Overview of Open Data Portal Solutions

Khaled SELLAMI
Director General, e-government Unit, Tunisia
Expert in Open Government and Open Data

National Workshop in Palestine on Open Government
And Open Data
29-30 January 2020, Ramallah, Palestine

(Presented remotely)

اللجنة الاقتصادية والاجتماعية لغربي آسيا

الأمـــم المتحـــدة

ESCWA

Definitions

- A **Data Portal** is a Web-based system that contains a data catalogue with descriptions of datasets and provides services enabling discovery and re-use of the datasets.

- A **Dataset** is a collection of data, published or curated by a single source, and available for access or download in one or more formats.
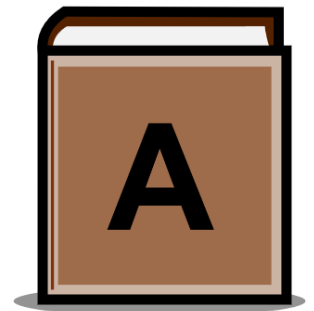
What is a data portal ?

Could mean :

- **Data Catalog**
  - **List of datasets**
  - Serach capacity
  - Display Metadata
  - Information on license
  - Link to each dataset
- **Platform**
  - Acess point to data catalog(s)
  - Offer a number of other services :
    - Technical assistance
    - Discussion forum
    - Documentation
    - Learning material
    - Blog
    - Examples of reuses

Questions about OD portals

- Which tools exist?

- What are the criteria ?

- what is the right tool for the right job?

- what type of data do they cover?

- What does the market for open data publishing tools look like?

- Is there a variety of commercial offerings?

- Are there open source communities?

Types of solutions for Open data portals

Two different types of solutions exist:

- **Open Source Software (OS):**

  - The source-code is available in an online and public repository under a liberal reuse license (such as the General Public License and its affiliates);

  - Sometimes known as Free and Open Source Software (FOSS),

  - Not all open source software is free, and not all free software is open source;

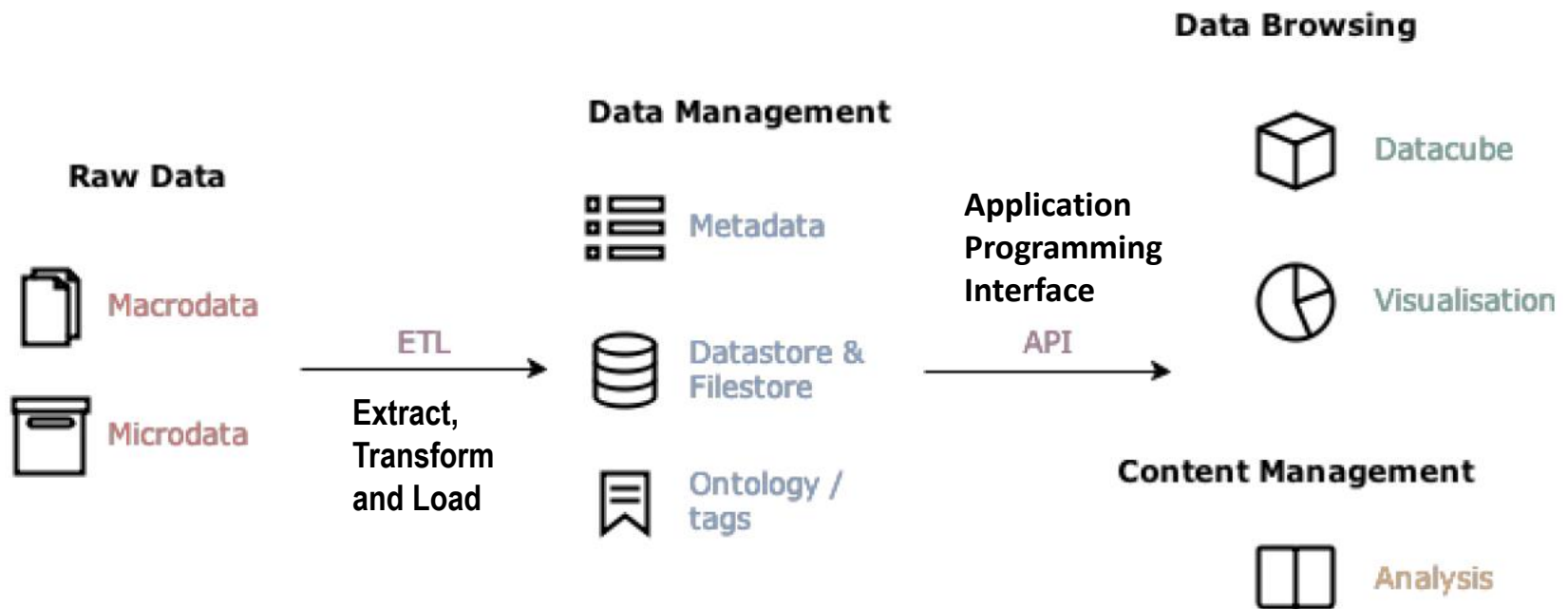 - Full customization and extensibility is guaranteed;

Types of solutions for Open data portals

- **Software-as-a-Service (SaaS):**

  - software is available online on a centralized hosted server via a subscription service instead of as a deployable software system with a single, static price;

  - upgrades, bug fixes and patches are consistently and regularly applied;

  - custom extensions of functionality can be achieved via an API but customization of the user interface is more limited;

Types of solutions for Open data portals

- Open source is also only useful if a developer community remains engaged and continues improving the software.

