



ODI Learning in Tanzania

Dr David Tarrant | @davetaz | The Open Data Institute

<http://tanzania.learndata.info>



Course aim

Build a solid foundation and experience in
publishing, consuming and building a
business in Open Data



Open data essentials

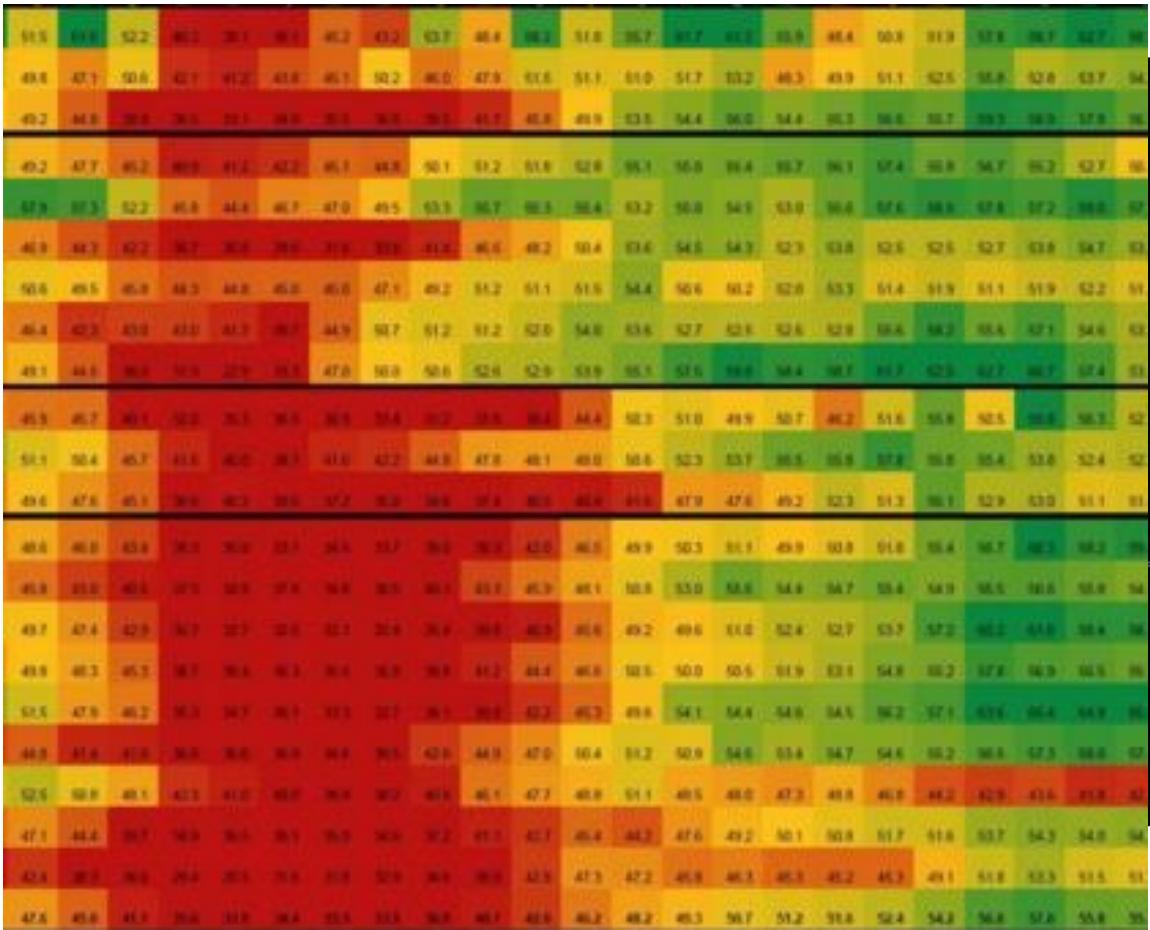
- ✓ Open data definitions
- ✓ Impact of open data
- ✓ Global impact
- ✓ Licensing
- ✓ Best practices
- ✓ Quality checking





Practical publication

- ✓ Data types
- ✓ Standards
- ✓ Platforms
- ✓ Publishing hands-on



Visulisation and communication

- ✓ Enriching data
- ✓ Filtering data
- ✓ Communication strategies
- ✓ Creating infographics

Today

Open data search patterns

Licensing

Platforms

Platforms and cleaning lab



Open data discovery patterns



Outcomes

Identify a number of different sources of open data on the web.

Create search patterns that enable easy discovery of new sources of open data.

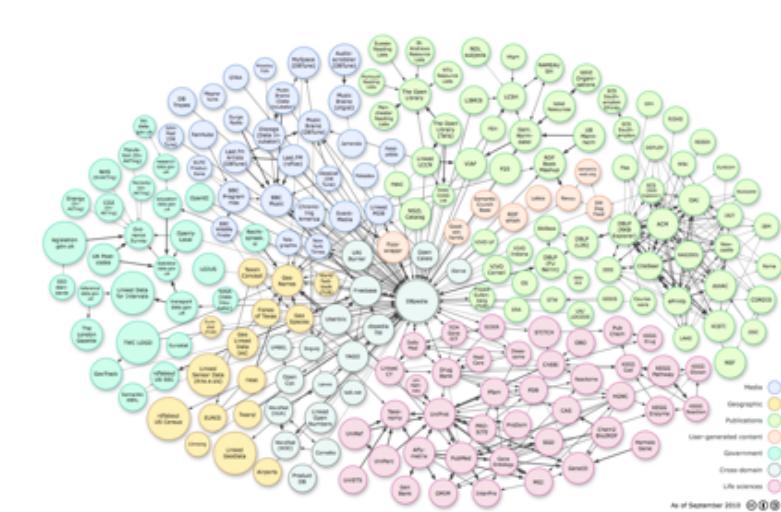


Approaches to publishing data

ON the web



IN the web



Finding data on the web (**of documents**)

