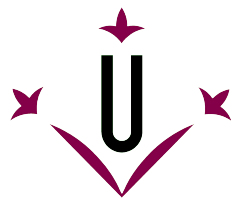

103088 - Massive Data Processing
P1A1 - TBL report

Jordi Rafael Lazo Florensa

4 March 2024

Master's degree in Computer Engineering



Universitat de Lleida
Escola Politècnica Superior

Contents

1	Introduction	2
2	Energy efficiency certificates for existing buildings in Lleida	2
3	List of accommodation facilities in the city of Barcelona	2
4	What is available openly at Lleida City Council	3
4.1	Foreign data 2019	3
4.2	Budget of the council of Lleida 2020	3
4.3	Population data 2021	3
5	Fuel consumption in Spain	4
6	License plate numbers in Spain.	4
7	References	4

1 Introduction

The “TBL classification” is a star-based system developed by Tim Berners-Lee [1], the inventor of the World Wide Web. This system is designed to evaluate the openness and linkability of data available on the web. It assigns a rating from one to five stars, with one star representing basic openness and five stars indicating a high level of openness and linkability.

In this assignment, we delve into the exploration and evaluation of various sources of open data. Our focus is on specific topics relevant to Spain, including energy efficiency certificates, accommodation facilities, city council information, fuel consumption, and registration numbers. The aim is to assess these data sources and understand their level of openness based on the TBL classification system.

The data will be classified according to TBL classification [2]:

- ★: make your stuff available on the Web (whatever format) under an open license.
- ★★: make it available as structured data (e.g., Excel instead of image scan of a table).
- ★★★: make it available in a non-proprietary open format (e.g., CSV instead of Excel).
- ★★★★: use URIs to denote things, so that people can point at your stuff.
- ★★★★★: link your data to other data to provide context.

2 Energy efficiency certificates for existing buildings in Lleida

Source [4]: https://icaen.gencat.cat/ca/energia/usos_energia/edificis/certificacio/observatori-de-la-certificacio-deficiencia-energetica-dedificis/index.html

TBL classification: ★

The data is available on the web. It leads to a webpage that contains information about energy efficiency certificates of buildings in Catalonia.

The user can filter the information according to their needs. In this case you can filter by all the buildings in Lleida according to their classification, however it is not possible to download the data in a CSV or XLS format. It is just a Microsoft Power BI that allows you to view the data but not access and download it.

3 List of accommodation facilities in the city of Barcelona

Source [5]: <https://opendata-ajuntament.barcelona.cat/data/ca/dataset/equipament-centres-d-informacio>

TBL classification: ★★★

On the Open Data BCN website of the Barcelona City Council, you can access the list of centers equipments in Barcelona. The dataset is in a structured and non-proprietary

format, which meets the criteria for a level 3 dataset according to the TBL classification system (even on the website you can see the 3 stars) .

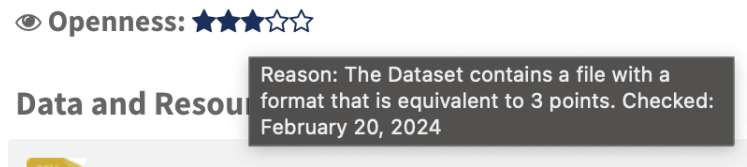


Figure 1: Openness of the data: List of information centers equipments in the city of Barcelona.

The dataset also includes a list of information centers in Barcelona, with various details about each center, such as their register ID, name, institution ID, and institution name, road and neighborhood information, and more. The dataset is detailed and complete enough to be considered a level 3 dataset.

4 What is available openly at Lleida City Council

4.1 Foreign data 2019

Source [6]: <https://aplicacionsweb.paeria.es/eOpenDataPublicWeb/faces/ca/cataleg/territori/demografia/dades-estrangers-2019/detalls>

TBL classification: ★★

The data can be downloaded in a structured CSV format containing the following columns: province code, province, municipality code, urban sector, nationality, men, women, total, % men and % women. The data is available in a non-proprietary format, it is structured data and available on the web in csv

4.2 Budget of the council of Lleida 2020

Source [7]: https://aplicacionsweb.paeria.es/eOpenDataPublicWeb/faces/ca/cataleg/economia_i_empresa/economia/pressupost-ajuntament-lleida-2020/detalls

TBL classification: ★★

The data can be downloaded in a structured CSV format where the episode, application, description and budget columns of each project appear. The data is available in a non-proprietary format, it is structured data and available on the web in csv.

