



Spatial Opendata

Technical Information Note 03/2018

May 2018

Contents

Executive summary	2
Open data	2
Geo-spatial	2
Uses of open-data	2
Data types	3
Projections	5
Licences	5
Software	6
Data sources	7
References & further reading	8

1. Introduction

In recent years government and other organisations around the world have made an increasing number of datasets freely and openly available. It is estimated that 80% of all data is spatial and it is no surprise that many of the open datasets are in a spatial format, that can be used in a Geographic Information System (GIS), web application or other system. This Technical Information Note (TIN) on Spatial open-data will enable landscape professionals to understand:

- what open spatial data is
- the various licences that open-data is released under that describe how it can be used
- the most commonly used data formats
- where to find the most useful and relevant datasets for landscape professionals

An accompanying spreadsheet catalogues of some of the most useful open spatial data resources for LI members and can be found on the LI website

Whilst every effort has been made to ensure this TIN is as up-to-date as possible, the pace of change in open data community is rapid. If you are aware of additional resources that would be useful to landscape professionals please contact the spatial opendata group via technical@landscapeinstitute.org

2. Geo-spatial

Wikipedia, one of the largest open-source projects, defines a geographic information system (GIS) as one that is “designed to capture, store, manipulate, analyse, manage, and present spatial or geographic data”. The range of data that can be used in a GIS is as broad as the user base across Business, Environment, Defence, Health, Agriculture, Meteorology, Sustainable Development and many more. Likewise, a broad range of spatial data is now openly available.

There are some key frameworks which enable the sharing of open-spatial data. First and foremost, the Open Geospatial Consortium (OGC) <http://www.opengeospatial.org/>, which is an international not for profit organization, sets open standards for the global geospatial community. These standards are “made through a consensus process and are freely available for anyone to use to improve sharing of the world's geospatial data”.

In the EU, the INSPIRE Directive (2007/2/EC) <http://inspire.ec.europa.eu/about-inspire/563> is the principal regulatory framework that enables the sharing of environmental spatial information for policy-making and public access to spatial information. The associated UK Regulations <https://data.gov.uk/location> place legal obligations on public authorities which hold geographical data about (things that affect the) environment. Many of these datasets are highly relevant to landscape professionals as they cover “species and habitats, population demographics, utility services, energy resources, marine features, metrological features and human health”.

3. Uses of open-data

Given the wide range of spatial datasets available, the potential for innovation is wide. Whether your interest is Green Infrastructure Planning, Neighbourhood Planning, Landscape Character

Assessment/Evaluation, Management planning or Interpretation/Education there are datasets out there that you can use. For inspiration here are some examples:

- The Soils Site Reporter <https://data.gov.uk/apps/cranfield-soil-site-reporter> is an “easy-to-use, online soil reporting tool which produces site-specific soils information with maps and soil descriptions”.
- Walkonomics <https://data.gov.uk/apps/walkonomics-find-walkable-route> analyses over 1.7 million street trees and thousands of urban parks to find beautiful routes through tree-filled streets and parks.
- NiceAreas <https://data.gov.uk/apps/niceareas-search-any-area-based-upon-criteria-you-set-including-crime-prices-travel-links> allows users to search for any area based upon criteria including crime, prices, travel links and schools .
- Ukactive <https://theodi.org/blog/why-dataenabled-insight-is-the-key-to-an-active-nation> explains how the benefits of open data can help tackle the global physical inactivity crisis
- GLA population projections <https://data.london.gov.uk/case-studies/population-projections/> is an interactive tool to view London’s population projections.
- Breathe Heathrow: <https://theodi.org/summer-showcase-breathe-heathrow> uses air quality sensors to help residents understand how the airport affects their area, bringing more data into the hands of communities, to address local needs.

4. Data types

Most GIS systems can open, visualise and process many data types. This following list is not exhaustive but instead focuses on those most commonly found on open-data portals.

CSV

A comma-separated values (CSV) file “stores tabular data (numbers and text) in plain text. Each line of the file is a data record. Each record consists of one or more fields, separated by commas” https://en.wikipedia.org/wiki/Comma-separated_values CSV files can be visualised in a GIS if there are columns with geographic co-ordinates, like Eastings/ Nothings or Latitude/ Longitude. If the CSV contains geographic identifiers, like postcodes, the information can be ‘geo-coded’, to append co-ordinates from another source.