- The lifespan for open source software is often not different from proprietary solutions.

- Migration to different service provider A combination of standards compliance and open APIs can ensure the ability to even where the software is not open source.

- Between standards compliance (i.e. offering a straightforward mechanism to migrate your data to another service) and open source, go for standards compliance.

Visual representation of the various data and software components which form part of an overall open data platform :

Data Browsing

Data Management

Raw Data

Macrodata

Microdata

ETL

Extract, Transform and Load

Metadata

Datastore & Filestore

Ontology / tags

Application Programming Interface

API

Datacube

Visualisation

Content Management

Analysis

Assessment criteria for Open data platforms

Components	Open Data Systems
Descriptive metadata	●
Machine-readable	●
Anonymous access	●
Data licenses	●
Data attribution	●
Search	●
Open API	●
Static URI	●
Harvesting	●
Federating	●
Documentation	●
Standards-based	●
Structural metadata	statistical data publication
OLAP hypercube	statistical data publication
Data endpoints	•
Visualisation	•
UX & S/W extensibility	•

● is of primary importance   • serves a secondary, or partial, role

Source : World Bank

- These criteria cover mainly Data Discovery and Publication, as the core requirements for any open data distribution software concerns.

- Some data publication platforms offer more advanced criteries such as Extract, Transform and Load, Business Intelligence and Content Management functionality.

Assessment criteria for Open data platforms

o **Descriptive metadata:** External metadata typically used for discovery and identification. Information used to search and locate datasets. Ex. author, subjects, keywords, publisher.

o **Machine-readable:** Data available as machine-readable structured data in a non-proprietary format can be easily read and used by software systems without human interpretation.

o **Anonymous access:** Users can search for and access data and metadata without having to identify themselves, create a user account, or receive advance permission.

o **Data licences:** Terms of use associated with each dataset are clearly presented to the user and permit reuse and republication of that data in any alternative form.

Assessment criteria for Open data platforms

o **Data attribution:** Users can cite, attribute, and link to datasets, and contact data owners if they have questions.

o **Search:** Search results should return focused summaries on datasets, along with keywords which aid classification, and the option of reviewing the data online to assess its content.

o **Application Programming Interface (API):** Platforms make their contents available to external systems by supporting programmatic queries and access to metadata and resources.

o **Uniform Resource Identifiers (URI):** Platforms make datasets available at persistent URIs that never change, allowing them to be externally referenced reliably.

Assessment criteria for Open data platforms

o **Harvesting:** An automated and autonomous mechanism for ETL of known data from known web addressable locations into a single database or datastore.

o **Federation:** A meta-database management system, which transparently maps multiple autonomous database systems into a single federated database, allowing discrete data publishing systems to be integrated, yet operate independently.

o **Public Documentation:** Data platforms provide comprehensive information for developers and the general public on how their platform works. Such documentation should be updated with each new software release.

Assessment criteria for Open data platforms

o **Standards-based:** Platforms are consistent with emerging standards recognised by the W3C especially as regards metadata, RDF, and hypercubes.

o **Structural metadata:** Corresponds to internal metadata about the structure of database objects such as tables, columns, keys and indexes.

o **Online Analytical Processing (OLAP) hypercube or cube:** Supports the ability to analyse multidimensional data interactively from multiple perspectives, including the ability to consolidate, drill-down, or slice and dice data.

Assessment criteria for Open data platforms

o **Data Endpoints:** Provides structured data endpoints return data in predictable ways. These can be as simple as a known type of serialisation format while more complex implementations permit the data to be queried, filtering or refining the dataset prior to download.

o **Visualisation:** Provides tools to present data as common charts, maps or perform more complex statistical analysis.

o **UX & S/W Extensibility:** Permits sufficient template and layout customisation to provide a consistent user-experience and provide a common look and feel across all online services.

Descriptive metadata

- Used for discovery and identification, as well as for data life-cycle management.

- There are a large number of metadata vocabularies and ontologies used in open data systems.

