5h...
....

Table 3: possible example for the "PHARMACY_ROME" data table.

ID	SourceID	Geom	MainAttribute	MinorAttributes
5	2	MULTIPOINTS((2.0,4.5), (7.2,4.4) ,(2.6,8.5),....)	AAA Pharmacies	Always Open
6	2	MULTIPOINTS((2.1,1.5), (7.4,2.4) ,(4.6,5.5),....)	ABB Pharmacies	Always Close during week end
....

6 Tools

6.1 Conversion tool ogr2ogr

The automatic conversion routines will be implemented through the ogr2ogr software. This tool is part of GDAL (Geospatial Data Abstraction Library) which is a translator library for raster and vector geospatial data formats that is released under an X/MIT style Open Source license by the Open Source Geospatial Foundation. The tool is used to convert features data between to file formats. In i-locate it has been used for importing data from different file formats directly into PostgreSQL and, in this process, to set the correct coordinate system.

The following parameters are passed as arguments to the tool:

- **-f** (format_name): in this case -f = "PostgreSQL", because the tool is used to import data directly into the database.
- **-append**: it is used when data had to be added to an existing data.
- **-overwrite**: it is used when old data has to be replaced with new data.
- **-skipfailures**: this option is used when some data entries in the input files are erroneous and therefore it is necessary for the program to continue importing data after a failure, skipping features not complying with requirements.

A non-exhaustive list of supported input formats include: ESRI Shapefile, CSV, KML, and PostgreSQL. Some of the supported output file formats are: ESRI Shapefile, PostgreSQL.

7 Open Data Repositories

This section details, for each of the repositories identified within D.2.1 (Data Survey Report), the technical solution identify to ensure integration within i-locate repository.

7.1 MH-OPN-02: Greece – Public Wi-Fi Hot-Spots

7.1.1 Technical capabilities

Short description	Map of Greece's Wi-Fi public access points
URI / File name	http://geodata.gov.gr/geodata/index.php?option=com_sobi2&sobi2Task=sobi2Details&catid=17&sobi2Id=98&Itemid=
Format(s) available (or analogue)	SHP, GML, KML
Owner / source	geodata.gov.gr
Copyright / license	Creative Commons Attribution (CC BY v3.0)
Geographical theme	Public Wi-Fi map
Geographical extent / coverage	Nation wide - Medium level of detail
Other relevant spatial data (e.g. EPSG code etc.)	WGS84 (16.76871,30.655429,33.188516,43.339883)
Publishing date	2010
Frequency of Update	Yearly
No. of (logically) different datasets	1
Data quality	4

7.1.2 Integration solution

Specifically, the data within the file available from download, contains a description of the Wi-Fi location, their location expressed by geographic coordinates (latitude and longitude). Additionally, the file contains information regarding the municipality to which each hotspot belongs to as well as the IP address of the Wi-Fi point.

The ogr2ogr tool can be used to import the SHP file to PostGIS/PostgreSQL. Appropriate tables will be created in the database corresponding to the shape file data. A GeoServer instance can be configured to use PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 12 of 39	Open Data Repositories



The connector component will download the SHP files from the source data and then call the ogr2ogr tool to import the data into the database.

7.1.3 Regulatory and privacy concerns

The data is published under a Creative Commons Attribution License. As an overview, this means: “You are free to: **share** - copy and redistribute the material in any medium or format; **adapt** - remix, transform, and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms. Under the following terms: **attribution** - you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.”

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 13 of 39	Open Data Repositories

7.2 MH-OPN-03: Greece – Public Buildings

7.2.1 Technical capabilities

Short description	Map of Greece's public buildings
URI / File name	http://geodata.gov.gr/geodata/index.php?option=com_sobi2&sobi2Task=sobi2Details&catid=17&sobi2Id=29&Itemid=
Format(s) available (or analogue)	SHP, CSV, GML, KML, XLS, ODS
Owner / source	Geodata.gov.gr
Copyright / licence	Creative Commons (CC BY v3.0)
Geographical theme	Public buildings map
Geographical extent / coverage	Nation wide - High level of detail
Other relevant spatial data (e.g. EPSG code etc.)	WGS84 (16.76871,30.655429,33.188516,43.339883)
Publishing date	2010
Frequency of update	Yearly
No. of (logically) different datasets	1
Data quality	4

7.2.2 Integration solution

The file contains information about the buildings including the region they lie in, the county they lie in, their size, type, address, number, to which municipality they belong to and the location of each building given through their geographic coordinates (latitude and longitude).