GeoJSON

GeoJSON is a geospatial data interchange format based on JavaScript Object Notation (JSON). It “defines several types of JSON objects and the manner in which they are combined to represent data about geographic features, their properties, and their spatial extents” <https://tools.ietf.org/html/rfc7946> . The GeoJSON format “differs from other [GIS](#) standards in that it was written and is maintained not by a formal standards organization, but by an Internet working group of developers” <https://en.wikipedia.org/wiki/GeoJSON>. It is no surprise that it is used in many web-applications as well as GIS.

GeoTIFF

TIFF is a very commonly used raster format and GeoTIFF is a metadata standard which allows geo-referencing information (like map projection, coordinate system, ellipsoid, datum) to be embedded within a TIFF file http://www.gdal.org/frmt_gtiff.html. This additional information includes everything necessary to establish the exact spatial reference for the file. It means a raster file can be displayed in GIS without the need for a separate file containing the projection information.

GML

The Geography Markup Language (GML) is an XML grammar for expressing geographical features. It “serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the Internet”. As with most XML based grammars, there are “two parts to the grammar – the schema that describes the document and the instance document that contains the actual data”. A GML document is described using a GML Schema which “allows users and developers to describe generic geographic data sets that contain points, lines and polygons” <http://www.opengeospatial.org/standards/gml>

KML

Google created Keyhole Markup Language (KML) which is a “XML language focused on geographic visualization, including annotation of maps and images” <http://www.opengeospatial.org/standards/kml>. It is the key data format use to visualise data in Google Earth but can be used in other GIS systems as well.

TAB

The MapInfo TAB file format https://en.wikipedia.org/wiki/MapInfo_TAB_format is a vector data format that was developed as a proprietary format by MapInfo (now Pitney Bowes). Each TAB comprises several files:

- .TAB (links between all other files and holds information about the type of data)
- .DAT (stores the attribute data)
- .IND (optional index file for tabular data, present if any columns are indexed)
- .MAP (stores the graphic and geographic information needed to display each vector feature on a map)
- .ID (stores information linking graphic data to the database information).

SHP

The ESRI Shapefile was developed by ESRI <https://www.esri.com> and, like the TAB format, is made up of several constituent parts, which are outlined below and described in detail here: <https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>

- **.shp** - shape file, contains the geometry
- **.shx** - shape index file, an index of the geometry

- **.dbf** - attribute file, contains attributes for the features
- **.prj** - a plain text file that contains the projection and coordinate system
- **.sbn** and **.sbx** - spatial index of the features
- **.shp.xml** - XML metadata

WFS

A Web Feature Service (WFS) was defined and is maintained by the Open Geospatial Consortium. It “provides an interface allowing requests for geographical features across the web using platform-independent calls” https://en.wikipedia.org/wiki/Web_Feature_Service. Effectively it provides data that that end-users can edit or spatially analyze . It contrasts with a Web mapping service (WMS) interface that only returns an image (like a raster tile). Further information on WFS here: <http://www.opengeospatial.org/standards/wfs#news>

WMS

Web Map Service (WMS) Interface Standard provides “a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases” <http://www.opengeospatial.org/standards/wms>. A WMS request defines the geographic layer(s) and area of interest and the response is one or more geo-registered map images (returned as JPEG, PNG, etc) that can be displayed in a browser application.

5. Projections

Any map representing part of the earth’s surface, which is a spheroid, has to make adjustments to represent a curved surface on a flat plane. The mathematical adjustments made to do this, and to add an x/y reference system, is a map projection. Each projection has an European Petroleum Survey Group (EPSG) code as well as their name. There are thousands of projections which approximate the spheroid in different ways. They are often designed to be especially accurate for a part of the world (e.g. country). In the UK the 2 projections that are most often used are:

- British National Grid (EPSG:27700)
- World Geodetic System 84, (EPSG:4326)

Ordnance Survey data and most terrestrial data use ESPG: 27700 while marine data and data originating from GPS is often encoded as WGS 84, which is a global standard. It is important to check which projection data is in before you start using it. You may need to re-project the data to a standard projection relevant to your project area. More information here: <http://epsg.io/about>

6. Licences

In Europe, you can have have copyright over “works (content) that you create and which are original to you” and a “database right over collections of data that you have put a substantial effort into obtaining, verifying or presenting” <https://theodi.org/guides/publishers-guide-open-data-licensing>.

If you have ownership of a work, and someone else wants to use it, they have to ask your permission. Licences are how you explicitly give someone else permission to use that work. There are various licences that open-data is often released under and that describe how it can be used. Some of the most common are outlined below.