4.3 Population data 2021

Source [8]: <https://aplicacionsweb.paeria.es/eOpenDataPublicWeb/faces/ca/cataleg/territori/demografia/dades-poblacionals-2021/detalls>

TBL classification: ★★

The data is organized by district, section, neighborhood, sex, and age groups. The dataset is in a CSV format, which is simple to read and use by machines. The data is detailed enough to reveal population trends and demographic shifts over time.

5 Fuel consumption in Spain

Source [9]: <https://www.epdata.es/datos/consumo-carburantes-gasolina-gasoleo-espana-estadisticas-datos/326/espana/106>

TBL classification: ★★

The web page has a brief introduction that explains the main trends and variations in the consumption of gasoline and diesel in Spain. The first graph shows the total consumption of gasoline and diesel in Spain in kilotons (thousands of tons) for each month from January 2010 to August 2019.

This data it has 3 stars because it is available in a non-proprietary format, it is structured data and you can download it either in csv or json.

6 License plate numbers in Spain.

Source [10]: <https://apps.fomento.gob.es/bdotle/visorBDpop.aspx?i=396>

TBL classification: ★★

The web page that contains information about the national park of vehicles by autonomous community, province, type of vehicle and type of fuel in Spain.

It also shows the total number of vehicles for each type of vehicle and type of fuel for the whole country, as well as the total number of all vehicles. The data can be viewed in a table or graphs.

The data on this web page are open to different degrees, depending on the format and structure. Some data are in .pdf or .docx files, which are not easily reusable. Some data are in .xls or .tiff files, which are non-proprietary but still not very accessible. The most open data are in .csv or .json files, which are non-proprietary, structured, and web-based.

7 References

- [1] Tim Berners-Lee
https://en.wikipedia.org/wiki/Tim_Berners-Lee
- [2] 5 stars - Open Data
<https://5stardata.info/en/>
- [3] Dades Obertes Catalunya - Certificats d'eficiència energètica d'edificis
https://analisi.transparenciacatalunya.cat/Energia/Certificats-d-eficiencia-energ-tica-d-edificis/j6ii-t3w2/about_data
- [4] Institut Català d'Energia - Observatori de la Certificació d'eficiència energètica d'edificis

https://icaen.gencat.cat/ca/energia/usos_energia/edificis/certificacio/observatori-de-la-certificacio-deficiencia-energetica-dedificis/index.html

- [5] Open Data BCN - Llista d'equipaments de centres d'informació de la ciutat de Barcelona
<https://aplicacionsweb.paeria.es/eOpenDataPublicWeb/faces/ca/cataleg/territori/demografia/dades-estrangers-2019/detalls>
- [6] Dades obertes de Lleida - DADES ESTRANGERS 2019
<https://opendata-ajuntament.barcelona.cat/data/ca/dataset/equipament-centres-d-informacio>
- [7] Dades obertes de Lleida - PRESSUPOST DE L'AJUNTAMENT DE LLEIDA 2020
https://aplicacionsweb.paeria.es/eOpenDataPublicWeb/faces/ca/cataleg/economia_i_empresa/economia/pressupost-ajuntament-lleida-2020/detalls
- [8] Dades obertes de Lleida - DADES POBLACIONALS 2021
<https://aplicacionsweb.paeria.es/eOpenDataPublicWeb/faces/ca/cataleg/territori/demografia/dades-poblacionals-2021/detalls>
- [9] EP data - Consumo de carburantes, gasolina y gasóleo en España, estadísticas y datos
<https://www.epdata.es/datos/consumo-carburantes-gasolina-gasoleo-espana-estadisticas-datos/326/espana/106>
- [10] Observatorio del transporte y logística de España - Parque nacional de vehículos por comunidad autónoma, provincia, tipo de vehículo y tipo de carburante
<https://apps.fomento.gob.es/bdotle/visorBDpop.aspx?i=396>