Government data

Private sector data

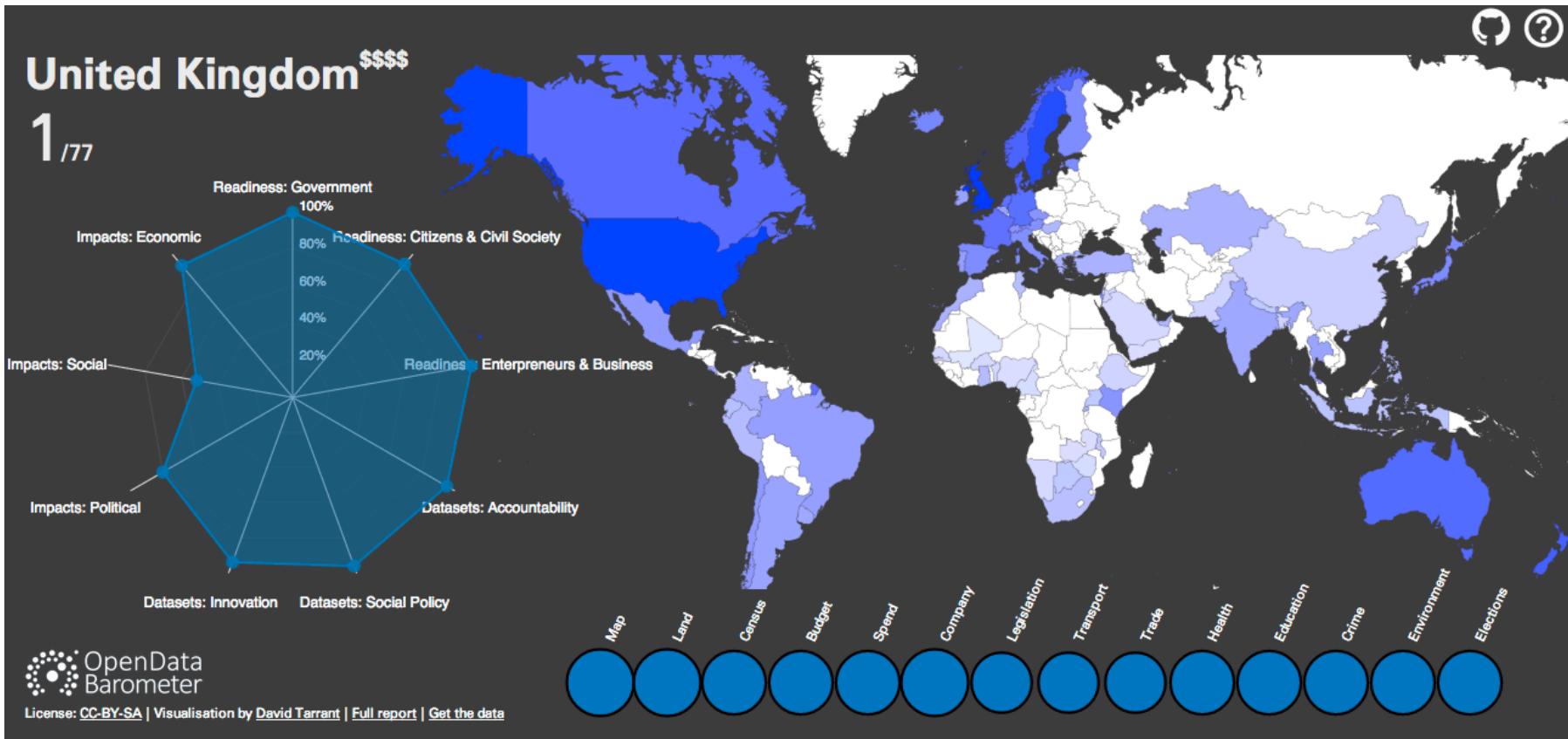
Google advanced

Aggregators and portals

Scraping



Government data



<http://www.opendatabarometer.org/>



data.gov.XX

The screenshot shows the DATA.GOV.UK homepage with a search bar and navigation menu. A large search result summary for "19266 Results" is displayed, with one item highlighted: "Live traffic information from the Highways Agency". Below this, there's a "SHOW ONLY..." section and a "Published datasets (15186)" link. The main content area features a green banner with "OpenData Burkina Faso" and a "THEMES" section listing various categories like Agriculture, Education, Diplomatie, Infrastructures, Eau et Environnement, TIC, Santé, Collectivités territoriales, Sécurité Publique, and Tourisme Culture.

The screenshot shows the DATOS.GOB.MX website. At the top, there's a header with the logo and the text "OPEN TO PARTICIPATE". Below it is a search bar with the placeholder "Buscar conjuntos de datos...". To the right of the search bar are buttons for "Datasets" and "Organizations". On the left, there's a map of the United States and Mexico with a "Filter by your suburb" button. The main content area displays a search result for "127 datasets found" under the heading "RATING SOCIAL PROGRAMS". The result details a database for evaluating social program results. The bottom section shows a news feed titled "Latest News and Events" with a message about no latest event.

Google advanced

The screenshot shows a Google search results page with the query "site:gov filetype:xls" entered into the search bar. The results are filtered to show only Microsoft Excel files from government websites. The first result is a link to "Code List or Concept (Acronym)" from www.acquisition.gov/short_codelistsTS.xls. The second result is a link to "Approps - Foreign Assistance.gov" from www.foreignassistance.gov/Full_ForeignAssistanceData.xls. The third result is a link to "TSB Monthly Cash Flow Projection" from www.dia.iowa.gov/tech/cashflow.xls. The search interface includes standard Google navigation links like Web, Images, Maps, Shopping, More, and Search tools, along with a note that approximately 4,150,000 results were found in 0.22 seconds.

site: Get results only from certain sites or domains

link: Find pages that link to a certain page

related: Find sites similar to one you already know

filetype: Find certain file types only

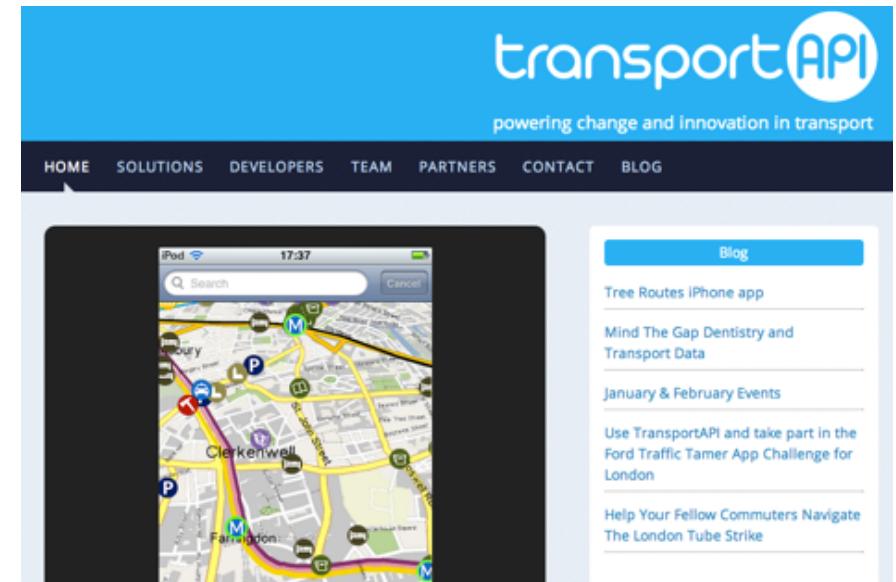


Aggregators and portals

Collect together data from across the web into one place.



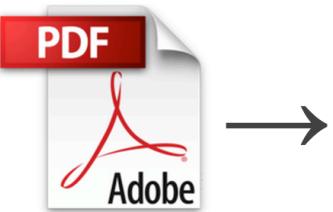
enigma.io



transportAPI

Scraping

If you can't obtain usable data (csv, xls) then you may have to resort to scraping.



Variety Name	Total Receipts	Total Receipts	Total Inedibles	Receipts
Aldrich	49,043,476	48,900,432	560,403	2.61%
Aviation	8,999,260	8,932,000	159,766	7.79%
Butte	144,865,828	147,182,233	1,186,402	7.70%
Butte/Padre	215,647,498	213,129,006	1,118,151	11.47%
Carmel	171,682,007	170,460,781	1,571,020	9.13%
Canyon	138,180,000	137,500,000	1,680,000	1.21%
Fritz	110,175,472	110,246,890	1,854,841	5.86%
Harvey	36,907	36,907	1,181	0.06%
Hashem	347,618	348,680	3,812	0.02%
LeGrand	9,797	9,557	261	0.26%
Livingston	7,394,450	7,500,200	159,210	0.39%
Marchini	418,442	418,442	9,096	0.02%
Mered	57,800	57,768	846	0.06%
Mission	17,265,270	17,174,884	78,967	0.92%

The image shows the import.io platform's user interface. At the top, there is a search bar with the placeholder "Enter a URL for a list page" and a pink "Extract Data" button. Below the search bar, there are four examples of scraped data:

- Reseller Ratings:** A screenshot of a web page showing a list of products with their ratings and reviews.
- Zoopla:** A screenshot of a real estate listing page from Zoopla.
- 500px:** A screenshot of a photo sharing website.