- The most common:

  1. **Data Catalog Vocabulary (DCAT):** designed to facilitate interoperability of data catalogues published online;

  2. **Data Document Initiative (DDI):** describing the complete research data life-cycle for the social, behavioural and economic sciences;

  3. **Dublin Core Metadata Element Set:** fifteen "core" properties : title, creator, subject, description, publisher, date, format, source, etc.

- As well as permitting data and metadata upload, user-friendly interfaces are required in which system administrators can associate metadata with each data resource.

Descriptive metadata

Data about data:

- **Descriptive Metadata**
  - Help easily search and identify a data set
  - Title, summary, author, key words
  - External to data

- **Structural Metadata**
  - Allow understanding data and cross different data sets between them
  - Describe data structure (tables, columns, indexes)
  - Intern to data

- **Administratives metadata**
  - Facilitate data set maintenance
  - Describe publication context :
    - who? what? where? when? how?
  - Describes access and reuse rights

Examples of descriptive/administrative metadata

Property	Description	Data type	Multiple	Required
Name	Name given to the directory.	Short text string.	N	Y
Description	Brief narrative summary of the directory.	Text string.	N	Y
Homepage	The homepage associated to the directory	URI for the directory homepage.	N	Y
Language	The language or languages used in the textual metadata for describing items in the directory.	Language code as defined by BCP47[10] and recorded at the IANA registry[11].	Y	Y
Themes	The full set of subjects and topics that are addressed by the directory.	Predefined texts from the set of Subjects and Topics specified by the directory taxonomy[12].	Y	Y
Geographic coverage	The geographical area covered by the directory.	Predefined texts from the set of Territories specified by the directory taxonomy.	N	N
Publisher	The entity or entities responsible for making the directory online.	Text string or a reference to them.	Y	Y

Examples of data sets with metadata

https://catalog.data.gov/dataset/fruit-and-vegetable-prices

Fruit and Vegetable Prices

📅 Metadata Updated: February 7, 2017

How much do fruits and vegetables cost? ERS estimated average prices for 153 commonly consumed fresh and processed fruits and vegetables.

Access & Use Information

🌐 **Public:** This dataset is intended for public access and use.
📄 **License:** Creative Commons CCZero

Downloads & Resources

📊 **Web page with links to Excel files** 📈 4309 views
fruit-and-vegetable-prices.aspx
⊘ Link is ok ⭐⭐☆☆☆ Openness score

⬇ Download

Dates

Metadata Created Date	April 1, 2014
Metadata Updated Date	February 7, 2017

Metadata Source

📄 **Data.json Metadata**
Download Metadata

```
{
  "accessLevel": "public",
  "bureauCode": ["005:13"],
  "contactPoint": {
    "@type": "vcard:Contact",  "fn": "Hayden Stewart",
    "hasEmail": "mailto:hstewart@ers.usda.gov"
  },
  "describedBy": "https://www.ers.usda.gov/data-products/fruit-and-vegetable-prices/documentation.aspx",
  "description": "How much do fruits and vegetables cost? ERS estimated average prices for 153 commonly consumed fresh and processed fruits and vegetables.",
  "distribution": [{
    "@type": "dcat:Distribution",
    "downloadURL": "https://www.ers.usda.gov/data-products/fruit-and-vegetable-prices.aspx",
    "mediaType": "application/vnd.ms-excel",
    "title": "Web page with links to Excel files"
  }],
  "identifier": "USDA-ERS-00071",
  "issued": "2011-02-07",
  "modified": "2014-03-19",
  "keyword": ["agricultural economics", "food costs"],
  "license": "https://creativecommons.org/publicdomain/zero/1.0/",
  "programCode": ["005:041"],
  "publisher": {
    "@type": "org:Organization",
    "name": "Economic Research Service, U.S. Department of Agriculture"
  },
  "title": "Fruit and Vegetable Prices"
}
```

Examples of data sets with metadata

https://data.gov.uk/dataset/road-accidents-safety-data

Home / Datasets / Road Safety Data

Road Safety Data

Published by **Department for Transport**. Licensed under **OGL** Open Government Licence.
Openness rating: ★★★☆☆   Open Data Certificate: Bronze Level

These files provide detailed road safety data about the circumstances of personal injury road accidents in GB from 1979, the types (including Make and Model) of vehicles involved and the consequential casualties. The statistics relate only to personal injury accidents on public roads that are reported to the police, and subsequently recorded, using the STATS19 accident reporting form.

All the data variables are coded rather than containing textual strings. The lookup tables are available in the "Additional resources" section towards the bottom of the table.

Please note that the 2015 data were revised on the 29th September 2016.

Accident, Vehicle and Casualty data for 2005 - 2009 are available in the time series files under 2014. Data for 1979 - 2004 are available as a single download under 2004 below.

Also includes: Results of breath-test screening data from recently introduced digital breath testing devices, as provided by Police Authorities in England and Wales

Results of blood alcohol levels (milligrams / 100 millilitres of blood) provided by matching coroners' data (provided by Coroners in England and Wales and by Procurators...

