M2B – Mapas y OpenData: herramientas de localización, visualización y análisis de geodatos

TECNOLOGÍAS SIG

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
School of Professional & Executive Development

Máster SIG

OpenData

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
School of Professional & Executive Development

Datos Abiertos (Open Data)

Para desarrollar las políticas, los gobiernos deben abrir sus datos:
- para ser transparentes con la ciudadanía (open Government).
- Para que las empresas puedan añadir valor y generar riqueza.

El objetivo es proporcionar toda la información pública a la ciudadanía, en un formato fácil de manipular, para que se puedan convertir en servicios públicos o privados con valor añadido. Datos meteorológicos, de equipamientos, estadísticas, de presupuestos, jurídicas y Judiciales...

De hecho, hay muchos datos publicados en webs y boletines oficiales, pero su formato de publicación actual hace más difícil el tratamiento y reaprovechamiento fácil.

El movimiento open data se está extendiendo por todo el mundo y la Administración se engloba dentro la cultura de cambio en la concepción, gestión y prestación del servicio público.

Fuente:Datos abiertos gencat. Dirección General de Atención Ciudadana y Difusión. Febrero de 2011

Datos Abiertos (Open Data)

En Europa y en Cataluña, open data = reutilización = RISP

Reutilización de la información del sector público (RISP): uso que pueden hacer las personas, organizaciones y empresas

Base legal:
- Ley 37/2007, de RISP (+ Directiva 2003/98/CE)

- Ley 20/2010 del uso de los medios electrónicos en el sector público de Cataluña

Fuente:Datos abiertos gencat. Dirección General de Atención Ciudadana y Difusión. Febrero de 2011

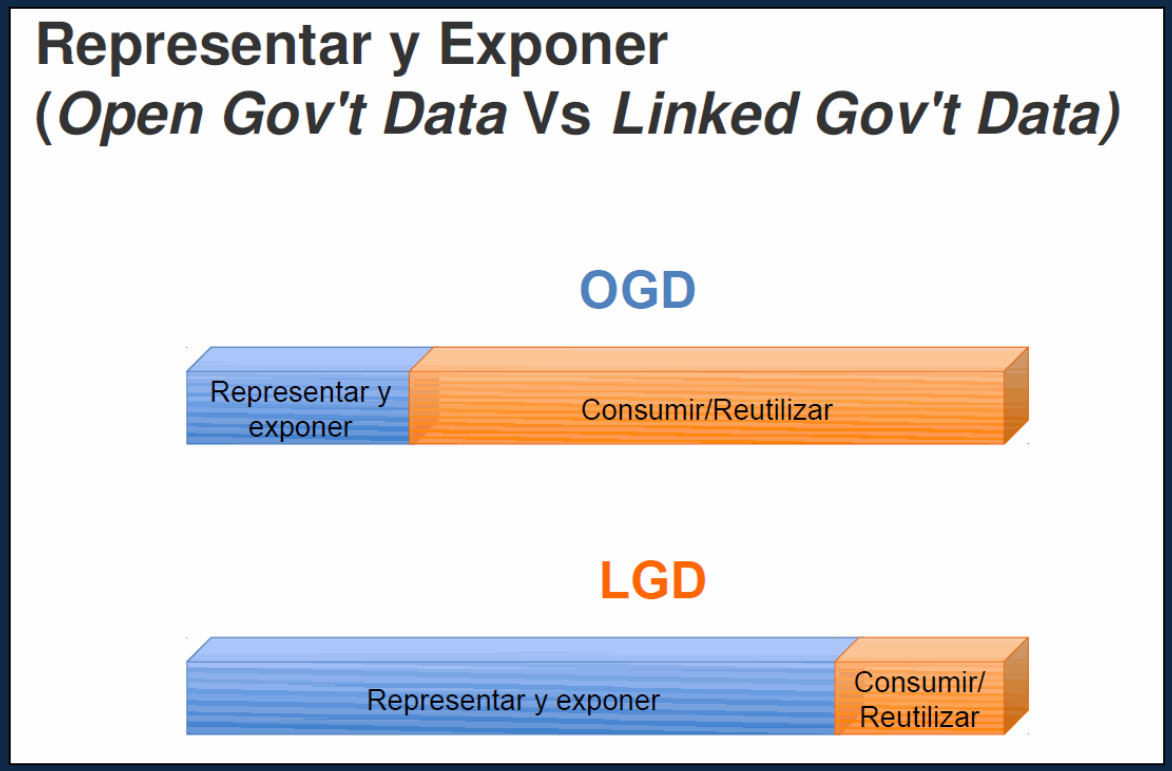
Datos Abiertos (Open Data)

Open data: Modelos

Modelo Kundra. Pone el foco en liberar conjuntos de datos,
en cualquier formato mínimamente estructurado.

Modelo Berners-Lee. Hace un esfuerzo adicional por contribuir a la generación de *Linked Data*

Datos Abiertos (Open Data)



http://www.w3c.es/Presentaciones/2011/0317-geoInteroperabilidad-MA/0317-geoInteroperabilidad-MA.pdf

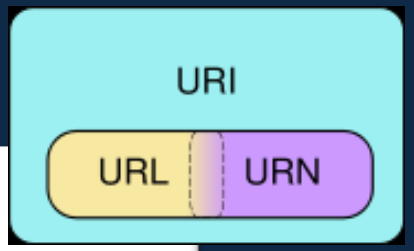
Datos Abiertos (Open Data): Linked Data

T. Berners-Lee (2006)

Define los “principios básicos de Linked Data”*

- Linked Data hace referencia a (técnica) como publicar
- datos utilizando la WEB
- Web de los datos vs Web de los documentos(tradicional)
- La Web como una base de datos
- Conectar –utilizando la Web- datos de diferentes dominios: publicaciones científicas, datos estadísticos, libros, compañías...

Principios básicos de Linked Data



1. Utilizar URIs per identificar recursos

2. Utilizar HTTP de las URIs per poder localizar los recursos

3. Proporcionar información útil sobre el recurso de la Uri, utilizando estándares W3C

- RDF(Resource Description Framework)
- SPARQL (Protocol and RDF Query Language)

4.Incluir enlaces a otras URI relacionadas con los recursos(nube de enlaces)

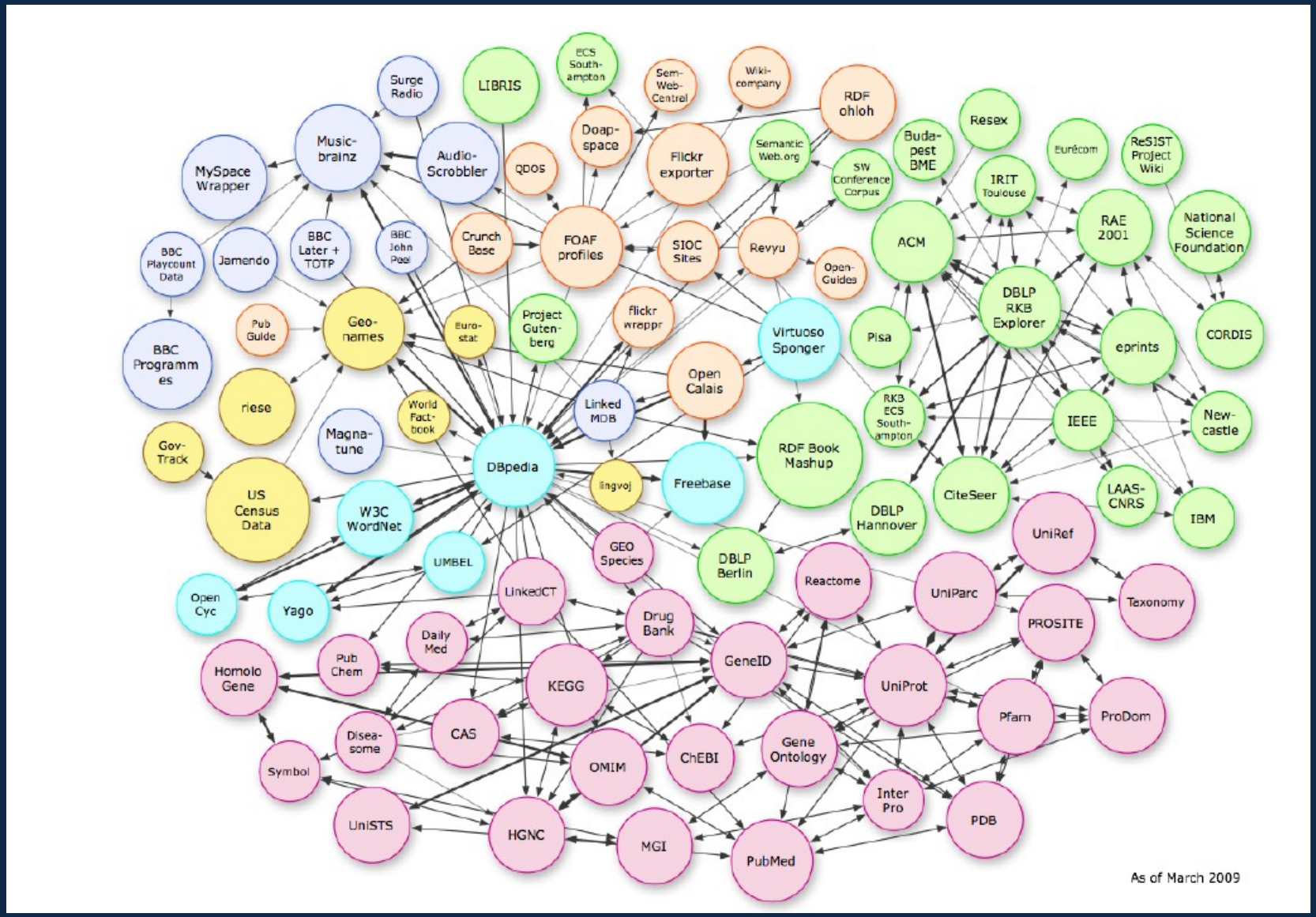
http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html