Below these examples, there are three more screenshots of different websites being scraped:

- Growth Hackers:** A screenshot of a blog or resource site.
- Udemy:** A screenshot of a course marketplace.
- Stack Exchange:** A screenshot of a Q&A website.

pdftables.com

magic.import.io

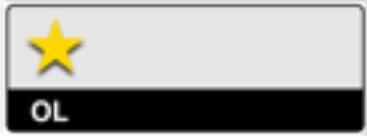


5-Stars



<http://5stardata.info/>

ON THE WEB



IN THE WEB

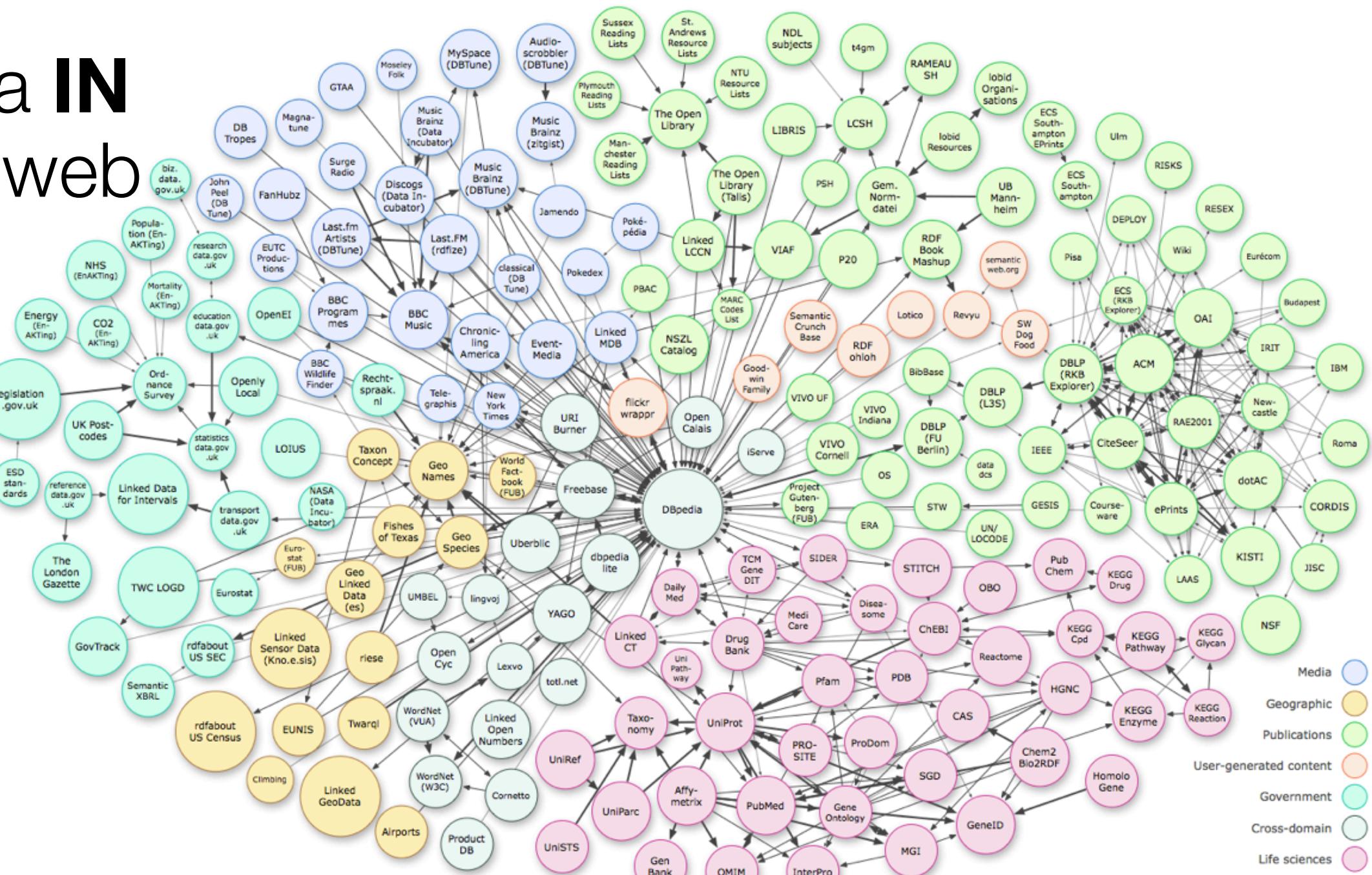


OPEN DATA



Content created by
The Open Data Institute

Data IN the web



Linked data

Amazing but challenging to publish and use

EEE Building

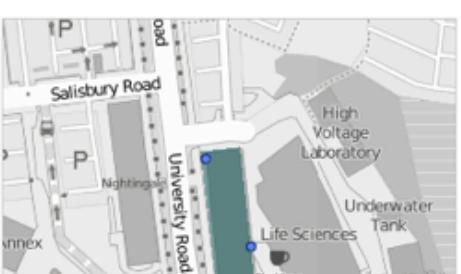
[http://id.southampton.ac.uk/building/32 ← This is the URI](http://id.southampton.ac.uk/building/32)

Detail	Facilities	Services	Energy
<p>Site: Highfield Campus</p> <p>Construction: 2006</p> <p>Architect: John McAslan & Partners</p> <p>Features: Building 32 is non-residential</p> <p>View Disability Report for this Building</p>			


©2010 Francois-Xavier Beckers (CC-BY)

Occupants

- Electronics & Computer Science
- Southampton Education School
- Agents, Interactions & Complexity
- Web & Internet Science
- Leadership School Improve &Effectiveness
- Lifelong & Work-Related Learning
- Mathematics & Science Education
- Social Justice & Inclusive Education
- Teaching Only Staff
- Deanery