�Read More⋎

Transport

ADDITIONAL INFORMATION

Added to data.gov.uk	20/09/2011
Theme	Transport
Mandate	https://www.gov.uk/government/publications/annual-data-requirement-fro http://www.legislation.gov.uk/ukpga/2011/13/section/92/enacted
Temporal coverage	1/1/1979 - 31/12/2016
Geographic coverage	Great Britain (England, Scotland, Wales)
Schema/Vocabulary	STATS19 Road accident with injury (police)
Update frequency	annual
Themes (secondary)	No value
Code list	No value
Service Level	No value

DATA RESOURCES (48 IN A TIME SERIES)

2016 View Less ∧

ZIP — 2016 Road Safety - Vehicles by Make and Model 2016
Details   Download ZIP (4.9 MB)

ZIP/CSV — 2016 Road Safety Data - Casualties 2016
Details   Download ZIP/CSV (1.2...

ZIP/CSV — 2016 Road Safety data -Vehicles 2016
Details   Download ZIP/CSV (2.7...

ZIP/CSV — 2016 Road Safety Data - Accidents 2016
Details   Download ZIP/CSV (4.7...

Examples of data sets with metadata

http://www.data.gouv.fr/fr/datasets/emissions-de-co2-et-de-polluants-des-vehicules-commercialises-en-france/
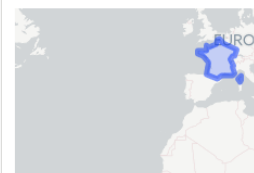
Examples of data sets with metadata (data.gov)

County-level Data Sets

🗓 Metadata Updated: May 2, 2019

Socioeconomic indicators like the poverty rate, population change, unemployment rate, and education levels vary across the nation. ERS has compiled the latest data on these measures into a mapping and data display/download application that allows users to identify and compare States and counties on these indicators.

Access & Use Information

🌐 Public: This dataset is intended for public access and use.
📄 License: Creative Commons CCZero

Downloads & Resources

Poverty 📈 1555 views
Query tool
[🔗 Visit page]

Web page with links to Excel files 📈 1945 views
[Open With ▾] [⬇ Download]

✅ Link is ok ⭐⭐☆☆☆ Openness score

References

https://www.ers.usda.gov/data-products/county-level-data-sets/documentation/

Dates

Metadata Created Date	April 1, 2014
Metadata Updated Date	May 2, 2019
Data Update Frequency	R/P1Y

Metadata Source

Data.json Metadata
Download Metadata

Harvested from USDA JSON

Additional Metadata

Resource Type	Dataset
Metadata Created Date	April 1, 2014
Metadata Updated Date	May 2, 2019
Publisher	Economic Research Service, Department of Agriculture
Unique Identifier	USDA-ERS-26471
Maintainer	Timothy Parker
Maintainer Email	tparker@ers.usda.gov
Public Access Level	public
Data Update Frequency	R/P1Y
Bureau Code	005:13
Metadata Context	https://project-open-data.cio.gov/v1.1/schema/catalog.jsonld
Schema Version	https://project-open-data.cio.gov/v1.1/schema
Catalog Describedby	https://project-open-data.cio.gov/v1.1/schema/catalog.json
Data Dictionary	http://www.ers.usda.gov/data-products/county-level-data-sets/documentation.aspx
Harvest Object Id	2b350aa4-4d94-4cf9-8c6f-b0c73b2f7ce0
Harvest Source Id	50ca39af-9ddb-466d-8cf3-84d67a204346
Harvest Source Title	USDA JSON
Data First Published	2012-06-28
License	https://creativecommons.org/publicdomain/zero/1.0/
Data Last Modified	2014-03-19
Program Code	005:041
Related Documents	https://www.ers.usda.gov/data-products/county-level-data-sets/documentation/
Source Datajson Identifier	True
Source Hash	0a369d8cd86e5e1d008732bc44b2cbad7b171ed0
Source Schema Version	1.1
Spatial	United States
Hide	

Didn't find what you're looking for? Suggest a dataset here.

Descriptive metadata

- **Metadata Vocabulary**
  - DCAT – http://www.w3.org/TR/vocab-dcat/
  - Dublin Core – http://dublincore.org/
  - ICalendar – http://en.wikipedia.org/wiki/ICalendar
  - …

- Using Metadata vocabulary to describe datasets in data catalogs will increase discoverability and enable applications easily to consume metadata from multiple catalogs.

- It further enables **decentralized publishing** of catalogs and facilitates **federated dataset** search across sites.

- **DCAT** is an RDF (Resource Description Framework) vocabulary designed to facilitate interoperability between data catalogs published on the Web.

Machine-readable datasets

- Extracting data from its source, transforming it into the required machine-readable format, and loading it into a database or Datastore. (e.g. Excel) can easily be converted to a **comma-separated-value** (CSV) file.