Ejemplo SPQRL

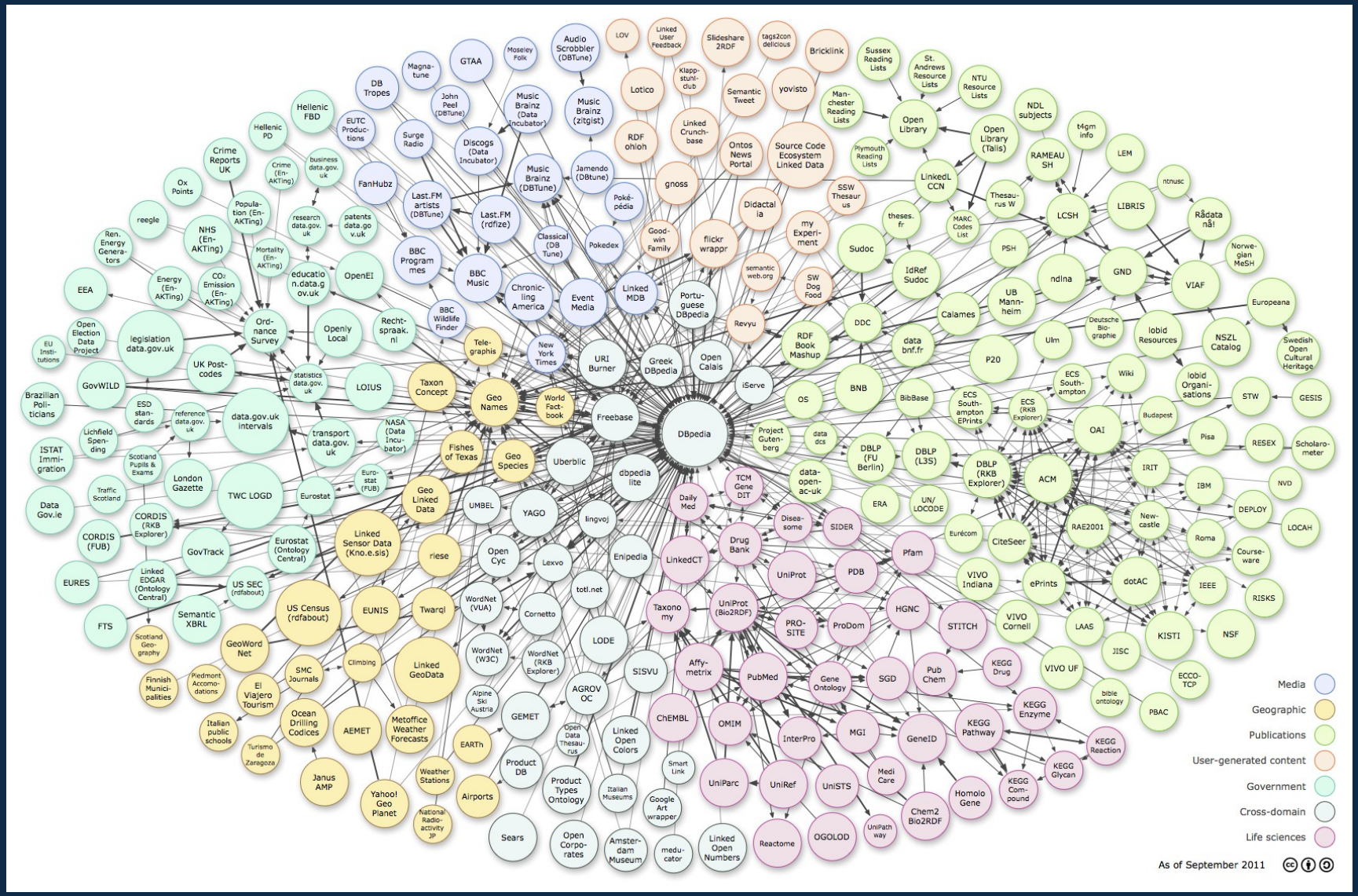
Todas la personas nacidas en Barcelona después de 1960 que están en la Wikipédia

http://dbpedia.org/snorql/?query=SELECT+%3Fname+%3Fbirth++%3Fperson+WHERE+%7B++++++%3Fperson+dbo%3AbirthPlace+%3ABarcelona+.++++++%3Fperson+dbo%3AbirthDate+%3Fbirth+.++++++%3Fperson+foaf%3Aname+%3Fname+.+++++%3Fperson+dbo%3AbirthDate+%3Fbirth+.++++++FILTER+%28%3Fbirth+%3E+%221960-01-01%22%5E%5Exsd%3Adate%29+.+%7D+ORDER+BY+%3Fname