Properties and Values

- soton.ac.uk/building/32 → rooms:Building, http://id.southampton.ac.uk/ns/UoSBuilding → "EEE Building"
- occupant → Electronics & Computer Science, Southampton Education School, Agents, Interactions Complexity, Web & Internet Science, ...show 8 more...
- construction → "32"^^http://id.southampton.ac.uk/ns/building-code-scheme
- locations:within → Highfield Campus
- <http://www.soton.ac.uk/estates/ourestate/buildings/highfield/32.html>
- "50.9364157"^^xsd:float
- "-1.395905"^^xsd:float
- southampton.ac.uk/ns/disabledGoPage → <http://www.disabledgo.com/en/access-guide/building-32>
- locations:easting → "442544"^^xsd:integer
- locations:northing → "115392"^^xsd:integer
- Organization → University of Southampton
- southampton.ac.uk/ns/ombielName → "Bldg 32 (EEE)"
- feature → Building 32 is non-residential
- southampton.ac.uk/ns/buildingDate → "2006"
- southampton.ac.uk/ns/buildingArchitect → John McAslan & Partners
- spatial → "POLYGON((-1.3961073411331264 50.93683868764933,-1.3958347895092957 50.9368567227702,-1.3956958407975968 50.936065737417,-1.3959558923017397 50.93603859197583,-1.3961073411331264 50.93683868764933))"
- southampton.ac.uk/ns/electricityTimeSeries → "elec/b32/ekw"
- ← is spatialrelations:within of ← 32 / 3077, 32 / 1015, Physical and Applied Science Faculty Deanery, Social and Human Sciences Faculty Deanery, ...show 54 more...
- ← is foaf:depicts of ← <http://data.southampton.ac.uk/image-archive/buildings/raw/32.jpg>, <http://data.southampton.ac.uk/image-archive/buildings/1000/32.jpg>, <http://data.southampton.ac.uk/image-archive/buildings/800/32.jpg>, ...show 5 more...
- ← is event:place of ← RaeS Solent Branch Christmas Special Lecture - The Red Arrows


a Building
Google



The Open Data Institute

Finding data on the web (**of data**)

1. Add random extensions (.xml, .json, .csv etc)
2. Look for alternative links (rss feeds etc)
3. Look for embedded data
4. Do some content negotiation
5. Spot the API
6. Scrape (or search google again)

IN THE WEB

How the web should work, but people forgot that Tim put this in when he invented it!



Content created by
The Open Data Institute

Duck typed data

If it looks like a duck
and quacks like a duck,
then it's probably a duck.

Basically, keep an eye out for tables,
lists and other stuff that looks like data.



1. Adding random extensions

The screenshot shows the GOV.UK Trade Tariff homepage. At the top, there's a navigation bar with the GOV.UK logo and a search bar. Below it, a breadcrumb trail reads "Home > Business and self-employed > Imports and exports". The main title "Trade Tariff" is displayed. A search form allows users to "Search the tariff" by name or code and includes a "Search" button. Below the search form, a note says "This tariff is for 6 August 2014" with a "change date" link. There are also links to "View all sections" and "A-Z Index". A table lists tariff sections from I to IX, each with a title:

Section	Chapters	Title
I	1 to 5	Live animals; animal products
II	6 to 14	Vegetable products
III	15	Animal or vegetable fats and oils and their cleavage products; prepared edible fats; animal or vegetable waxes
IV	16 to 24	Prepared foodstuffs; beverages, spirits and vinegar; tobacco and manufactured tobacco substitutes
V	25 to 27	Mineral products
VI	28 to 38	Products of the chemical or allied industries
VII	39 to 40	Plastics and articles thereof; rubber and articles thereof
VIII	41 to 43	Raw hides and skins, leather, furskins and articles thereof; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silkworm gut)
IX	44 to 46	Wood and articles of wood; wood charcoal; cork and articles of cork; manufactures of straw, of esparto or of other plaiting materials; basket-ware and wickerwork

UK Trade Tariff

Try using the following: .csv .json .xml .rss .rdf

The screenshot shows the BBC Doctor Who website. At the top, there's a navigation bar with the "one" BBC logo and a search bar. The main title "DOCTOR WHO" is prominently displayed with the TARDIS logo. Below the title, a menu bar offers links to "Home", "Episodes", "Clips", "Galleries", "Latest News", "Characters", "Monsters", "Fun and Games", and "More". The page features several promotional images and sections: "On iPlayer" (with a note about availability after broadcast), "It's Tomorrow... Get the Latest on the Launch!", and "On TV" (listing "The Day of the Doctor" for Saturday 19:00 on BBC THREE). There's also a "All upcoming (0 NEW AND 1 REPEAT)" section.

BBC Music and Programmes





2. Look for alternative links

The screenshot shows the NewsAsia website homepage from Wednesday, August 06, 2014. The top navigation bar includes 'Business Insight' (highlighted in red), 'NEWS', 'TV', 'WATCH LIVE', and various news categories like 'ASIA PACIFIC', 'SINGAPORE', 'WORLD', 'BUSINESS', 'SPORT', 'ENTERTAINMENT', 'TECHNOLOGY', 'HEALTH', 'LIFESTYLE', 'VIDEOS', 'WEATHER', and 'MORE'. A large blue box with the text 'Scroll down!' is overlaid on the left side. A large black arrow points downwards from this box towards the bottom of the page. The main content area features several news stories:

- 12% for Home Team officers, bonuses of up to S\$30,000** (with a photo of two men shaking hands)
- Pay rise, special bonus for about 23,000 nurses** (with a photo of a nurse)
- 50,000 openings on Jobs Bank for Singaporeans, PRs** (with a photo of a job application interface)
- NUS University Town Identified as a high-risk dengue cluster** (with a photo of a university building)

The sidebar on the left lists other news categories: 'SINGAPORE', 'WORLD', 'BUSINESS', 'SPORT', 'LIFESTYLE', 'HEADLINES', 'VIDEOS', and 'ENRICHMENT'.





2. Look for alternative links

 CHANNEL NEWSASIA MediaCorp News Group. © 2014 MediaCorp Pte Ltd. All Rights Reserved. Terms and Conditions Privacy Policy About MediaCorp Pte Ltd	NEWS Asia Pacific Singapore World Business Sport Entertainment Technology Health Lifestyle Videos Photos Special Reports Archives	TV Live TV TV Videos TV Schedule SERVICES Weather ADVERTISE WITH US Online Advertising Mobile Advertising TV Advertising Contact Sales	ABOUT US About Channel NewsAsia Our Logo Our Coverage Our Tagline Presenters and Correspondents Contact Us GET OUR NEWS  
---	---	---	---

RSS

3. Look for embedded data

The screenshot shows the 'Hidden data extractor' tool. At the top, there's a dark header bar with the text 'Hidden data extractor' in blue on the left and 'ODI Experiment' in orange on the right. To the right of the header is the ODI logo. Below the header is a light gray main area. In the center, it says 'Hidden data extractor' and 'Enter the URL of any webpage to see what JSON data is hidden within it.' There is a large input field for pasting a URL, followed by a blue 'Submit' button. Below the input field, under the heading 'Try these', are two links: 'Products from Marks and Spencer UK' and 'Products from ASOS'.

<http://odinprac.theodi.org/hidden-data-extractor/>



Finding data on the web (**of data**)

1. Add random extensions (.xml, .json, .csv etc)
2. Look for alternative links (rss feeds etc)
3. Look for embedded data
4. Do some content negotiation
5. Spot the API
6. Scrape (or search google again)

IN THE WEB

How the web should work, but people forgot that Tim put this in when he invented it!



Content created by
The Open Data Institute

Exercise

Find a data set using one of the routes we've just looked at.....

Ask yourself – (and discuss in groups)

Is it usable?

What makes it usable?

What more do you need to know?



Outcomes

Identify a number of different sources of open data on the web.

Create search patterns that enable easy discovery of new sources of open data.



Outcomes

Identify a number of different sources of open data on the web.

Create search patterns that enable easy discovery of new sources of open data.

Understand the difference between “data on the web” and the “web of data”.