The ogr2ogr tool can be used to import the SHP file to PostGIS/PostgreSQL. Appropriate tables will be created in the database corresponding to the shape file data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

The connector component will download the SHP files from the source and then call the ogr2ogr tool to import the data into the database.

7.2.3 Regulatory and privacy concerns

The data is published under a Creative Commons Attribution License. As an overview, this means: "You are free to: **share** - copy and redistribute the material in any medium or format; **adapt** - remix,

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 14 of 39	Open Data Repositories



transform, and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms. Under the following terms: **attribution** - you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.”

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 15 of 39	Open Data Repositories

7.3 UMC-OPN-02 – Utrecht – Building DataBase

7.3.1 Technical capabilities

Short description	Roads of Utrecht
URI / File name	http://www.nationaalgeoregister.nl/geonetwork/srv/dut/search# 3d2825a2-269a-486b-933b-4726b124ad9b
Format(s) available (or analogue)	XML
Owner / source	NGR Nationaal Georegister
Copyright / licence	Creative Commons CC0 License
Geographical theme	Roads
Geographical extent / coverage	City - High level of detail
Other relevant spatial data (e.g. EPSG code etc.)	N/A
Publishing date	Unknown
Frequency of update	Yearly
No. of (logically) different datasets	1
Data quality	5

7.3.2 Integration solution

The file contains geographic points that constitute roads. The resulting roads that are represented also include lane separation.

The ogr2ogr tool can be used to import the XML file, more specifically encoded as GML file, to PostGIS/PostgreSQL. Appropriate tables will be created in the database corresponding to the XML data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

The connector component will download the XML files from the source location and then call the ogr2ogr tool to import the data into the database.

7.3.3 Regulatory and privacy concerns

The data is published under a Creative Commons CC0 License. As an overview, this means: “The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 16 of 39	Open Data Repositories



neighboring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.”

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 17 of 39	Open Data Repositories

7.4 RO-OPN-02 - Rovereto – roads

7.4.1 Technical capabilities

Short description	Roads of Rovereto
URI / File name	http://www.dati.gov.it/catalog/dataset/pat_comune-di-rovereto-strade
Format(s) available (or analogue)	SHP
Owner / source	dati.gov.it
Copyright / licence	Creative Commons CC0 License
Geographical theme	Roads
Geographical extent / coverage	City –High level of detail
Other relevant spatial data (e.g. EPSG code etc.)	UTM32 - EUREF89
Publishing date	2013
Frequency of update	Not available
No. of (logically) different datasets	1
Data quality	5

7.4.2 Integration solution

The specific dataset contains several parameters regarding the streets in the city including: codes, geographic coordinates, type of the street (whether it is a regular street, a square, an alley, an avenue, a path etc.) together with the name of the streets.

The ogr2ogr tool can be used to import the SHP file to PostGIS/PostgreSQL. Appropriate tables will be created in the database corresponding to the shape file data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

The connector component will download the SHP files from the source location and then call the ogr2ogr tool to import the data into the database.

7.4.3 Regulatory and privacy concerns

The data is published under a Creative Commons CC0 License. As an overview, this means: “The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 18 of 39	Open Data Repositories



neighboring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.”

7.5

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 19 of 39	Open Data Repositories

7.6 RO-OPN-03 - Rovereto – Impaired people facilities

7.6.1 Technical capabilities

Short description	Impaired people facilities of Rovereto
URI / File name	http://www.contextaware.provincia.tn.it/PortaleCxA/csv?comune=Rovereto
Format(s) available (or analogue)	CSV
Owner / source	dati.gov.it
Copyright / licence	Creative Commons Attribution
Geographical theme	Denomination of the service, address, website, geographical position, altitude, different categories, accessibility code, description of the service, opening hours and other values.
Geographical extent / coverage	City – High level of detail
Other relevant spatial data (e.g. EPSG code etc.)	N/A
Publishing date	2013
Frequency of update	Not available
No. of (logically) different datasets	1
Data quality	5

7.6.2 Integration solution

The data contains the name of the service, street address, zip code, municipality, province, country to which it belongs to, the Internet site, altitude, some accessibility codes and some of their properties (description of the service), opening hours, last modified date and geographic coordinates.

A software component will be developed to download the CSV files from the site, parse it and then import the resulting data into PostgreSQL. Appropriate tables will be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

7.6.3 Regulatory and privacy concerns

The data is published under a Creative Commons Attribution License. As an overview, this means: “You are free to: **share** - copy and redistribute the material in any medium or format; **adapt** - remix, transform, and build upon the material for any purpose, even commercially. The licensor cannot

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 20 of 39	Open Data Repositories



revoke these freedoms as long as you follow the license terms. Under the following terms:
attribution - you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.”