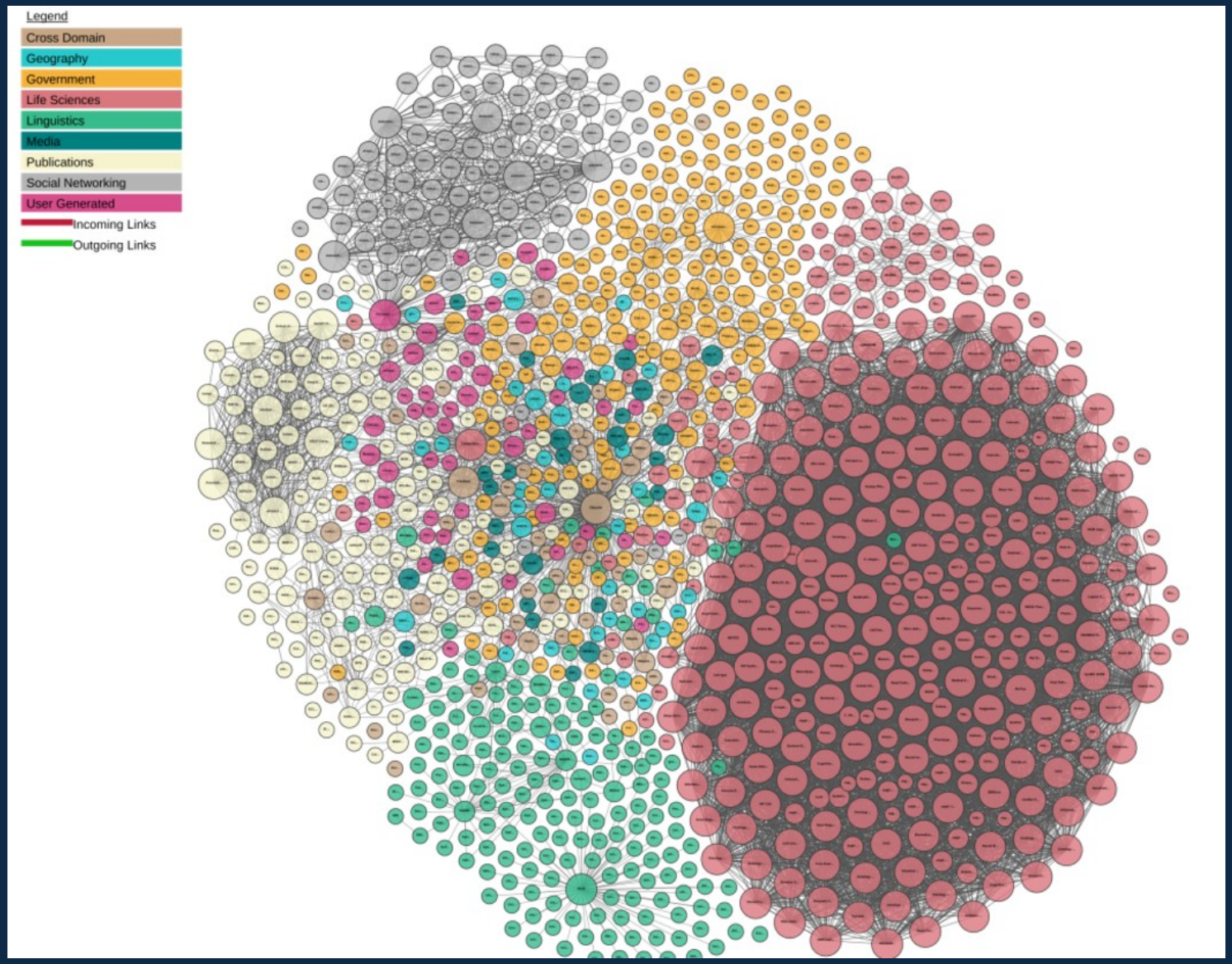
Datos Abiertos (Open Data): Linked Data



Datos Abiertos (Open Data): Linked Data



The Linking Open Data cloud diagram



http://lod-cloud.net/

“Otros“ Datos Abiertos: OpenStreetMap

OpenStreetMap empezó en 2004 el Reino Unido en respuesta a las duras restricciones de copyright del Ordnance Survey.

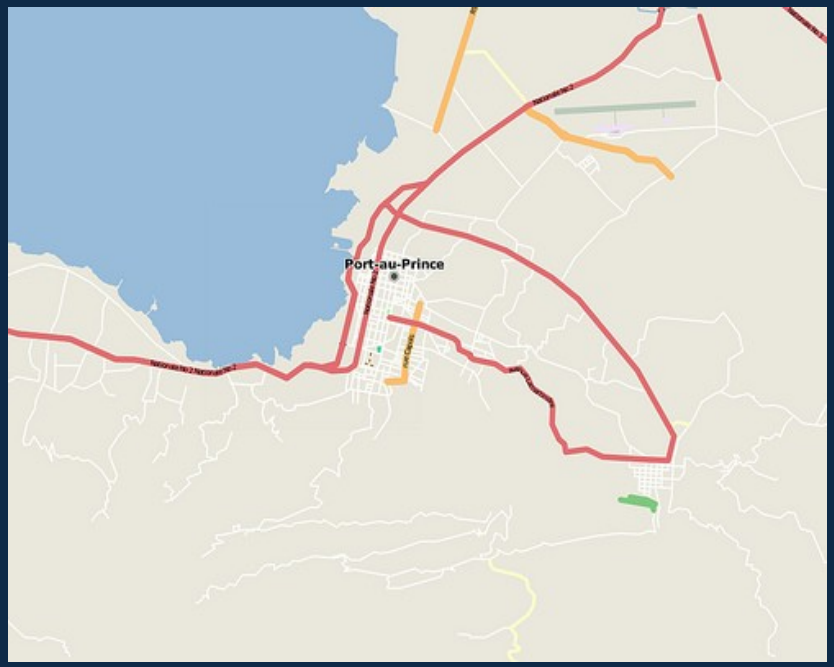
Es un proyecto colaborativo para crear mapas libres y editables.

Los mapas se crean utilizando información geográfica capturada con dispositivos GPS móviles, ortofotografías y otras fuentes libres. Esta cartografía, tanto las imágenes creadas como los datos vectoriales almacenados en su base de datos, se distribuye bajo licencia abierta Licencia Abierta de Bases de Datos (en inglés ODbL).(wikipedia)

