

# *OpenData*

*Máster en Business Intelligence e  
Innovación Tecnológica*

## Que es OpenData?


El concepto datos abiertos (open data, en inglés) es una filosofía y práctica que persigue que determinados tipos de datos estén disponibles de forma libre para todo el mundo, sin restricciones de derechos de autor, de patentes o de otros mecanismos de control. Tiene una ética similar a otros movimientos y comunidades abiertos, como el software libre, el código abierto (open source, en inglés) y el acceso libre (open access, en inglés).


## Acceso a los Datos

Los datos suelen estar publicados para ser explotados desde elementos externos. Algunos modos de publicación:

- **Descarga de ficheros.** El método más tradicional, consiste en la descarga de los ficheros en formato de texto plano, por ejemplo csv.
- **API's REST.** Accediendo a los datos mediante servicios web tipo los microservicios, la mensajería se realiza mediante peticiones y respuestas JSON.
- **SPARQL.** Algunos sitios publican los datasets para ser accedidos mediante SPARQL, un lenguaje de consulta para base de datos de tipo grafos que permite acceder de forma remota.

# Amazon Data Sets

 Menu



Getting Started

Products

Solutions

Pricing

Software

English ▼

DATA SET CATEGORIES

Astronomy >

Biology >

Chemistry >

Climate >

Economics >

Encyclopedic >

Geographic >

Mathematics >

RELATED LINKS

Amazon Machine Images (AMIs)

Articles & Tutorials

Actualmente podemos mostrar este contenido sólo en inglés. Trabajamos constantemente para agregar contenido en español. Gracias por su paciencia.

## AWS Public Data Sets

Public Data Sets on AWS provides a centralized repository of public data sets that can be easily integrated into AWS cloud-based applications. AWS is hosting the public data sets for the research community, and like all AWS services, users pay only for the compute and storage used in their applications. Learn more about [Public Data Sets on AWS](#) and visit the [Public Data Sets on AWS](#) page.

### CCAFS-Climate Data

High resolution climate data to help assess the impacts of climate change primarily on agriculture. The climate projections will help researchers make climate change impact assessments.

Last Modified: February 9, 2016

## Ayuntamiento BCN

The screenshot displays the OpenDataBCN website, which is the open data portal for the Ayuntamiento de Barcelona. The header includes the city's logo and navigation links in Catalan. The main navigation bar contains links to '¿Qué es Open Data?', 'Catálogo de datos', 'Visualizador', 'Destacamos', 'Enlaces', 'Actualidad', and 'eGovernament'. The 'Catálogo de datos' link is active, leading to the 'Catálogo de Datos Abiertos' page. This page features a search bar with a 'Buscar' button and a checkbox for 'Mostrar descripció estesa'. Below the search bar, it states there are 331 data sets in the catalog and provides a link to download the catalog in RDF format. A featured dataset, 'Accidentes gestionados por la Guardia Urbana en la ciudad de Barcelona', is highlighted with a description and available formats (XLSX, ODATA, CSV, XML). A sidebar on the right offers filters by 'Temas' and 'Frecuencia'.

**Ajuntament de Barcelona**

**OpenDataBCN**  
Servicio de datos abiertos del Ayuntamiento de Barcelona

> ¿Qué es Open Data? > Catálogo de datos > Visualizador > Destacamos > Enlaces > Actualidad > eGovernament

Inicio » Catálogo de datos

### Catálogo de Datos Abiertos

Introduzca las palabras clave de búsqueda. Puede utilizar también los filtros laterales

☐ Mostrar descripció estesa

Hay un total de 331 conjuntos de datos en el catálogo [Descarga el catálogo en formato RDF](#)

#### Accidentes gestionados por la Guardia Urbana en la ciudad de Barcelona

Listado de los accidentes gestionados por la Guardia Urbana en la ciudad de Barcelona. El sistema en el que se expresan las coordenadas es el UTM31/ED50. Incorpora el número de lesionados segun gravedad, el número de vehiculos y el punto de impacto.