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 21 of 39	Open Data Repositories

7.7 RO-OPN-04 - Rovereto – Transportation facilities

7.7.1 Technical capabilities

Short description	Transport facilities of Rovereto
URI / File name	http://www.dati.gov.it/catalog/dataset/pat_trasporti-pubblici-del-trentino-formato-gtfs
Format(s) available (or analogue)	CSV
Owner / source	dati.gov.it
Copyright / licence	Creative Commons Attribution
Geographical theme	The data includes the urban public transport data and the suburban public transport data. For both are provided 8 files: <ul style="list-style-type: none"> • Agency.txt: providers of the service; • Stops.txt: stops data; • Routes.txt: transport lines; • Trips.txt: bus rides data; • Stop_times.txt: timetable for each bus stop; • Calendar.txt and Calendar_date.txt: frequencies of bus rides;
Geographical extent / coverage	City – High level of detail
Other relevant spatial data (e.g. EPSG code etc.)	N/A
Publishing date	2013
Frequency of update	Weekly
No. of (logically) different datasets	2
Data quality	5

7.7.2 Integration solution

The data provided comes as a ZIP package. Each file contains transit information like the provider of this data (the agency), timetable, stops and routes as GTFS format. More information about the GTFS format are available online from: <https://developers.google.com/transit/gtfs/>.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 22 of 39	Open Data Repositories



A software component will be developed in order to automatically download the CSV files from the site, parse it and then import the data into PostgreSQL. Appropriate tables will be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS. The PDF which can be found at the above URI explains the above CSV files.

7.7.3 Regulatory and privacy concerns

The data is published under a Creative Commons Attribution License. As an overview, this means: “You are free to: **share** - copy and redistribute the material in any medium or format; **adapt** - remix, transform, and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms. Under the following terms: **attribution** - you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.”

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 23 of 39	Open Data Repositories

7.8 VE-OPN-02 - Province of Rome– Road Map

7.8.1 Technical capabilities

Short description	Roads Map of Rome Province
URI / File name	http://www.dati.gov.it/catalog/dataset/provincia-di-roma_strade-di-propriet-dell-amministrazione-provinciale
Format(s) available (or analogue)	KMZ, XML, CSV, TSV
Owner / source	dati.gov.it
Copyright / licence	Creative Commons CC0 Public Domain
Geographical theme	Roads
Geographical extent / coverage	Province – High level of detail
Other relevant spatial data (e.g. EPSG code etc.)	WGS84
Publishing date	2012
Frequency of update	Weekly
No. of (logically) different datasets	1
Data quality	5

7.8.2 Integration solution

The data available for download contains street codes, names, categories. The ogr2ogr tool can be used to import the KML file to PostGIS/PostgreSQL. Appropriate tables will be created in the database corresponding to the KML data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

The connector component will download the KMZ file from the source address, unzip it to extract the corresponding KML file and then call the ogr2ogr tool to import it into the database.

7.8.3 Regulatory and privacy concerns

The data is published under a Creative Commons CC0 License. As an overview, this means: “The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.”

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 24 of 39	Open Data Repositories

7.9 VE-OPN-03 - Province of Rome– Viability

7.9.1 Technical capabilities

Short description	Viability of Rome Province
URI / File name	http://www.dati.gov.it/catalog/dataset/provincia-di-roma_monitoraggio-traffico-stradale
Format(s) available (or analogue)	XML, CSV, TSV
Owner / source	dati.gov.it
Copyright / licence	Creative Commons CC0 Public Domain
Geographical theme	Viability of roads of Rome's province. For each monitored road, the following is provided: name of the road, coordinates of the road, timestamp of measurement, average speed and number of vehicles per hour.
Geographical extent / coverage	Province – High level of detail
Other relevant spatial data (e.g. EPSG code etc.)	Name of the roads, coordinates of the road
Publishing date	2013
Frequency of update	Every 10 minutes
No. of (logically) different datasets	1
Data quality	5

7.9.2 Integration solution

A software component could be developed to download the CSV files from the site, parse it and then import the data into PostgreSQL. Appropriate tables could be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

However, after a first test, it has been discovered that the files downloaded from the above link are empty. More specifically, they contain only column names like the name of the street, coordinates of the streets, date and time of the collection, the flow of the vehicles. Without proper data the connector for this repository it will not be possible to properly tested its functioning.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 25 of 39	Open Data Repositories



7.9.3 Regulatory and privacy concerns

The data is published under a Creative Commons CC0 License. As an overview, this means: “The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.”