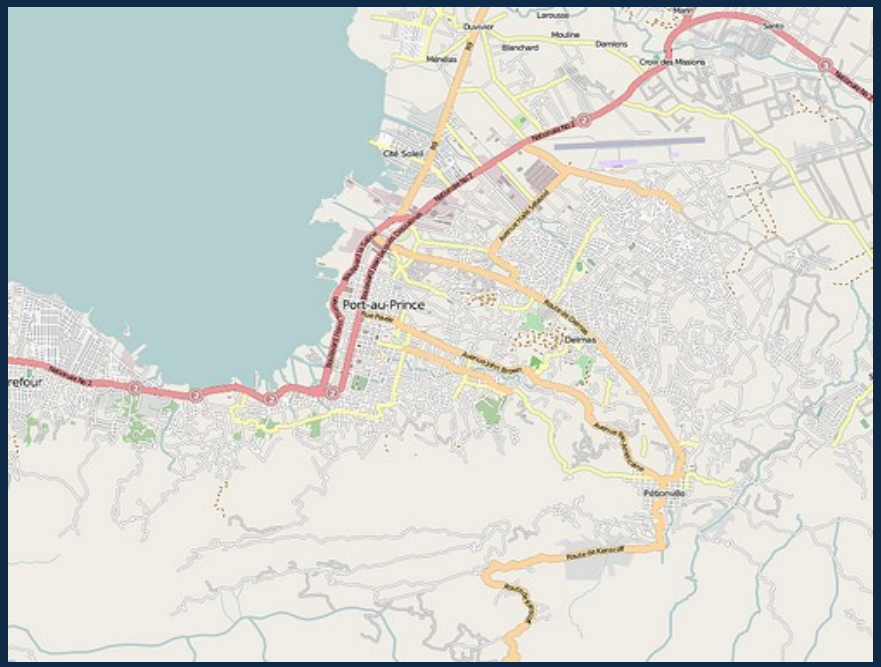
https://www.openstreetmap.org

“Otros“ Datos Abiertos: OpenStreetMap

Terremoto de Haití de 2010



antes



después

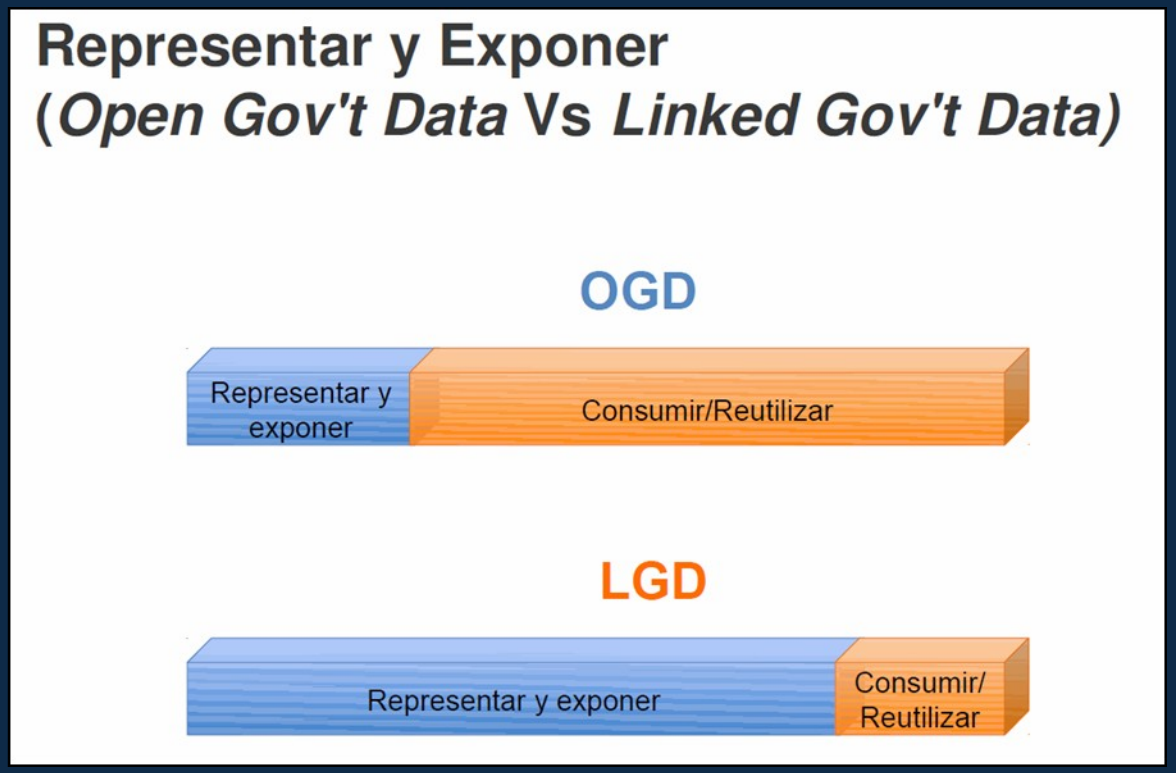
“Otros“ Datos Abiertos: OpenStreetMap

- Cualquier dato de OSM puede ser descargado y reutilizado
- Cobertura mundial
- Muchas empresas han creado un modelo de negocio basado en estos datos.
- Modelo de datos flexible
  - Los **nodos** (*nodes*). Son puntos que recogen una posición geográfica dada.
  - Las **vías** (*ways*). Son una lista ordenada de nodos que representa una linia opoligono(cuando una polilínea empieza y finaliza en el mismo punto).
  - Las **relaciones** (*relations*). Son grupos de nodos, vías y otras relaciones a las que se pueden asignar determinadas propiedades comunes.
  - Las **etiquetas** (*tags*). Se pueden asignar a nodos, caminos o relaciones y constan de una clave (*key*) y de un valor (*value*). Por ejemplo: highway=trunk

Conceptos (Open Data)

Muchos servicios y/o información relativa a Smart City, son expuestos cómo OpenData (Datos Abiertos) en portales de administraciones públicas.

Son las llamada “Plataformas “ para la publicación y gestión de OpenData

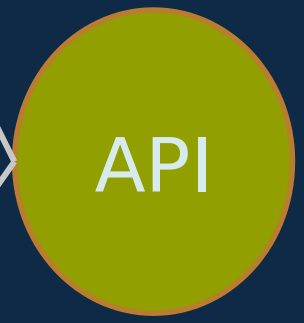


Las nuevas Plataformas de publicación OpenData equilibran la barra entre los esfuerzos entre representar/exponer y consumir/reutilizar.

Arquitectura Plataformas de publicación

Estructura Front-End (web)

Generación de portales
Herramientas gestión interna de datos
Herramientas visualizaciones de datos.
Plugins, ejemplo para Geodatos



Un portal de un cliente puede:

Residir en la nube
En servidores proveedor
En servidores cliente

Plataformas Open Data

¿Puedo desarrollar nuevos Plugins para las Plataformas?

¿Podría crear una API de APIs?

Plataformas Open Data

VAMOS INTRODUCIRNOS EN 3 PLATAFORMAS ...

- **Socrata**
- **OpenDatasoft**
- **Ckan**

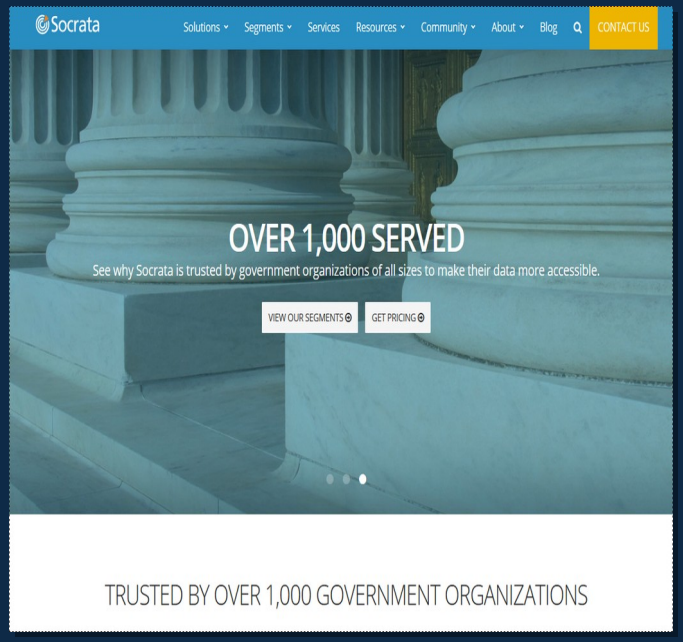
Socrata

Sede: Seattle, Washington, USA

Fundada: 2007

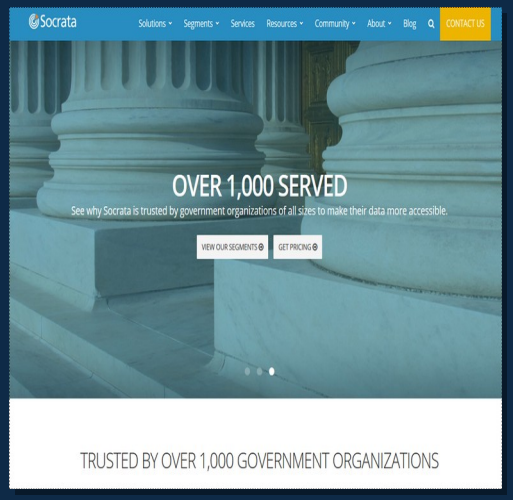
Modelo negocio:
Software-as-a-Service bajo licencia.

API y SDK's (acceso a datos) OpenSource.



Socrata is a company that provides cloud-based data visualization and analysis tools for opening government data. Originally called Blist, Socrata was founded in February 2007. Socrata targets non-technical Internet users who want to view and share government, healthcare, energy, education, or environment data. Its products are issued under a proprietary, closed, exclusive license