Today

Open data search patterns

Licensing

Platforms

Value propositions



Increase confidence in using and understanding open data licences



List the characteristics of an open licence

Identify instances of an open licence

Identify when you can apply an open licence to a data set for publication



Law and licensing

Please note:

I am not a lawyer and this should not be treated as legal advice.



Why licence?

A licence provides clarity

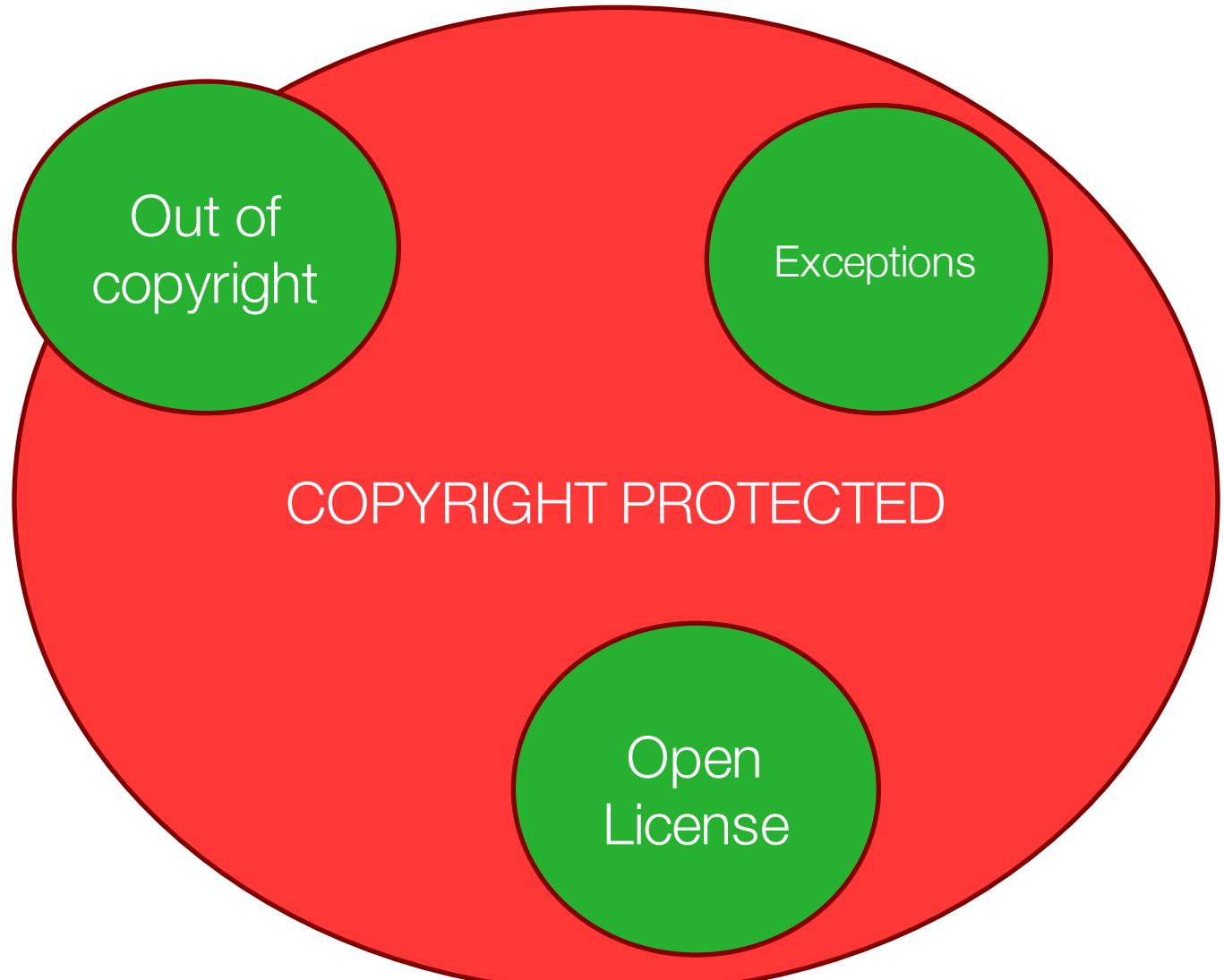
It sets out exactly what users and re-users are permitted to do with your dataset



The content world

You can use it if [green]:

- It's out of copyright [public domain]
- It's openly licensed
- Your use is covered by a copyright exception



Content created by
The Open Data Institute

Copyright

Applies automatically to –

- literary works
- artistic works
- dramatic works
- musical works
- layouts of works
- sound and film works
- broadcasts



Copyright

Applies automatically to –

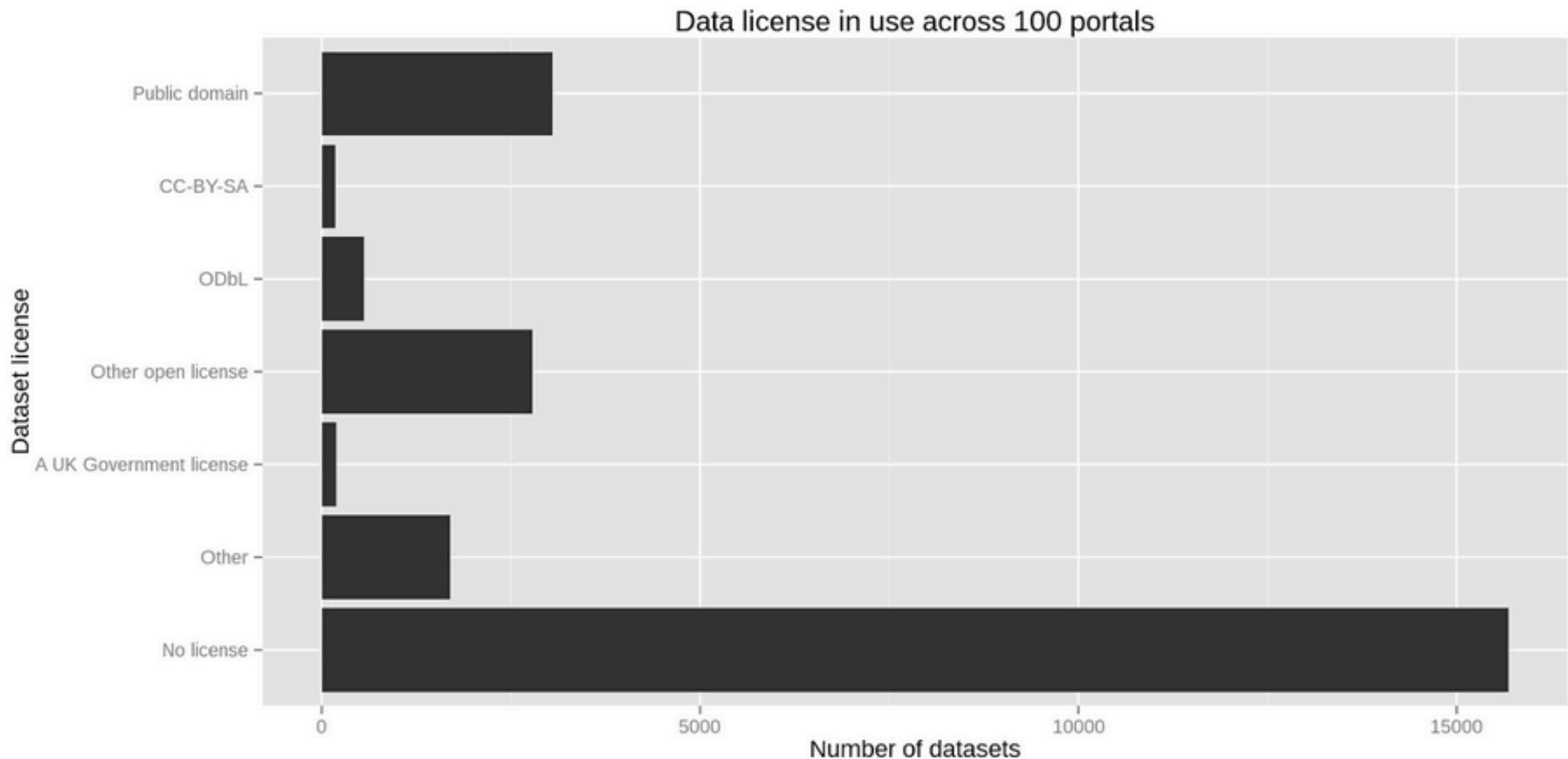
Anything where there is an intellectual, creative
and original process in creating a thing.