- Tim Berners-Lee 5 stars data quality:

  ★   Available on the web (in any format) but with an open licence;

  ★★   Available as machine-readable structured data in a proprietary format (e.g. Excel instead of image scan of a table);

  ★★★   Available as machine-readable structured data but in a non-proprietary format (e.g. CSV instead of Excel);

  ★★★★   All the above plus standards from W3C (RDF and SPARQL) to identify data uniquely so that others can access and reference this live data;

  ★★★★★ All the above plus cross-link data to other data to provide context;

Anonymous access

- users can search for and access data and metadata without having to identify themselves, create a user account, or receive advance permission.

- Data managers often want to know the profile of users and what they use data for : both would not be addressed by user-registration.

- Creating opportunities for engagement, and making the process interactive will make users tell themselves what they're doing with the data.

Data reuse licenses

- Data publication software must offer a mechanism by which the license associated with each dataset is clearly presented to the user.

- Datasets should be clearly labelled as released under standard open data licenses such:
  - **Creative Commons (CC-By, CC-0)**,
  - **Open Government Licence (OGL)**,
  - **Open Database Licence (ODbL)**,
  - **Open Intergovernmental Organisation License (IGO)**

- Licenses which permit data discovery but not liberal reuse are all but useless.

Different éléments of a license
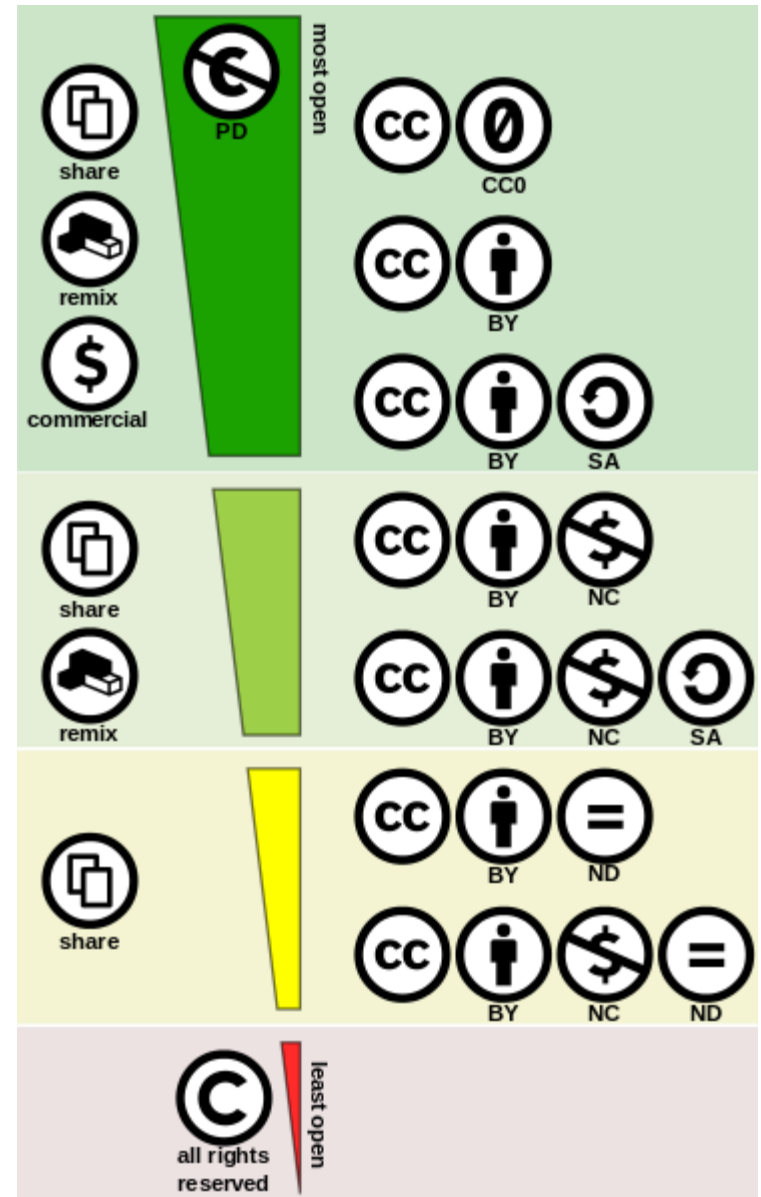
o Public Domain