Fuente:
https://en.wikipedia.org/wiki/Socrata



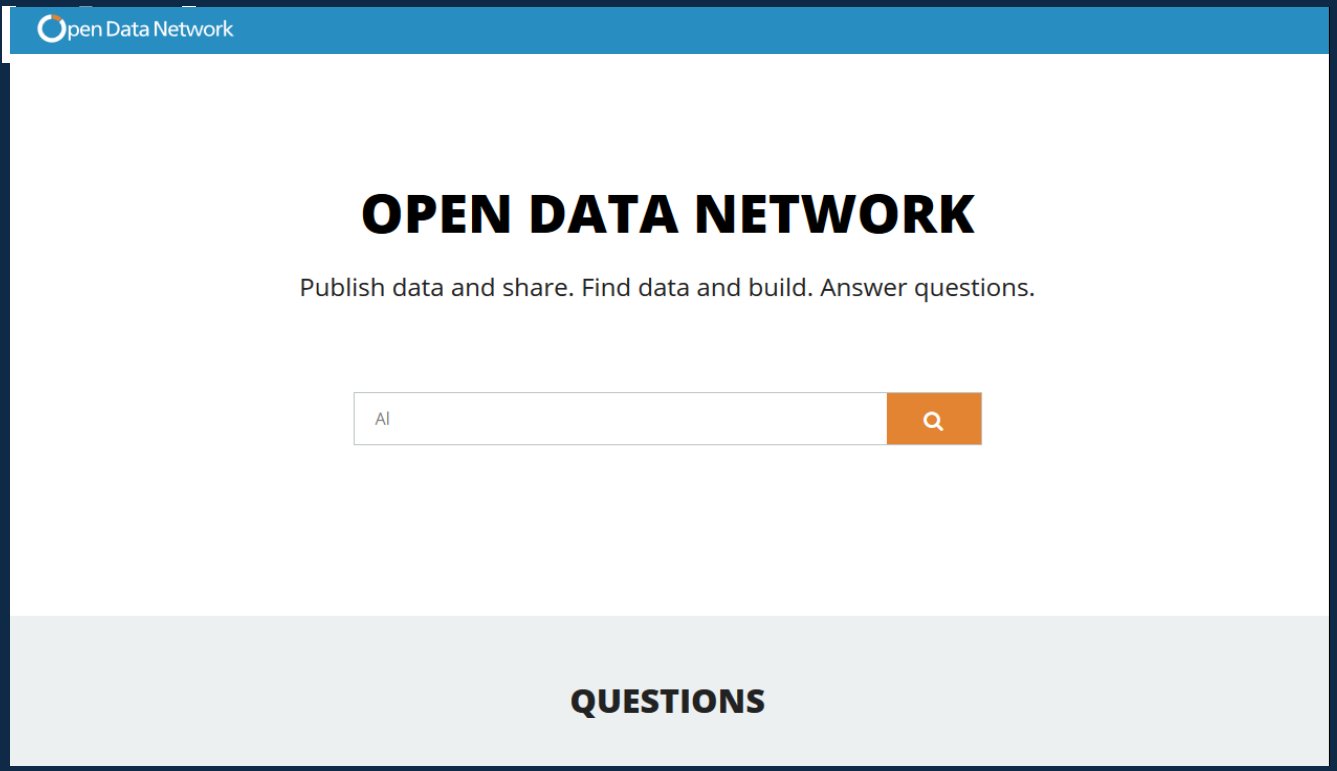
Links

https://socrata.com/

https://socrata.com/solutions/publica-open-data-cloud/

https://github.com/socrata

SOCRATA : Portal g



https://www.opendatanetwork.com/

https://socrata.com/blog/socrata-introduces-open-data-network/

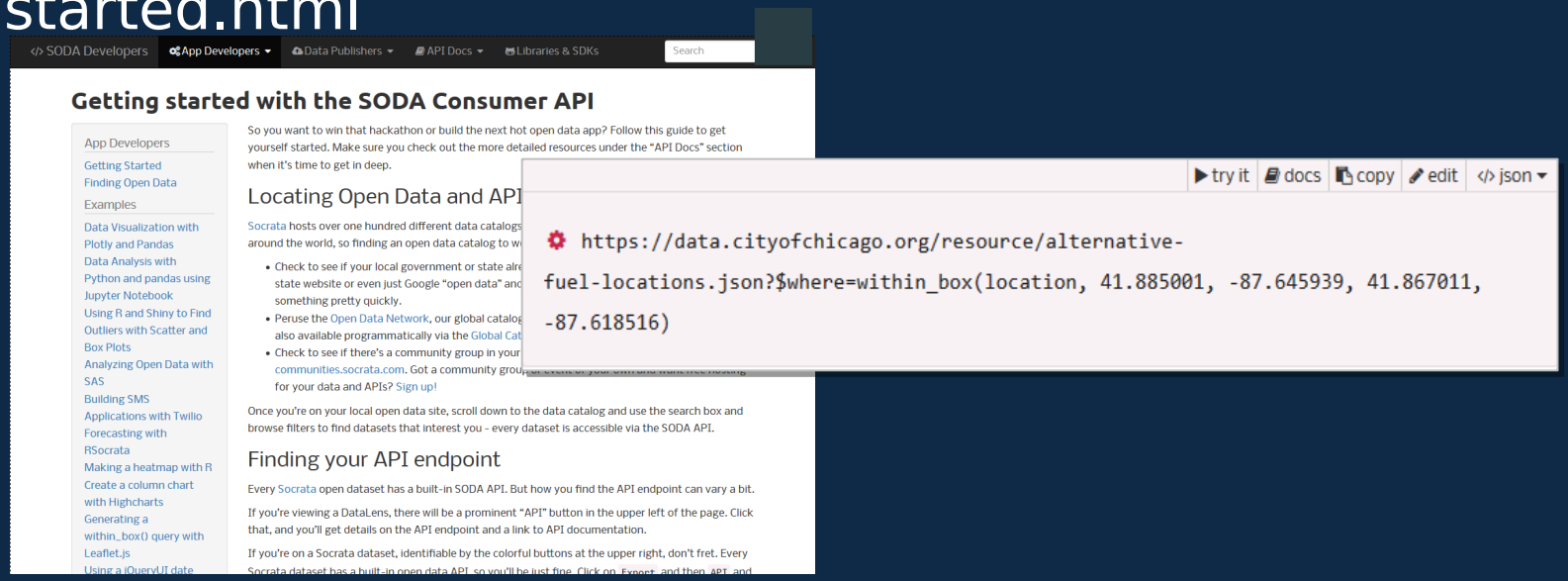
SOCRATA APIs:

- **DISCOVERY API**

  http://docs.socratadiscovery.apiary.io/

- **API (SODA)**

https://dev.socrata.com/consumers/getting-started.html



Discovery API

Discovery API

Socrata tiene la API SODA que me permite interrogar recursos –datasets- que están en los portales basados en Socrata

Para encontrar estos portales puedo utilizar
https://www.opendatanetwork.com/

{O
R }

Utilizar la Discovery API (Apiary) de SOCRATA

http://docs.socratadiscovery.apiary.io

Discovery API

Discovery API nos permite listar y buscar en todos los dominios (portales) de Estados Unidos o resto de el mundo.

Podemos Filtrar por

- **Search_context / domain: {domain}**
- **Categories: {Education}**
- **Assets: {Datasets, Maps, Charts}**
- **Attributions: {organizations}**
- **q: {free text}**
- **Derived_view: {id_dataset}**
- **limit: {num results}**

http://api.us.socrata.com/api/catalog/v1?q=crime&limit=3

Discovery API: US y EU

http://api.eu.socrata.com/api/catalog/v1?q=crime&limit=3

http://api.us.socrata.com/api/catalog/v1?q=crime&limit=3

Discovery API: Respuesta


```
▾ resource: {
  ▾ view_count: {
      page_views_total: 3950,
      page_views_total_log: 11.948002130977672,
      page_views_last_week_log: 4.0,
      page_views_last_month: 35,
      page_views_last_week: 15,
      page_views_last_month_log: 5.169925001442312
    },
    obe_fxf: null,
    description: "Seattle Part 1 Crime stats by precinct\nThe precinct and beat data can be found at\nhttp://data.seattle.gov/Government/Seattle
    Police-Department-Beats/nnxn-434b",
    name: "Crime",
  ▾ parent_fxf: [
      "3xqu-vnum",
      "hapq-73pk"
    ],
    nbe_fxf: null,
    attribution: "City of Seattle, Department of Information Technology",
    provenance: "community",
  ▸ columns_field_name: [ … ],
    download_count: 167,
  ▸ columns_name: [ … ],
  ▸ page_views: { … },
    updatedAt: "2016-08-30T22:  00Z",
    type: "chart",
    id: "pvkf-73v3",
    createdAt: "2010-11-18T00:30:51.000Z",
  ▾ columns_description: [
      "Description of the crime type",
      "Number of times crime occurred in a beat for reported month",
      "The police beat for reported crimes.",
      "Type 1 Crime committed"
    ]
  },
▸ classification: { … },
▾ metadata: {
    domain: "data.seattle.gov"
  },
  permalink: https://data.seattle.gov/d/pvkf-73v3,
  link: https://data.seattle.gov/Public-Safety/Crime/pvkf-73v3,
  preview_image_url: https://data.seattle.gov/views/pvkf-73v3/files/5c49b088-1863-4ebd-adaf-a9cc5d2d5fcb
}
],
```


Parientes

Tipo

Id

Dominio

Vinculo Dataset

Discovery API: Respuesta


```
 "type": "dataset",
   "id": "pvkf-73v3",
"metadata": {
data.seattle.gov    "domain": ""

},
"permalink": "https://data.seattle.gov/d/pvkf-73v3",
"link": "https://data.seattle.gov/Public-Safety/Crime/pvkf-73v3
```