It does not apply when the thing created would be
the same whoever created it.

“it’s obvious”



Open data is hardly ever appropriately licensed.



Exercise

Build your own open data licence



Some drafting tips

“Under this licence **you can:**”

[what is permitted?]

“**You must**, where you do any of the above:”

[do you have any restrictions?]

“**You must not**”

[explicit restrictions?]



The UK Open Government Licence

Version 3.0 released in October 2014

[http://www.nationalarchives.gov.uk/doc/open-government-
licence/version/3/](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/)



What does it permit?



Using the Information under this licence

Use of copyright and database right material expressly made available under this licence (the 'Information') indicates your acceptance of the terms and conditions below.

The Licensor grants you a worldwide, royalty-free, perpetual, non-exclusive licence to use the Information subject to the conditions below.

This licence does not affect your freedom under fair dealing or fair use or any other copyright or database right exceptions and limitations.

You are free to:

- copy, publish, distribute and transmit the Information;
- adapt the Information;
- exploit the Information commercially and non-commercially for example, by combining it with other Information, or by including it in your own product or application.



Are there restrictions on its use?

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;

If the Information Provider does not provide a specific attribution statement, you must use the following:

Contains public sector information licensed under the Open Government Licence v3.0.

If you are using Information from several Information Providers and listing multiple attributions is not practical in your product or application, you may include a URI or hyperlink to a resource that contains the required attribution statements.

These are important conditions of this licence and if you fail to comply with them the rights granted to you under this licence, or any similar licence granted by the Licensor, will end automatically.



The main alternative



Open and community developed; not controlled by a central agency or body.



This is a human-readable summary of (and not a substitute for) the [license](#).

[Disclaimer](#)



You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material
for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:



Attribution — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.



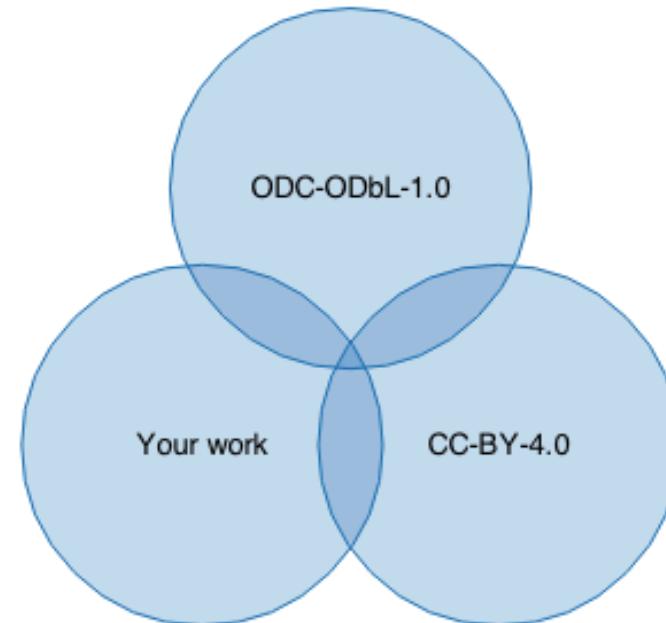
ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.

Checking license compatibility

Type of remix:

My new work is a:

- "Strong" adaptation. I remix or tightly integrate significant portions of the original works directly into my own.
- "Weak" adaptation of individual files. All of my changes and additions are contained in separate files from those of the original works.
- "Weak" adaptation of software libraries. My own work merely links against the original works, using them as software libraries. I've made no changes to the source code of the original works.
- Collection. I simply package the original works together, without changing any of the individual works themselves.



<http://clipol.org/tools/compatibility>



Identifying open licenses

Recognising Open Licences

When discovering data online, it is important to be able to check the licence to understand whether the data is open and if so, under what conditions can you access, use and share it

Take a look at each of the following licences and try to determine whether they are open licences



Licensing your data set for publication



What do you need to think about?

! Exemptions

This licence does not cover:

- personal data in the Information;
- Information that has not been accessed by way of publication or disclosure under information access legislation (including the Freedom of Information Acts for the UK and Scotland) by or with the consent of the Information Provider;
- departmental or public sector organisation logos, crests and the Royal Arms except where they form an integral part of a document or dataset;
- military insignia;
- third party rights the Information Provider is not authorised to license;
- other intellectual property rights, including patents, trade marks, and design rights; and
- identity documents such as the British Passport



Be Careful!

Personal data

Protected by the data protection act

Information that has been accessed unlawfully

You ripped it from a protected medium (DVD)

Third party IP rights

The data was obtained from a copyrighted source.



Personal Data

Does this data contain personal identifiable data?

Could this data be used combined with another data set to create personal identifiable data?

Can it be anonymised?



Third party rights

A data set could also contain other copyright works:

- Images (i.e. satellite pictures)
- Headlines (if original enough)
- Excerpts of text



List the characteristics of an open licence

Identify instances of an open licence

Identify when you can apply an open licence to a data set for publication



Open data publication platforms



Outcomes

Evaluate a number of different approaches for publishing open data.

Publish a dataset on one of the many available platforms.