o Possible restrictions

   o Attribution : producer, date

   o Share Alike : use same license

   o Non commercial

   o No Derivatives
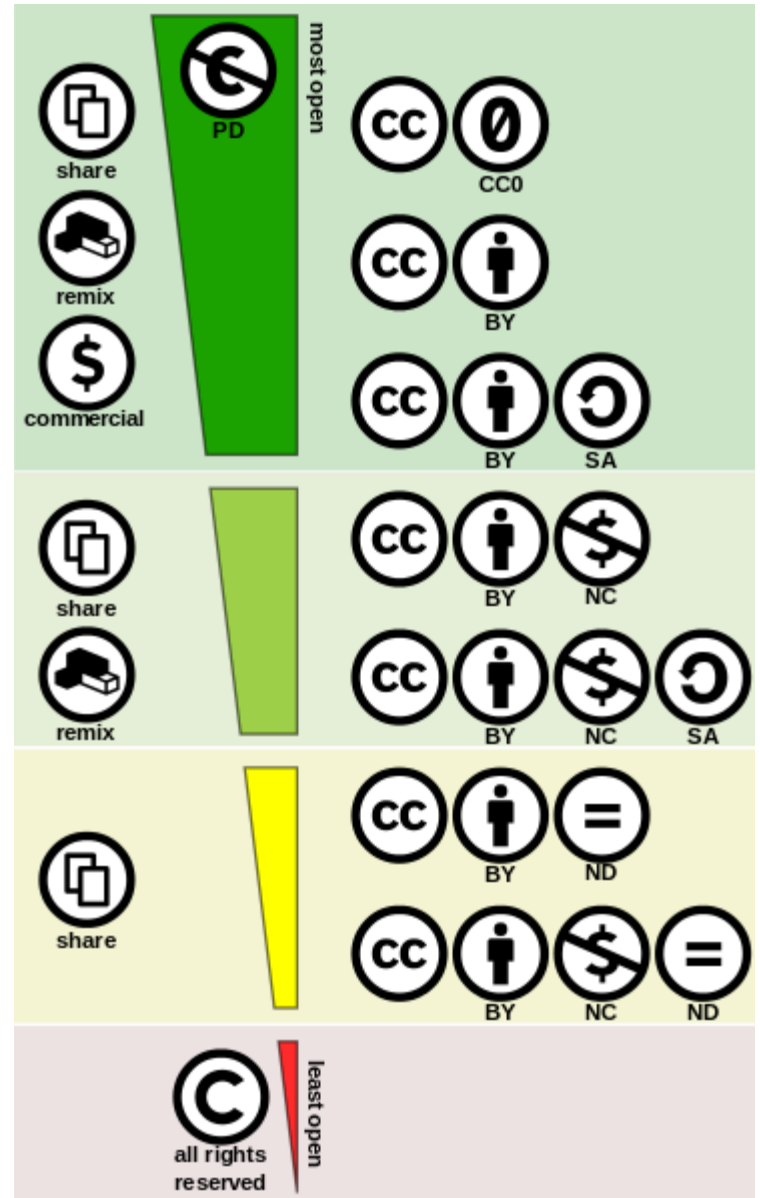
Popular licenses

- Creative Commons
- Open Data Commons
  - Open Database Licence: CC-by-SA for DB
  - Attribution
  - Public Domain
- Specific to certain countries:
  - Open Government Licences (UK)
  - Licence Ouverte (France)

Data attribution to source

Any dataset should present a clear set of information which permits a user to:

- link directly to the data;

- cite the data in their reuse of that data;

- attribute the data creator, either as an individual or as an organization;

- contact a data owner should they have any queries;

Search for data

- Search engine : find relevant data quickly and easily, then access that data in a useable format for the wide range of research activities which they may wish to perform.

- Minimum number of steps to find appropriate data, verify the data, and then access that data in a format which permits them to use it.

- Search results return focused summaries on datasets, along with keywords which aid classification, and the option of reviewing the data

Example of Search

Open API

- Interoperability requires an **Application Programming Interface (API)**

- APIs provide standardized commands to an external system to query the data, metadata and other attributes in a database.

- APIs can also be used by site administrators to automate data harvesting, uploading or similar processes.

- APIs need to be made available to public

Live roadworks API

Published by:	Leeds City Council
Last updated:	24 November 2018
Topic:	Not added
Licence:	[Open Government Licence](#)

Summary

This link provides access to the live data held behind
www.roadworks.org. Information such as live roadworks and gritting
routes can be found here.

More from this publisher

[All datasets from Leeds City Council](#)

Related datasets

[Live car park spaces API](#)

[Live Flood Warnings and Alerts API Data](#)

[Bradford gritting](#)

[Planned road works on the HE road network](#)

Search

Data links

Link to the data	Format	File added	Data preview
[Live Roadworks API](#)	TEXT/HTML; CHARSET=UTF-8	24 November 2018	Not available

CancelPush

Method used for cancelling a registr...
http://schemas.elgin.gov.uk/sdep/w...   ...essage that states whether
the cancellation was successful.