URL derivadas de la respuesta JSON

https://data.seattle.gov/d/pvkf-73v3

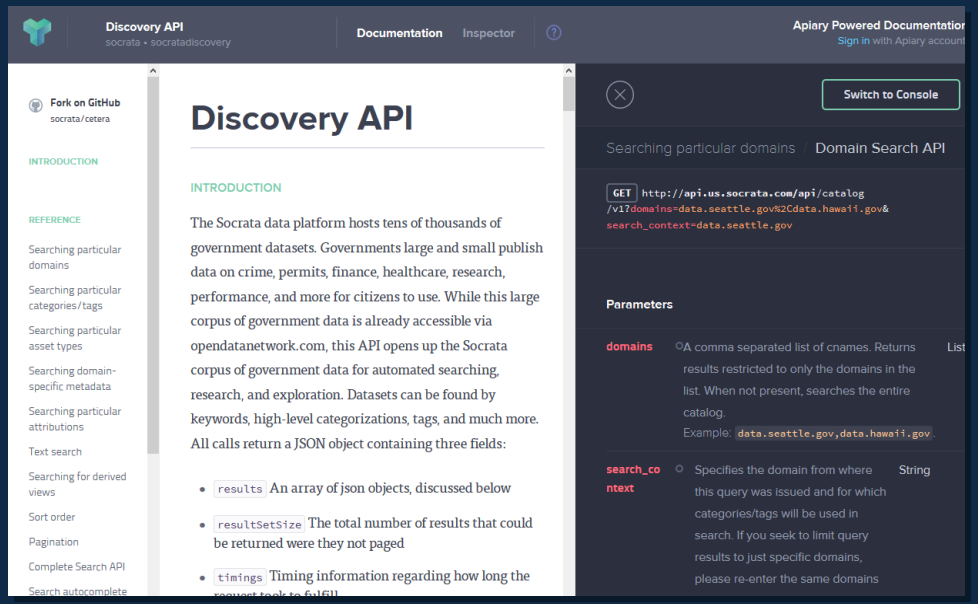
https://data.seattle.gov/api/views/pvkf-73v3/rows.csv?accessType=DOWNLOAD

http://api.us.socrata.com/api/catalog/v1/?derived_from=pvkf-73v3

https://dev.socrata.com/foundry/data.seattle.gov/pvkf-73v3

Vínculo con API SODA

https://docs.socratadiscovery.apiary.io



API
SODA

API SODA

https://dev.socrata.com/foundry/data.seattle.gov/pvkf-73v3

Simple Filters

Filtering data is very straightforward. SODA APIs are self-describing — the schema and contents of the dataset itself determines how you can query it. Any field within the data can be used as a filter, simply by appending it to the API endpoint as a GET parameter.

SoQL Queries

The “Socrata Query Language” (SoQL) is a simple, SQL-like query language specifically designed for making it easy to work with data on the web. The language is both powerful and easy to learn, and everything works via GET parameter

API SODA (SoQL)

Parameter	Description	Default	In $query
$select	The set of columns to be returned, similar to a SELECT in SQL	All columns, equivalent to $select=*	SELECT
$where	Filters the rows to be returned, similar to WHERE	No filter	WHERE
$order	Column to order results on, similar to ORDER BY in SQL	Unspecified order	ORDER BY
$group	Column to group results on, similar to GROUP BY in SQL	No grouping	GROUP BY
$having	Filters the rows that result from an aggregation, similar to HAVING	No filter	HAVING
$limit	Maximum number of results to return	1000 (2.0 endpoints: maximum of 50,000; 2.1: unlimited »)	LIMIT
$offset	Offset count into the results to start at, used for paging	0	OFFSET
$q	Performs a full text search for a value.	No search	N/A
$query	A full SoQL query string, all as one parameter	N/A	N/A
$$bom	Prepends a UTF-8 Byte Order Mark to the beginning of CSV output	false	N/A

https://dev.socrata.com/docs/queries/

API SODA (SoQL): Ejemplos

Simple feature
https://finances.worldbank.org/resource/45tv-a6qy.json?region=AFR

SoQL
https://finances.worldbank.org/resource/45tv-a6qy.json?$where=region = 'AFR'

https://finances.worldbank.org/resource/45tv-a6qy.json?$where=region = 'AFRICA' AND contract_signing_date > '2015-01-01T14:00:00'

Discharger CSV
https://finances.worldbank.org/resource/45tv-a6qy.csv

Web
https://finances.worldbank.org/Procurement/Major-Contract-Awards/kdui-wcs3

https://dev.socrata.com/docs/queries/

API SODA (SoQL): Funciones espaciales

within_box(...)	Returns the rows that have geodata within the specified box, defined by latitude, longitude corners
within_circle(...)	Returns the rows that have locations within a specified circle, measured in meters

https://data.cityofchicago.org/resource/alternative-fuel-locations.json?$where=**within_box(location, 41.885001, -87.645939, 41.867011, -87.618516)**

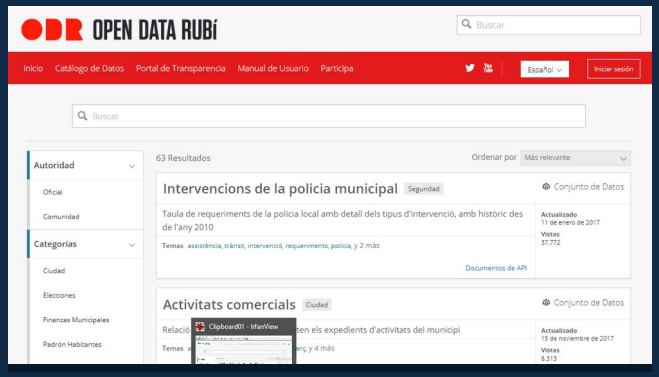
https://data.seattle.gov/resource/3k2p-39jp.json?$where=**within_circle(incident_location, 47.59815, -122.334540, 500)**

https://dev.socrata.com/docs/queries/

Portales con Socrata:

https://analisi.transparenciaatalunya.cat/

https://opendata.rubi.cat/es/browse



OPENDATASOFT

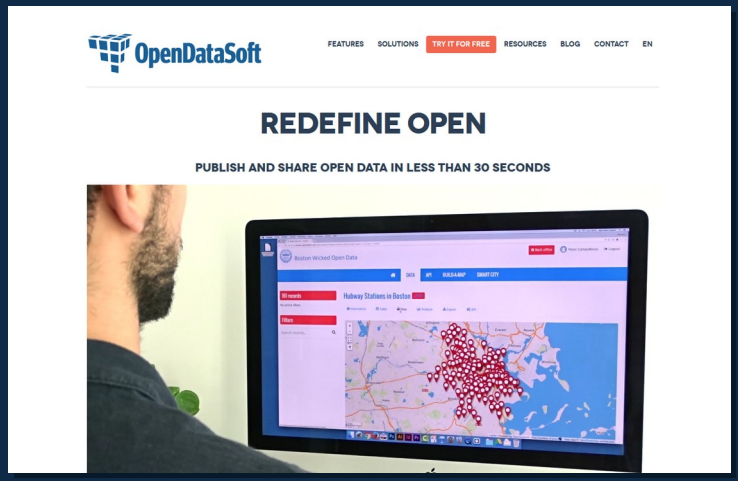
Sede: Paris, Fr

Fundada: 2011

Modelo negocio:
Software-as-a-Service bajo licencia.

API y SDK's (acceso a datos)
 OpenSource

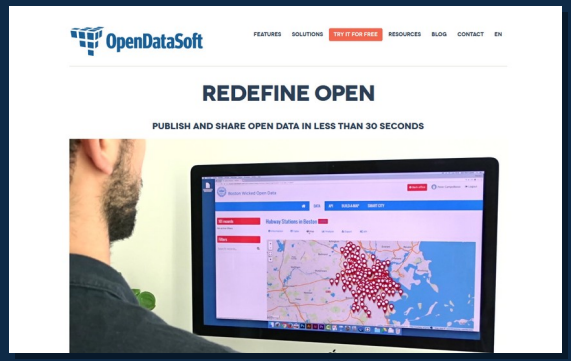
Gratuito para "nonprofit" y ONG



OpenDataSoft is a private software company specialized in transforming structured data into API and visualizations. Founded in 2011, OpenDataSoft targets non-technical users who wish to share and visualize government, health, energy and environmental data. OpenDataSoft allows restricted and open sharing ecosystems like open data portals.