**Formats disponibles**

**Temas**

- ☐ Ad
- ☐ Cit
- ☐ Eco
- ☐ Po
- ☐ Te

**Frecue**

- Inme
- Diari
- Semi
- Mens
- Trim



## TheGamesDB.net



## Otros

# The Linking Open Data cloud diagram

insight



NUI Galway

Institut

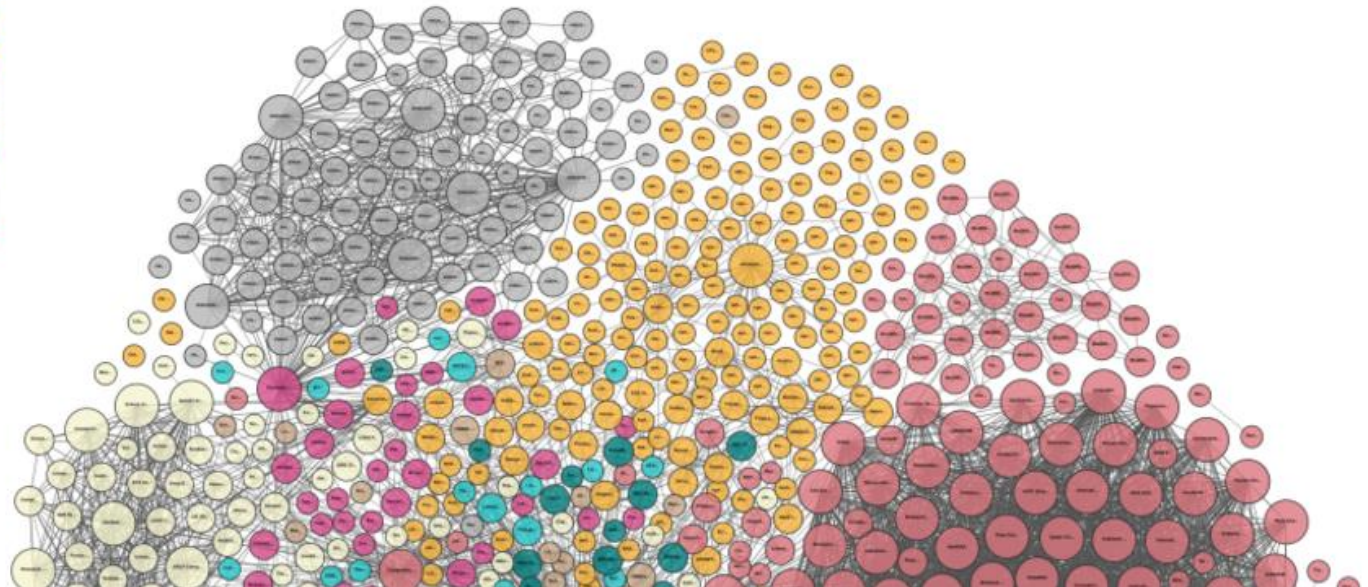
MANNHEIM

This web page is the home of the *LOD cloud diagram*. This image shows datasets that have been published in [Linked Data](#) format, by contributors to the [Linking Open Data](#) community project and other individuals and organisations. It is based on metadata collected and curated by contributors to the [Data Hub](#). Clicking the image will take you to an interactive SVG version, where each dataset is a hyperlink to its entry in Datahub.

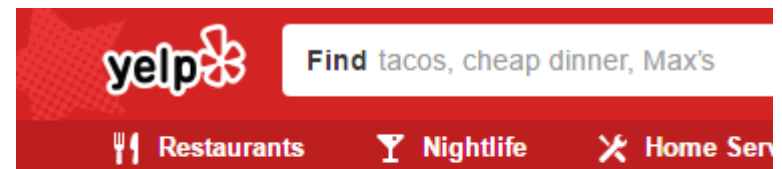
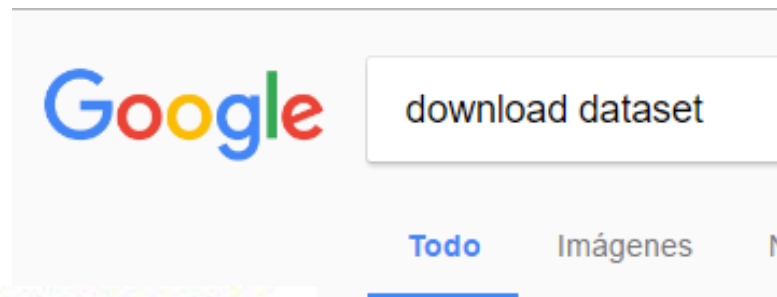
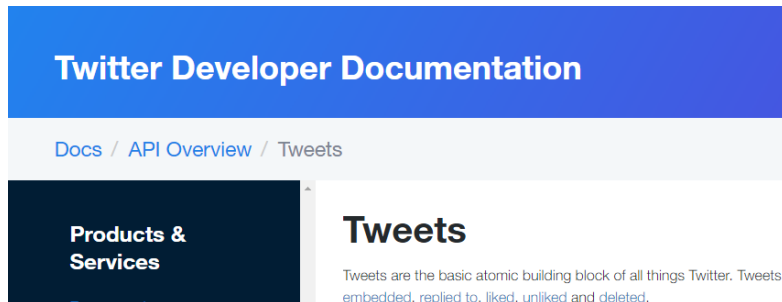
The diagram is maintained by [Andrejs Abele](#) and [John McCrae](#) ([Insight Centre for Data Analytics](#) at NUI Galway). For any questions and comments, please email [andrejs.abele@insight-centre.org](mailto:andrejs.abele@insight-centre.org) and [John.McCrae@insight-centre.org](mailto:John.McCrae@insight-centre.org). The original version was developed by Richard Cyganiak and Anja Jentzsch.

Last updated: 2017-02-20

### Legend



## Cualquier otro



## Yelp Dataset Challenge



## Requisitos

A partir de vuestro dataset favorito, utilizad cualquiera de las herramientas vistas en el curso para cargarlo y manipularlo:

- MongoDB
- Cassandra
- Python
- Hive
- Pig
- MapReduce
- Spark

Podéis hacerlo en grupos de 3 compañeros y utilizar una o varias de estas tecnologías.

## Entregable

- El entregable debe ser únicamente un documento
- Identificad el dataset y su origen
- Explicad cómo lo habéis cargado: pasos, dificultades, decisiones, etc.
- Cargad los datos y tratadlos: explotad la información para analizarla y presentar conclusiones
- Detallad el modelo de datos con el que habéis trabajado
- Se valorará la complejidad de la solución, el ingenio, trabajo con el dataset, su explotación,...