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 26 of 39	Open Data Repositories

7.10 VE-OPN-04 - Wi-Fi map of province of Rome

7.10.1 Technical capabilities

Short description	Wi-Fi map of province of Rome
URI / File name	http://www.dati.gov.it/catalog/dataset/provincia-di-roma_provinciawifi
Format(s) available (or analogue)	XML, CSV, TSV, JSON, GeoJSON
Owner / source	dati.gov.it
Copyright / licence	Creative Commons CC0 Public Domain
Geographical theme	Every access point has a name, the position in decimal degrees (SRID 4326), the address and the type of access point. There are two types: public access points and private ones shared with the public administration.
Geographical extent / coverage	Province – High level of detail
Other relevant spatial data (e.g. EPSG code etc.)	Position of Wi-Fi hotspot in decimal degrees (SRID 4326)
Publishing date	2013
Frequency of update	Daily
No. of (logically) different datasets	1
Data quality	5

7.10.2 Integration solution

All the files that can be downloaded from the above URI contain information related to the Wi-Fi hotspots in the province of Rome. The data available contain the name of the access points, their geographic coordinates, street address, the municipality where they belong to and the type (if public or private). It should be noted that some of the entries do not contain information within some columns.

A software component will be developed to download the CSV and JSON files from the site, parse them and then import the data into PostgreSQL. Appropriate tables will be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 27 of 39	Open Data Repositories



7.10.3 Regulatory and privacy concerns

The data is published under a Creative Commons CC0 License. As an overview, this means: “The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.”

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 28 of 39	Open Data Repositories

7.11 RJK-OPN-02 - Rijeka – City Streets

7.11.1 Technical capabilities

Short description	City Streets of Rijeka
URI / File name	http://ulice-kbr.rijeka.hr/Default.aspx
Format(s) available (or analogue)	PDF
Owner / source	Municipality of Rijeka
Copyright / licence	Public availability
Geographical theme	Street addresses
Geographical extent / coverage	City – Very low level of detail
Other relevant spatial data (e.g. EPSG code etc.)	None
Publishing date	Not available
Frequency of update	Monthly
No. of (logically) different datasets	1
Data quality	3

7.11.2 Integration solution

The data is only available in PDF format, or through an on-line map viewer. This format is not suitable for data import. As a consequence, this data is not well suited for the i-locate database. As alternative option, its use as low-quality “plain” rendering background (made accessible as WMS) will be considered.

7.11.3 Regulatory and privacy concerns

The licensing and privacy information is detailed from the website of the city administration at the address: <http://www.rijeka.hr/UvjetiKoristenja> .

7.12 ITA-OPN-01 - Italy – Public Health Agencies

7.12.1 Technical capabilities

Short description	Public Health Agencies
URI / File name	http://www.dati.salute.gov.it/dati/dettaglioDataset.jsp?menu=dati&idPag=2
Format(s) available (or analogue)	CSV
Owner / source	Italian Health Authority
Copyright / licence	Italian Open Data Licence v2.0
Geographical theme	Location of public health agencies (“ASL”)
Geographical extent / coverage	Nation wide - Average level of detail
Other relevant spatial data (e.g. EPSG code etc.)	Location provided as “street address” (regional code, name of the region, address, zip code, municipality, province)
Publishing date	2012
Frequency of update	Monthly
No. of (logically) different datasets	1
Data quality	4

7.12.2 Integration solution

The data provided comes as a plain CSV file. Each row includes the year, the ID of the region and its name (E.g. Piedmont), the code of the health agency, the name of the health agency, and its full address.

A software component will be developed in order to automatically download the CSV files from the site, parse it and then import the data into PostgreSQL. In case of the address, since no geographical position is available, the process will have to invoke a geocoding service (to retrieve the geographical position corresponding to the address). Appropriate tables will be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS. The PDF which can be found at the above URI explains the above CSV files.