Types

Specialist Solution

- + Easy to get setup and maintain.
- + Open Data focused
- + Clear workflows for publishing open data
- + Visualisation tools
- + Data mashing tools
- + Best for transactional data

Integrated Solution

- + No new platform to learn
- + Data is provided in parallel to web pages
- + No separation from authoritative data
- + Easy discovery of data
- + Best for reference data
- + Best for Linked Open Data



Key characteristics of specialist solution

Separate from your main org website

Designed to publish open data, not to fulfill other org goals

Examples: data.gov.uk | data.cityofchicago.org | data.sncf.com



Key characteristics of integrated solution

It is your main website

Publishes data alongside everything else that the organisation does

Examples: bbc.co.uk/programmes | southampton.ac.uk | gov.uk



Merging specialist and integrated

Method 1: *Build the functionality of your current website into a new open data platform.*

Example: [West Sussex Country Council](#)

Method 2: *Hide the specialist solution behind your main website and use it as a loosely coupled CMS.*

Example: [southampton.ac.uk](#)



The sliding scale of specialist solutions

1. Catalogue: Point to data (leave it at source)
2. Re-present: Provide data services (leave it at source)
3. Host the data: Be the source
4. Control the data: Be the authority
5. Be the hub: Host the data and processor



Specialist Solutions



comprehensive knowledge
archive network

<http://www.flickr.com/photos/okfn>



Content created by
The Open Data Institute



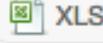
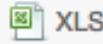
Overview

Licence

Open Government Licence **OGL**

OPEN DATA

Data Resources 6

-  Tariff index
-  Tariff data
-  Tariff data - overview
-  Tariff section index
-  Tariff section notes
-  Tariff chapter notes

<http://demo.ckan.org/>



Open Knowledge Foundation Supported

Data Catalogue

Open Source

Feels like a record manager

Simple API and search

Lots of community tools



Content created by
The Open Data Institute



Evolution of CKAN (v1.4)

Early) Dataset catalogue (data.gov.uk)

no data hosted or searched

Mid) Data and dataset catalogue

no data hosted but it is searchable

Now) Integrated data driven web site

data platform is integrated with data, search and content





Features

Publish, Store and Manage Data and Metadata

Visual and Geospatial

Social

Full Stored History

Federate Your Data With Other Organizations

Rich RESTful JSON API for Developers



Overview

Prix des carburants - J-7

Information Table Map Analyze Export API Share ▾

Nom	Marque	Carburant	Prix Gazole	Prix SP95	Prix SP98
1 MME POGGIO	Total	Gazole	1.35	1.56	1.63
2 RELAIS HAGONDANGE	Total	Gazole/E10	1.367		
3 AGIP CAVAILLON CLEMENCEAU	Agip	Gazole/E10/SP98	1.36		1.59
4 Carrefour Market	Carrefour Market	Gazole/SP95	1.337	1.589	1.58
5 INTERMARCHE RABASTENS	Intermarché	Gazole	1.317	1.519	1.539
6 DAINVILDIS	Leclerc	SP95		1.476	
7 LES BERJALLIENS	Total	Gazole/E10/SP98	1.36		1.59
8 Avia lecourbe	Avia	Gazole	1.46	1.7	1.72
9 Sarl STATION KLEBER	Elan	Gazole/SP95/SP98	1.79	1.94	1.96
10 BP PARIS PAUL DOUMER	BP	Gazole/E10	1.439		
11 RELAIS TOTAL MARECHAL DAVOUT	Total	Gazole/E10/SP98	1.365		1.648
12 ESSO DU MIN	Esso Express	E10	1.222		1.502
13 INTERMARCHE LILLEBONNE	Intermarché	Gazole/SP95/SP98	1.259	1.499	1.529
14 CARREFOUR GRUCHET	Carrefour	Gazole/SP95	1.189	1.429	1.529
15 SARL PICHON	Avia	Gazole/SP95/SP98	1.34	1.53	1.57



Data as a Service (DaaS)

Hosted enterprise solution

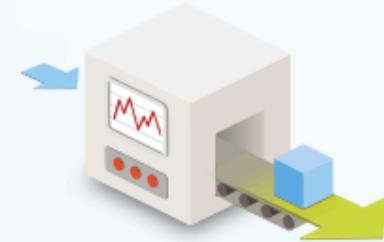
Query based API (3-Star)

EU Based

Closed source



Features



Prepare data

- Collect from any source
- Read every kind of formats
- Understand all data types
- Make datasets findable & reusable



Reuse data

- Full-featured secured API
- Standard access formats
- Interactive & shareable visualization
- Web extensions & open source



Scale services

- Big data technologies
- 100% cloud-based
- API factory
- Monitoring

Overview

The screenshot shows the Socrata platform interface. At the top is a search bar with a magnifying glass icon. Below it is a 'View Types' section with icons and labels: Datasets (blue), Charts (orange), Maps (green), Calendars (teal), Filtered Views (light blue), External Datasets (light green), Files and Documents (light orange), Forms (light purple), and APIs (pink). At the bottom is a navigation bar with icons for Manage (with a red '9' badge), More Views (with a red '9' badge), Filter, Visualize, Export, Discuss, Embed, and About.



Data as a Service (DaaS)

Hosted enterprise solution

Allows user created content

Full linked API (5-Star)

US Based

Closed source



Features



Data Publishing, Optimized for Business Users

Flexible Metadata Management

Federate Your Data With Other Organizations

Metrics of the Success of Your Initiative in Real-time

Anyone Can Create **Maps and Charts**

Data Becomes **Social**

Developers Are Supported Every Step of the Way



The sliding scale of specialist solutions

1. Catalogue: Point to data (leave it at source)



2. Re-present: Provide data services (leave it at source)



3. Host the data: Be the source

4. Control the data: Be the authority



5. Be the hub: Host the data and processor

Integrated solutions

Integrated solutions expose data using the current infrastructure (web pages).

Data driven web site

Best for reference and live data



The developers secret



[Home](#) > [Business and self-employed](#) > [Imports and exports](#)

Trade Tariff

Search the tariff [Search](#)

This tariff is for 6 August 2014 [change date](#)

[View all sections](#) [A-Z Index](#)

Trade between the UK and All countries [change country](#)

Section	Chapters	Title
I	1 to 5	Live animals; animal products
II	6 to 14	Vegetable products
III	15	Animal or vegetable fats and oils and their cleavage products; prepared edible fats; animal or vegetable waxes
IV	16 to 24	Prepared foodstuffs; beverages, spirits and vinegar; tobacco and manufactured tobacco substitutes
V	25 to 27	Mineral products
VI	28 to 38	Products of the chemical or allied industries
VII	39 to 40	Plastics and articles thereof; rubber and articles thereof
VIII	41 to 43	Raw hides and skins, leather, furskins and articles thereof; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silkworm gut)
IX	44 to 46	Wood and articles of wood; wood charcoal; cork and articles of cork; manufactures of straw, of esparto or of other plaiting materials; basket-ware and wickerwork

```
--  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
  
{"id": 1,  
 "position": 1,  
 "title": "Live animals; animal products",  
 "numeral": "I",  
 "chapter_from": "01",  
 "chapter_to": "05",  
 "section_note_id": 1}  
  
{"id": 2,  
 "position": 2,  
 "title": "Vegetable products",  
 "numeral": "II",  
 "chapter_from": "06",  
 "chapter_to": "14",  
 "section_note_id": 6}  
  
{"id": 3,  
 "position": 3,  
 "title": "Animal or vegetable fats and oils and their cleavage products; prepared edible fats; animal or vegetable waxes",  
 "numeral": "III",  
 "chapter_from": "15",  
 "chapter_to": "15",  
 "section_note_id": null},  
{  
 "id": 4,  
 "position": 4,  
 "title": "Prepared foodstuffs; beverages, spirits and vinegar; tobacco and manufactured tobacco substitutes",  
 "numeral": "IV",  
 "chapter_from": "16",  
 "chapter_to": "24",  
 "section_note_id": null},  
{  
 "id": 5,  
 "position": 5,  
 "title": "Mineral products",  
 "numeral": "V",  
 "chapter_from": "25",  
 "chapter_to": "27",  
 "section_note_id": null},  
{  
 "id": 6,  
 "position": 6,  
 "title": "Products of the chemical or allied industries",  
 "numeral": "VI",  
 "chapter_from": "28",  
 "chapter_to": "38",  
 "section_note_id": null},  
{  
 "id": 7,  
 "position": 7,  
 "title": "Plastics and articles thereof; rubber and articles thereof",  
 "numeral": "VII",  
 "chapter_from": "39",  
 "chapter_to": "40",  
 "section_note_id": null},  
{  
 "id": 8,  
 "position": 8,  
 "title": "Raw hides and skins, leather, furskins and articles thereof; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silkworm gut)",  
 "numeral": "VIII",  
 "chapter_from": "41",  
 "chapter_to": "43",  
 "section_note_id": null},  
{  
 "id": 9,  
 "position": 9,  
 "title": "Wood and articles of wood; wood charcoal; cork and articles of cork; manufactures of straw, of esparto or of other plaiting materials; basket-ware and wickerwork",  
 "numeral": "IX",  
 "chapter_from": "44",  
 "chapter_to": "46",  
 "section_note_id": null}
```