Fuente: https://en.wikipedia.org/wiki/opendatasoft

Links



https://www.opendatasoft.com/
https://www.opendatasoft.com/open-data-solutions/
https://github.com/opendatasoft

ESCOLA CIUTADANA DE DADES

OPENDATASOFT :Portal Global

OpenDataSoft | Explore | Cartograph | API | Login

1,905 datasets

Sort by Last modification

Filters

Find a dataset...

View

Analyze	1,671
Map	799
Image	10
Calendar	5
Custom view	4

Modified

2011	36
2012	35
2013	108
2014	226
2015	473
2016	952

› *More*

Publisher

Agences Régionales de Santé (ARS)	253
Municipalidad de San Isidro	165
Quandl	132
INSEE	102
City of Portland, Oregon	99
Atelier Parisien d'Urbanisme	69

History of the SuperBowl

The Super Bowl is an annual American football game that determines the champion of the National Football League (NFL). The game culminates a season that begins in the previous calendar year, and is the conclusion of the NFL playoffs. The contest is held in an American city, chosen three to four years beforehand, usually at warm-weather sites or domed stadiums. Since January 1971, the winner of ...

Publisher data.world
License CC BY

Super | Bowl | NFL | Football

Table | Analyze | Export | API

Chicago Sunrise Data

This dataset contains colors and brightness levels recorded during a sunrise in Chicago on October 3rd, 2016. Brightness was recorded as luminosity (lux), with the colors of the sky and the water recorded as RGB values. Color names are also included for a large portion of the dataset. The dataset was collected as part of a joint class between the School of the Art Institute of Chicago and North...

Publisher Valentino Constantinou
License Public Domain

Sunrise | Chicago

Table | Analyze | Export | API

Commodity Prices

Time series of major commodity prices and indices including iron, cooper, wheat, gold, oil. Data comes from the International Monetary Fund (IMF).

Publisher Open Knowledge Foundation
License http://www.imf.org/external/terms.htm

Price | Index | Commodity

Table | Analyze | Export | API

Global Temperature Time Series

Global Temperature Time Series. Data are included from the GISS Surface Temperature (GISTEMP) analysis and the global component of Climate at a Glance (GCAG).

Publisher Open Knowledge Foundation
License PDDL

GCAG | GISTEMP | Temperature

Table | Analyze | Export | API

Annual Consumer Price

Table | Analyze

Corruption Perceptions

Table | Analyze

https://public.opendatasoft.com/explore/?sort=modified/

OPENDATASOFT : API

OpenDataSoft | Explore | Cartograph | API | Login

V1 | V2 (beta)

Console | Documentation

Concepts Overview
Connexion and Authentication

Datasets API

Search
Lookup

Records API

Search
Download
Analyze
GeoCluster

Appendices

How to identify a Dataset ?
How to use facets ?
Facets in Datasets API
Sorting in Datasets API
Query language
Examples

Concepts Overview

You'll find below the main concepts that are used in the remaining of this documentation.

Dataset	A Dataset is a logical data entity. It contains a set of Records. It can be seen as a table in a relational database. A Dataset also contains a set of metadata that describes it further (for instance, the publication date, the ownership, tags, themes, ...). Thus, a Dataset is fully defined by the list of Fields of the Records it contains and by its metadata.
Record	A Record is simply a row of values associated with their Fields. It is similar to a row in an Excel spreadsheet.
Domain	A Domain contains users and Datasets and defines a set of services allowing to manage and access these objects (for instance, the search API, the exploration conso A Domain can be public or private. In the latter cas be granted access to one or several Domains.

Connection and Authentication

Access to Domain APIs can be either public or protected depe
access is protected, two solutions can be used to authenticate

- HTTP Basic Authentication using the user login and pas
- An API key, passing the key as a simple HTTP parameter
  page when connected.
  `http://<DOMAIN>/api/datasets/1.0/search/?apik`

Both HTTP and HTTPS may be used. When the call is authenti

dataset: chicago-sunrise-data
Dataset ID

q
Full-text query

lang
2-letters language code for linguistic text features

rows: 10
Number of rows in the result (default: 10)

start
Index of the first result to return (use for pagination)

sort
Sort expression (field or -field)

facet: time, location, water_name, sky_name
Name of facets to enable in the results

refine: Facet name


```
{
    "nhits": 357,
    "parameters": {
        "dataset": [
            "chicago-sunrise-data"
        ],
        "timezone": "UTC",
        "rows": 10,
        "format": "json",
        "facet": [
            "time",
            "location",
            "water_name",
            "sky_name"
        ]
    },
    "records": [
        {
            "datasetid": "chicago-sunrise-data",
            "recordid": "
            "fields": {
                "water_g": 146,
                "sky_r": 211,
                "illuminance_lux": 153,
                "water_b": 135,
                "water_name": "heathered gray",
                "sky_name": "pastel gray",
                "sky_b": 189,
                "sky_g": 205,
                "water_r": 153,
                "time": "4:41:00",
                "location": "nelson"
            },
            "record_timestamp": "2017-02-03T10:27:01+00:00"
```

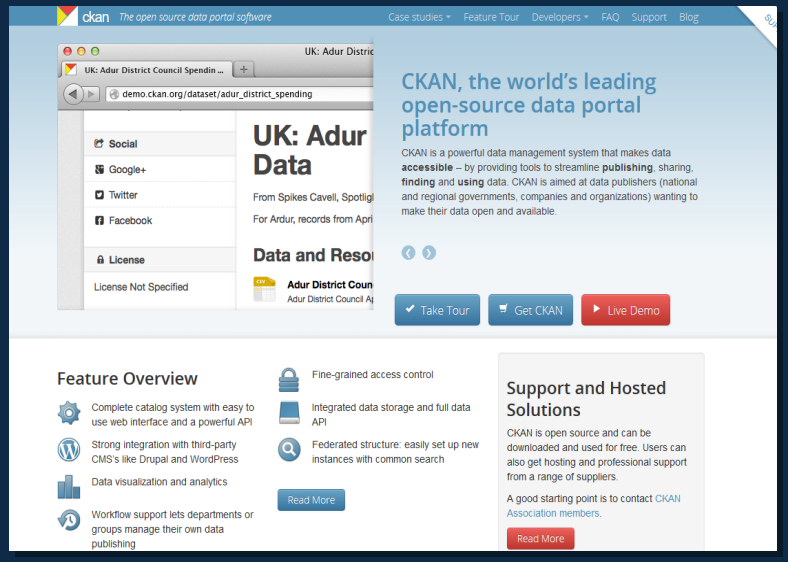

https://public.opendatasoft.com/api/v1/console/datasets/1.0/search/

CKAN

Sede: Cambrigde, UK

Fundación : 2004

Modelo negocio:
Fundación Open Knowledge International
Todo OpenSource pero también ofrecen servicio (de pago) de Hosting.

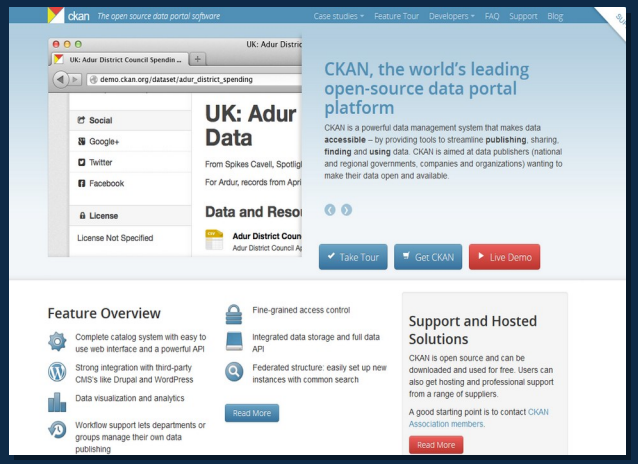


Open Knowledge International (OKI) (known as the Open Knowledge Foundation (OKF) until April 2014,[2] then Open Knowledge until May 2016[3]) is a global non-profit network that promotes and shares information at no charge, including both content and data.[4] It was founded by Rufus Pollock on 24 May 2004[5] in Cambridge, UK.