7.12.3 Regulatory and privacy concerns

The data is published under the IODL – Italian Open Data License. As an overview, this means: the user “can reproduce, publicly distribute, lease, present and demonstrate in public, communicate to the public, make available to the public including, transmit and retransmit in any

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 30 of 39	Open Data Repositories



way, include in collective works and / or composite publish, extract and reuse the information [...]” provided that the user “1) indicates the source of the information and the name of Licensor, including, if possible, a copy of this License or a shortcut (link) to it: 2) does not reuse the information in a way that suggests that they have an official character or that Licensor approves the use you make of the Information”.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 31 of 39	Open Data Repositories

7.13 ITA-OPN-02 - Italy – Automatic pharmaceutical vending machines

7.13.1 Technical capabilities

Short description	Automatic pharmaceutical vending machines
URI / File name	http://www.dati.salute.gov.it/dati/dettaglioDataset.jsp?menu=dati&idPag=4
Format(s) available (or analogue)	CSV
Owner / source	Italian Health Authority
Copyright / licence	Italian Open Data Licence v2.0
Geographical theme	Location of pharmaceutical vending machines; the dataset contains the list of medicines vending machines. It provides: <ul style="list-style-type: none"> • identification code of the site (UUID format); • name of the site; • information about the legal entity; • type of vending machine and description; • geolocation of vending machine
Geographical extent / coverage	Nation wide - Average level of detail
Other relevant spatial data (e.g. EPSG code etc.)	Location provided as “street address” and WSG84
Publishing date	2012
Frequency of update	Weekly
No. of (logically) different datasets	1
Data quality	4

7.13.2 Integration solution

The data provided comes as a plain CSV file. Each row includes the ID of the vending machine, the name of the place of the company where the machine is located and its full address, the VAT number of the company, the city (including name of the city and unique ID), validity date range, type of vending machine, role of the company holding the vending machine (e.g. leaser), position in latitude and longitude.

A software component will be developed in order to automatically download the CSV files from the site, parse it and then import the data into PostgreSQL. Appropriate tables will be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS. The PDF which can be found at the above URI explains the above CSV files.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 32 of 39	Open Data Repositories

7.13.3 Regulatory and privacy concerns

The data is published under the IODL – Italian Open Data License. As an overview, this means: the user “can reproduce, publicly distribute, lease, present and demonstrate in public, communicate to the public, make available to the public including, transmit and retransmit in any way, include in collective works and / or composite publish, extract and reuse the information [...]” provided that the user “1) indicates the source of the information and the name of Licensor, including, if possible, a copy of this License or a shortcut (link) to it: 2) does not reuse the information in a way that suggests that they have an official character or that Licensor approves the use you make of the Information”.

7.14 ITA-OPN-03 - Italy – Public para-pharmacies

7.14.1 Technical capabilities

Short description	Public para-pharmacies
URI / File name	http://www.dati.salute.gov.it/dati/dettaglioDataset.jsp?menu=dati&idPag=7
Format(s) available (or analogue)	CSV
Owner / source	Italian Health Authority
Copyright / licence	Italian Open Data Licence v2.0
Geographical theme	Location of public para- pharmacies; the dataset provides: <ul style="list-style-type: none"> • identification code (UUID); • name; • complete address; • start selling date; • geolocation of shop.
Geographical extent / coverage	Nation wide - Average level of detail
Other relevant spatial data (e.g. EPSG code etc.)	Location provided as “street address” and WSG84
Publishing date	2012
Frequency of update	Weekly
No. of (logically) different datasets	1
Data quality	4

7.14.2 Integration solution

The data provided comes as a plain CSV file with each line including: the unique ID, the name of the shop, the full address, its VAT number, other information regarding the code of the city where it is located, starting date when the was licensed and eventually the position in latitude and longitude.

A software component will be developed in order to automatically download the CSV files from the site, parse it and then import the data into PostgreSQL. Appropriate tables will be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS. The PDF which can be found at the above URI explains the above CSV files.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 34 of 39	Open Data Repositories

7.14.3 Regulatory and privacy concerns

The data is published under the IODL – Italian Open Data License. As an overview, this means: the user “can reproduce, publicly distribute, lease, present and demonstrate in public, communicate to the public, make available to the public including, transmit and retransmit in any way, include in collective works and / or composite publish, extract and reuse the information [...]” provided that the user “1) indicates the source of the information and the name of Licensor, including, if possible, a copy of this License or a shortcut (link) to it: 2) does not reuse the information in a way that suggests that they have an official character or that Licensor approves the use you make of the Information”.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 35 of 39	Open Data Repositories