Test

The test form is only available for

SOAP 1.1

The following is a sample SOAP 1...

```
POST /sdep-publication/v5.2/sdep.asmx HTTP/1.1
Host: services.roadworks.org
Content-Type: text/xml; charset=utf-8
Content-Length: length
SOAPAction: "http://schemas.elgin.gov.uk/sdep/webservice/CancelPush"

<?xml version="1.0" encoding="utf-8"?>
<soap:Envelope xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Header>
    <Authentication xmlns="http://schemas.elgin.gov.uk/sdep/webservice">
      <User>string</User>
      <Password>string</Password>
    </Authentication>
  </soap:Header>
  <soap:Body>
    <CancelPush xmlns="http://schemas.elgin.gov.uk/sdep/webservice">
      <CancelPushParameter xmlns="http://schemas.elgin.gov.uk/sdep/roadevents">
        <ServiceURL>string</ServiceURL>
      </CancelPushParameter>
    </CancelPush>
  </soap:Body>
</soap:Envelope>
```

```
HTTP/1.1 200 OK
Content-Type: text/xml; charset=utf-8
Content-Length: length

<?xml version="1.0" encoding="utf-8"?>
<soap:Envelope xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Body>
    <CancelPushResponse xmlns="http://schemas.elgin.gov.uk/sdep/webservice">
      <CancelPushReceipt Processed="boolean" xmlns="http://schemas.elgin.gov.uk/sdep/roadevents" />
    </CancelPushResponse>
  </soap:Body>
</soap:Envelope>
```

Datasets are reachable via persistent URI

- Uniform Resource Identifiers (URIs) which ensure that resources, content or data are always located through one address for any and all users or software-driven applications.

- The most familiar URI are the Uniform Resource Locators (URLs) or links to websites

- Software should be capable of providing a clear and easy to find permanent URLs for every dataset being served by the platform.

Additional Metadata

Resource Type	Dataset
Metadata Created Date	April 1, 2014
Metadata Updated Date	May 2, 2019
Publisher	Economic Research Service, Department of Agriculture
Unique Identifier	USDA-ERS-26471
Maintainer	Timothy Parker
Maintainer Email	tparker@ers.usda.gov
Public Access Level	public
Data Update Frequency	R/P1Y
Bureau Code	005:13
Metadata Context	https://project-open-data.cio.gov/v1.1/schema/catalog.jsonld
Schema Version	https://project-open-data.cio.gov/v1.1/schema
Catalog Describedby	https://project-open-data.cio.gov/v1.1/schema/catalog.json
Data Dictionary	http://www.ers.usda.gov/data-products/county-level-data-sets/documentation.aspx
Harvest Object Id	2b350aa4-4d94-4cf9-8c6f-b0c73b2f7ce0
Harvest Source Id	50ca39af-9ddb-466d-8cf3-84d67a204346
Harvest Source Title	USDA JSON
Data First Published	2012-06-28
License	https://creativecommons.org/publicdomain/zero/1.0/
Data Last Modified	2014-03-19
Program Code	005:041
Related Documents	https://www.ers.usda.gov/data-products/county-level-data-sets/documentation/
Source Datajson Identifier	True
Source Hash	0a369d8cd86e5e1d008732bc44b2cbad7b171ed0
Source Schema Version	1.1
Spatial	United States
Hide	

Didn't find what you're looking for? Suggest a dataset here.

Automated data harvesting

- The manual process of creating datasets, entering metadata, and uploading data, is slow and labour-intensive.

- Some software systems allow set up and manage a large number of automated processes for data upload. Such as service is known as data **harvesting**.

- Some data, however, are updated regularly as part of a data release cycle. Datasets already exist and merely need to be updated.

Federation of multiple data sites

- **Federation** is the mechanism by which dataset metadata are polled from different platforms and copied to a centralized software service or database.

- The original data usually continues to be hosted on the original platform but the metadata, and links to the data resources, are now accessible and discoverable via the platform's search engine.

- Different departments may wish to manage their own data life-cycle. Local authorities often enjoy significant autonomy and they will be used to operating their own systems.

- Not all software which permits harvesting will support federation. Federated data captures only the metadata.
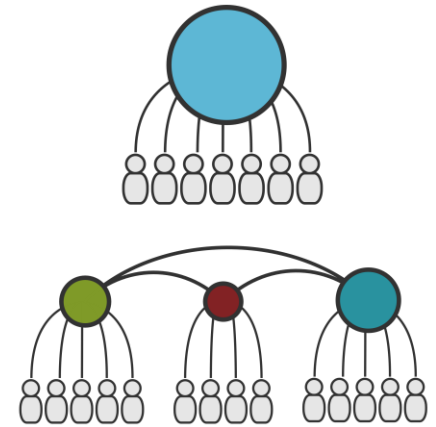
Public documentation

- Even experienced analysts will struggle in the absence of clear documentation.

- Software documentation is the responsibility of the platform vendor.

- Such documentation should be updated with each new software release.

Compliance with generally accepted standards

- Open data systems are still emerging as best-practice is agreed.