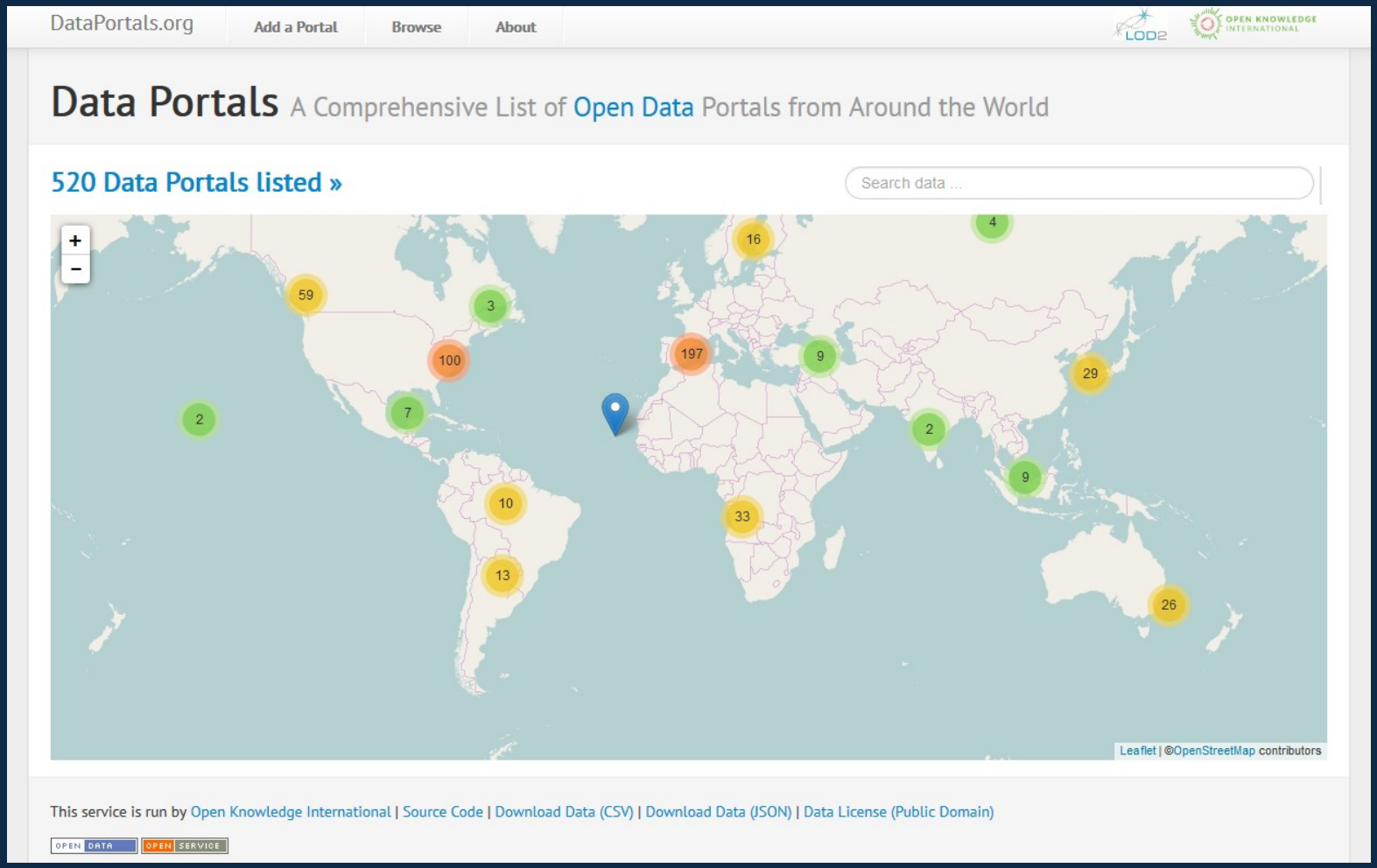
Fuente:
https://en.wikipedia.org/wiki/Open_Knowledge_International

Links

https://ckan.org
https://github.com/ckan/ckan
https://datahub.io



CKAN : Portal global



http://dataportals.org/

https://datahub.io/

CKAN : API

CKAN
latest

Search docs

User guide
Sysadmin guide
Maintainer's guide
API guide
Legacy APIs
Making an API request
Example: Importing datasets with the CKAN API
API versions
Authentication and API keys
GET-able API functions
JSONP support
API Examples
Action API reference
Extending guide
Theming guide
Contributing guide
Changelog

WRITE THE DOCS

Docs » API guide

Edit on GitHub

API guide

This section documents CKAN's API, for developers who want to write code that interacts with CKAN sites and their data.

CKAN's **Action API** is a powerful, RPC-style API that exposes all of CKAN's core features to API clients. All of a CKAN website's core functionality (everything you can do with the web interface and more) can be used by external code that calls the CKAN API. For example, using the CKAN API your app can:

- Get JSON-formatted lists of a site's datasets, groups or other CKAN objects:

  http://demo.ckan.org/api/3/action/package_list

  http://demo.ckan.org/api/3/action/group_list

  http://demo.ckan.org/api/3/action/tag_list
- Get a full JSON representation of a dataset, resource or other object:

  http://demo.ckan.org/api/3/action/package_show?id=adur_district_spending

  http://demo.ckan.org/api/3/action/tag_show?id=gold

  http://demo.ckan.org/api/3/action/group_show?id=data-explorer
- Search for packages or resources matching a query:

  http://demo.ckan.org/api/3/action/package_search?q=spending

  http://demo.ckan.org/api/3/action/resource_search?query=name:District%20Names
- Create, update and delete datasets, resources and other objects
- Get an activity stream of recently changed datasets on a site:

http://docs.ckan.org/en/latest/api/index.html

CKAN : API NIVELES

GLOBAL "API":
http://dataportals.org/api/data.json/

Domain "API":
http://demo.ckan.org/api/3/action/package_search?q=museu
http://opendata-ajuntament.barcelona.cat/data
/api/3/action/package_search?q=museu

http://demo.ckan.org/api/3/action/resource_search?query=description:Museu

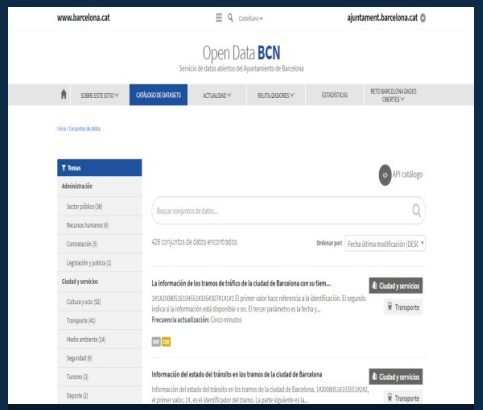
http://opendata-ajuntament.barcelona.cat/data/api/3/action/resource_search?query=description:barri

Dataset "API": (FileStore y DataStore):

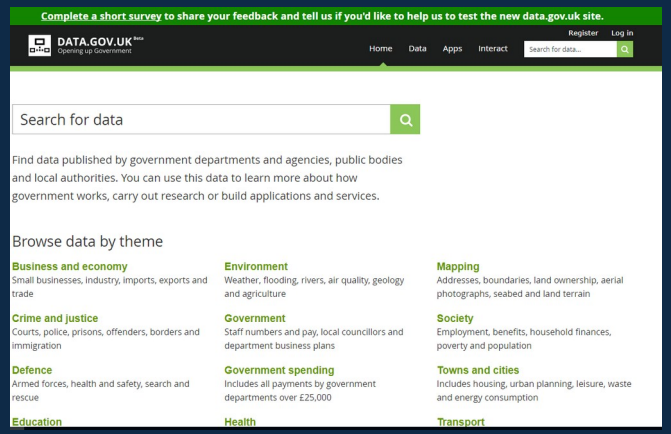
Create	http://demo.ckan.org/api/action/datastore_create
Update / Insert	http://demo.ckan.org/api/action/datastore_upsert
Query	http://demo.ckan.org/api/action/datastore_search
Query (via SQL)	http://demo.ckan.org/api/action/datastore_search_sql

Portales CKAN

http://opendata-ajuntament.barcelona.cat/data/es/dataset



https://data.gov.uk/



Conclusiones

- CKAN es OpenSource
- Puedo instalarlo en mi servidor o utilizar servicio web
- Puedo automatizar la gestión de datasets
- Todos los Datasets están expuestos via API