7.15 ITA-OPN-04 - Italy – Public pharmacies

7.15.1 Technical capabilities

Short description	Public pharmacies
URI / File name	www.dati.salute.gov.it/dataset/farmacie.jsp
Format(s) available (or analogue)	CSV
Owner / source	Italian Health Authority
Copyright / licence	Italian Open Data Licence v2.0
Geographical theme	Location of public pharmacies; the dataset provides: <ul style="list-style-type: none"> • identification code (UUID); • name; • complete address; • start selling date; • geolocation of shop.
Geographical extent / coverage	Nation wide - Average level of detail
Other relevant spatial data (e.g. EPSG code etc.)	Location provided as “street address” and WSG84
Publishing date	2012
Frequency of update	Weekly
No. of (logically) different datasets	1
Data quality	4

7.15.2 Integration solution

The data provided comes as a plain CSV file with each line including: the unique ID, the name of the shop, the corresponding health agency ID, the full address, its VAT number, other information regarding the code of the city where it is located, starting date when the was licensed and eventually the position in latitude and longitude.

A software component will be developed in order to automatically download the CSV files from the site, parse it and then import the data into PostgreSQL. Appropriate tables will be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS. The PDF which can be found at the above URI explains the above CSV files.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 36 of 39	Open Data Repositories

7.15.3 Regulatory and privacy concerns

The data is published under the IODL – Italian Open Data License. As an overview, this means: the user “can reproduce, publicly distribute, lease, present and demonstrate in public, communicate to the public, make available to the public including, transmit and retransmit in any way, include in collective works and / or composite publish, extract and reuse the information [...]” provided that the user “1) indicates the source of the information and the name of Licensor, including, if possible, a copy of this License or a shortcut (link) to it: 2) does not reuse the information in a way that suggests that they have an official character or that Licensor approves the use you make of the Information”.

7.16 SJH-OPN-02 - Malta – Facilities

7.16.1 Technical capabilities

Short description	Facilities of Malta cities
URI / File name	http://opendatamalta.com/
Format(s) available (or analogue)	CSV, XML, XLS, RDF, JSON
Owner / source	Open Data Initiative
Copyright / licence	Creative Commons Attribution
Geographical theme	Accommodation, Bus Routes, Cruises, Events, Historical Places, Hospitals, Hotels, Pharmacies, Restaurants, Shopping Malls, Streets
Geographical extent / coverage	Nation wide – Average level of detail
Other relevant spatial data (e.g. EPSG code etc.)	Street Addresses and WGS84
Publishing date	2014
Frequency of update	Monthly
No. of (logically) different datasets	11
Data quality	5

7.16.2 Integration solution

A software component will be developed to download the CSV and JSON files from the site, parse them and then import the data into PostgreSQL. Appropriate tables will be created in the database corresponding to the parsed data. A GeoServer instance can be configured to have PostgreSQL as a data source, thus exposing the data through standard interfaces such as OGC WMS and OGC WFS.

The available datasets are:

- Tourists expenditures: it does not contain any geographic information so it does not present any interest for the project.
- Hotels in Malta: The CSV file contains latitude and longitude information. A CSV parser will be implemented. The data contains information about the hotels: geographic coordinates, hotel ranking, their name, their address, the towns they are situated in, their fax and telephone number, e-mail and website.
- Accommodation in Malta: it does not have geographic information so it does not present any interest for the project.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 38 of 39	Open Data Repositories

- Restaurants in Malta: the JSON file contains geographic information therefore a JSON parser will be implemented. The data contains information about the restaurants: their name, address, geographic coordinates, telephone number, website, spoken language and opening hours.
- Events in Malta: it does not have geographic information so it does not present any interest for the project.
- Heritage Sites: the CSV file contains geographic information so a CSV parser will be implemented. The data contains information about the heritage sites: description, type (museum, garden, temples, church etc.), period, location, geographic coordinates, opening days, the number of the bus that can be took, contact, telephone number, e-mail address, spoken language.
- Maltese Bays: the CSV file contains geographic information so a CSV parser will be implemented. The data contains information about bays: name, geographic coordinates and spoken language.

7.16.3 Regulatory and privacy concerns

The data is retrieved from the opendatamalta.com site. Regarding intellectual property and privacy, this is the information on the site: “Open data is a world-wide policy making Public Sector Information (PSI), or in other words public data and information, available and exploitable in order to ensure transparency with citizens. Open data is the idea that certain data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control.”

The data is published under a Creative Commons Attribution License. As an overview, this means: You are free to: **share** - copy and redistribute the material in any medium or format; **adapt** - remix, transform, and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms. Under the following terms: **attribution** - you must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

File: D.2.2 - Analysis of open data repositories	D2.2
Page: 39 of 39	Open Data Repositories