- The leading open data software systems tend to have similar approaches to metadata and in data presentation.

- Software which deviates from those approaches is more likely struggle with interoperability and general compliance.

- Many of those approaches are becoming standards recognised by the W3C especially as regards metadata.

Software supporting online data publication

Software platform	Descriptive Metadata	Machine-readable	Anonymous access	Data licences	Data attribution	Search	Open API	Static URI	Harvesting	Federating	Public Documentation	Standards-based	Structural Metadata	OLAP Hypercubes	Data Endpoints	Visualisation	UX & S/w	Extensibility
CKAN	●	●	●	●	●	●	●	●	•	•	●	●			•	•	●	
DevInfo	●	●	●	●	●	●	●	●			●		●	●	•	●	•	
DKAN	●	●	●	●	●	●	●	●	•	•	●				•	•	●	
Junar	●	●	●	●	●	●	●	●	•	•	●	●				•	●	●
NADA	●		•	●	●	●		●			•	●	●					●
Nesstar	●	●	●	•	●		●			•				●			●	•
OpenDataSoft	●	●	●	●	●	●	●	●	•		●				●	●	●	●
PC-Axis and PX-Web	●	●	●	•	●	•	●	●	•	•	•	•	●	●	●	●	●	●
Prognoz	●	●	●	•	●	●	●				●			●		●	●	•
Semantic MediaWiki	●	●	●	●	●	●	●	●			●	●			•			●
Socrata	●	●	●	●	●	●	●	●	•	•	●	●			●	●	●	●
Swirrl	●	●	●	●	●		●	●	•	•	●	●			•			●

● offers a complete solution

• offers a partial, or incomplete, solution

Blue are open data plateforms

Source : World Bank

URL	http://www.ckan.org/
S/W Licence	Affero GPL, open source
Language	Python, Javascript
SaaS	http://ckanexpress.com/
Demo	http://demo.ckan.org/
Examples	data.gov
	data.surrey.ca
	publicdata.eu

CKAN is an open-source, data discovery system making data accessible and usable by streamlining publishing, sharing, finding and using data. As well as harvesting, cataloguing, and advanced searching, it can store data and provides rich data APIs, and simple visualization and exploration tools.

URL	http://www.nucivic.com/
S/W Licence	Open source, GNU GPL
Language	PHP, JavaScript
SaaS	http://nucivic.com/data
Demo	http://demo.getdkan.com/
Examples	https://healthdata.gov/
	abrepr.org
	www.offenedaten-koeln.de

DKAN is an open source solution built on Drupal, aligned with the CKAN data portal software.

DKAN is a data discovery system making data accessible and usable, as well as harvesting, cataloguing, and advanced searching, it can store data and provides rich data APIs, visualization and exploration tools.

DKAN is a distribution of Drupal and is a complete CMS offering comprehensive tools to manage content.

There is also the option of enterprise-level SaaS.

URL	http://www.junar.com/
S/W Licence	Proprietary
Language	Java, Django / Python
SaaS	http://www.junar.com/
Demo	http://www.junar.com/
Examples	data.sanjoseca.gov
	datosabiertos.gob.go.cr
	data.cityofsacramento.org

Junar is a specifically Software-as-a-service platform offering one of the leading open data platforms. The system is able to import and use a wide variety of data formats and, as with all SaaS systems, is useful to users looking **for rapid deployment** and the ability to develop and present insight from their data very rapidly.

Source : World Bank

opendatasoft

URL	http://www.opendatasoft.com/
S/W Licence	Proprietary
Language	
SaaS	http://www.opendatasoft.com/
Demo	http://public.opendatasoft.com/
Examples	opendata.brussels.be
	opendata.paris.fr
	data.sncf.com

OpenDataSoft offers a comprehensive suite of **open data and visualisation tools**. Their search functionality is straightforward with faceting / filtering and well-structured results listings . OpenDataSoft is one of the more sophisticated open data platforms and well designed to serve that need.

URL	http://www.socrata.com/
S/W Licence	Mixed proprietary and OS
Language	Scala, Javascript, Ruby
SaaS :	socrata.com/products/open-data-portal/
Demo	nycopendata.socrata.com
Examples	data.undp.org
	data.cityofchicago.org
	opendata.go.ke

Not open source, Socrata's Open Data Portal **SaaS** provides one of the more comprehensive open data services, with a range of extensions for dashboards, live reports and the ability to manipulate and update existing data live in the portal.
Beyond open data, they offer business intelligence and visualisation functionality permitting data visualisation, analysis and sharing via social media.

References :

- Technical Assessment of Open Data Platforms for National Statistical Organisations, The World Bank, 2014

- Data.gov

- Data.gov.uk

- Data.gov.fr

- https://creativecommons.org/

Thank You

Economic and Social Commission for Western Asia

UNITED NATIONS

ESCWA