Open Government Licence

The OGL is used by the UK government allows you to:

- copy, distribute and transmit the data;
- adapt the data; and
- exploit the data commercially, whether by sub-licensing it, combining it with other data, or including it in your own product or application.

However, you must (where you do any of the above):

- acknowledge the source of the Information in your product or application by including or linking to any attribution statement specified by the Information Provider(s) and, where possible, provide a link to this licence

Full details of the OGL are here: <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>

Creative Commons

There are a range of Creative Commons (CC) licences which vary in how they allow you to redistribute, use the work commercially and non-commercially, and how you are permitted to remix, tweak, and build upon your work. The various CC licences are listed below and described in detail here: <https://creativecommons.org/licenses/>.

- Attribution - CC BY
- Attribution-ShareAlike - CC BY-SA
- Attribution-NoDerivs - CC BY-ND
- Attribution-NonCommercial - CC BY-NC
- Attribution-NonCommercial-ShareAlike - CC BY-NC-SA
- Attribution-NonCommercial-NoDerivs - CC BY-NC-ND

7. Software

To be able to use (visualise, analyse and present) spatial open-data you will need a GIS or web application. It is not appropriate in this note to comment on the merits of different commercial software, however it is worth noting that there are now viable open-source (free) alternatives to the commercial systems, some of which are introduced below.

QGIS

QGIS is a freely available open-source desktop Geographic Information System (GIS) that matches the functionality and usability of commercial products like MapInfo and ArcGIS. It has been adopted by governments, businesses and NGOs around the world.

The software is developed collaboratively by many people and all the source code is freely available in the open source world. The software, too is available to anyone free of charge. A key feature of the success of the QGIS open source project is a very high standard of project management and governance led by a team called the Project Steering Committee (PSC).

Development on QGIS proceeds very rapidly. The latest release is always available from this web page: <https://www.qgis.org/en/site/forusers/download.html>. The QGIS team aim to make the latest developments available to the user community as soon as possible which means that new releases are frequent (under the control of the QGIS Release Team). With such rapid production of new releases, there is a risk of a new release containing bugs. To guard against this, the QGIS team also maintain a stable long term release which many corporate users may prefer to use since it is much less likely to be 'buggy'.

CARTO

CARTO (formerly CartoDB) is a Software as a Service (SaaS) cloud computing platform that provides GIS and web mapping tools for display in a web browser. CARTO Builder is a web-based drag and drop GIS tool. Users can use the free platform or deploy their own instance of the open source software. CARTO is offered as freemium service, where accounts are free up to a certain size. For larger accounts, a fee is applied. Further information here: <https://carto.com/>

Leaflet

Leaflet is the leading open-source JavaScript library for creating mobile-friendly interactive maps and contains many useful mapping features for developers. It works across all major desktop and mobile platforms, can be extended with lots of plugins, has a well-documented API and a simple, readable source code. Further information here: <http://leafletjs.com/>

PostGIS

PostGIS is an open-source spatial database extender for PostgreSQL object-relational database. It adds support for geographic objects allowing location queries to be run in SQL. It also offers many features rarely found in other competing spatial databases such as Oracle Locator/Spatial and SQL Server. Further information here: <http://postgis.net/>

8. Data Sources

An accompanying spreadsheet on the LI website details where to find the most useful and relevant datasets for landscape professionals. It catalogues data in the UK, internationally, 'tags' data by theme and provides links to access the data and licence.

9. References and further reading

Carto: <https://carto.com/>

Creative Commons: <https://creativecommons.org/licenses/>

ESRI <https://www.esri.com>

Global Open Data Index <https://index.okfn.org/place/>

GovLab Index: <http://thegovlab.org/govlab-index-on-open-data-2016-edition/>

INSPIRE Directive: <http://inspire.ec.europa.eu/about-inspire/563>

Leaflet: <http://leafletjs.com/>

Open data barometer: http://opendatabarometer.org/?_year=2016&indicator=ODB

Open Data Institute: <https://theodi.org/>

Open Definition <http://opendefinition.org/>

Open Geospatial Consortium (OGC) <http://www.opengeospatial.org/>

Open Government Partnership: <https://www.opengovpartnership.org/about/about-ogp>

Open Knowledge International <https://okfn.org/>

PostGIS: <http://postgis.net/>

QGIS: <https://www.qgis.org/en/>

Web foundation: <https://webfoundation.org/>

Authored by Matt Davies

© May 2018

Landscape Institute

Charles Darwin House

107 Gray's Inn Road

London WC1X 8TZ

www.landscapeinstitute.org

Document history

Edited for publication – Simon Odell CMLI 21 May 2018