Linked data

Amazing but challenging to publish and use

EEE Building

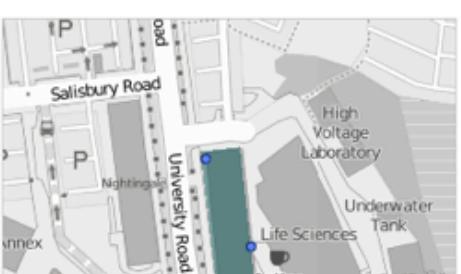
[http://id.southampton.ac.uk/building/32 ← This is the URI](http://id.southampton.ac.uk/building/32)

Detail	Facilities	Services	Energy
<p>Site: Highfield Campus</p> <p>Construction: 2006</p> <p>Architect: John McAslan & Partners</p> <p>Features: Building 32 is non-residential</p> <p>View Disability Report for this Building</p>			

 ©2010 Francois-Xavier Beckers (CC-BY)

Occupants

- Electronics & Computer Science
- Southampton Education School
- Agents, Interactions & Complexity
- Web & Internet Science
- Leadership School Improve &Effectiveness
- Lifelong & Work-Related Learning
- Mathematics & Science Education
- Social Justice & Inclusive Education
- Teaching Only Staff
- Deanery



Properties and Values

- soton.ac.uk/building/32 → rooms:Building, http://id.southampton.ac.uk/ns/UoSBuilding → "EEE Building"
- occupant → Electronics & Computer Science, Southampton Education School, Agents, Interactions Complexity, Web & Internet Science, ...show 8 more...
- construction → "32"^^http://id.southampton.ac.uk/ns/building-code-scheme
- locations:within → Highfield Campus
- <http://www.soton.ac.uk/estates/ourestate/buildings/highfield/32.html>
- "50.9364157"^^xsd:float
- "-1.395905"^^xsd:float
- southampton.ac.uk/ns/disabledGoPage → <http://www.disabledgo.com/en/access-guide/building-32>
- locations:easting → "442544"^^xsd:integer
- locations:northing → "115392"^^xsd:integer
- Organization → University of Southampton
- southampton.ac.uk/ns/ombielName → "Bldg 32 (EEE)"
- feature → Building 32 is non-residential
- southampton.ac.uk/ns/buildingDate → "2006"
- southampton.ac.uk/ns/buildingArchitect → John McAslan & Partners
- spatial → "POLYGON((-1.3961073411331264 50.93683868764933,-1.3958347895092957 50.9368567227702,-1.3956958407975968 50.936065737417,-1.3959558923017397 50.93603859197583,-1.3961073411331264 50.93683868764933))"
- southampton.ac.uk/ns/electricityTimeSeries → "elec/b32/ekw"
- ← is spatialrelations:within of ← 32 / 3077, 32 / 1015, Physical and Applied Science Faculty Deanery, Social and Human Sciences Faculty Deanery, ...show 54 more...
- ← is foaf:depicts of ← <http://data.southampton.ac.uk/image-archive/buildings/raw/32.jpg>, <http://data.southampton.ac.uk/image-archive/buildings/1000/32.jpg>, <http://data.southampton.ac.uk/image-archive/buildings/800/32.jpg>, ...show 5 more...
- ← is event:place of ← RaeS Solent Branch Christmas Special Lecture - The Red Arrows

 a Building
Google Images Ltd & Bluesky, The GeoInformation Group

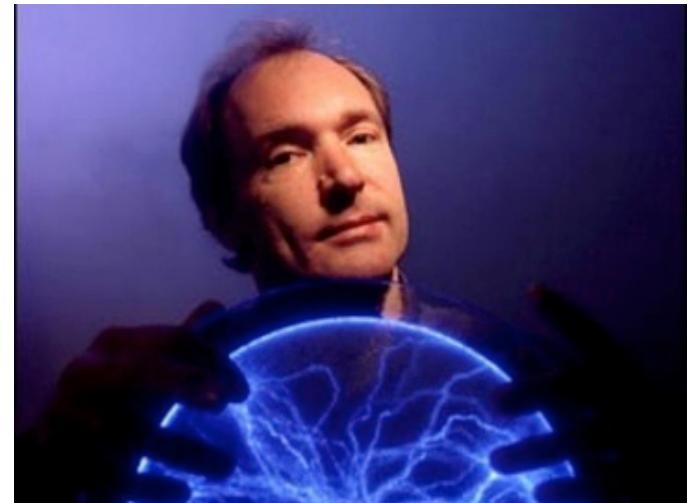


The Open Data Institute

Formats, Structures and Files

Data formats are complex and have ~~suffered~~ benefited from years of development in many different domains.

One thing is bringing them all together.



Recap

Specialist Solution

- + Easy to get setup and maintain.
- + Open Data focused
- + Clear workflows for publishing open data
- + Visualisation tools
- + Data mashing tools
- + Best for transactional data

Integrated Solution

- + No new platform to learn
- + Data is provided in parallel to web pages
- + No separation from authoritative data
- + Easy discovery of data
- + Best for reference data
- + Best for Linked Open Data

Both great for open data

Integrated solutions more suited for building a web of linked data



Exercise

Each table will have one of the publication platforms to explore.

Your task is to publish a prescribed dataset on this platform, with an associated open data certificate.



Outcomes

Understand the difference between “data on the web” and the “web of data”

Evaluate a number of different approaches for publishing open data.

Publish a dataset on one of the many available platforms.



Recap session



Today

The characteristics of open data

Open data discovery patterns

Open data publication

Quick big data break

Practical publication hands-on



The characteristics of open data

Identify a number of different characteristics of data

Explain the justifications for publishing different types of data

Evaluate the current open data ecosystem and future opportunities



Open data discovery patterns

Identify a number of different sources of open data on the web.

Create search patterns that enable easy discovery of new sources of open data.

Analyse the usability of available data and formulate plans for usage.

Understand the difference between “data on the web” and the “web of data.”



Open data publication

Understand the difference between “data on the web” and the “web of data”

Evaluate a number of different approaches for publishing open data.

Publish a dataset on one of the many available platforms.

