



<http://training.theodi.org/InPractice>

David Tarrant · @davetaz

Session 3

Data Publication Platforms



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.



Types

Specialist Solution	Integrated Solution
<ul style="list-style-type: none">+ Easy to get setup and maintain.+ Open Data focused+ Clear workflows for publishing open data+ Visualisation tools+ Data mashing tools+ Best for transactional data	<ul style="list-style-type: none">+ 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

1. Separate from your main org website
2. Designed to publish open data, not to fulfill other organisation goals



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



Key characteristics of integrated solution

1. It is your main website
2. 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

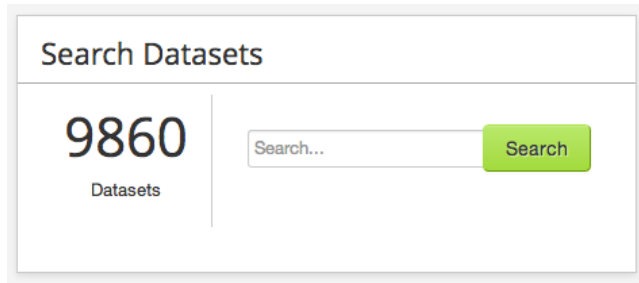
Specialist Solutions

1 2  comprehensive knowledge
archive network
<http://www.flickr.com/photos/okfn>

2 3  **OpenDataSoft**

4 5  **Socrata**











Licence

Open Government Licence **OGL** **OPEN DATA**

Data Resources **6**

-  XLS Tariff index
-  XLS Tariff data
-  DOC Tariff data - overview
-  TXT Tariff section index
-  DOC Tariff section notes
-  DOC Tariff chapter notes



Open Knowledge Foundation Supported

Data Catalogue

Open Source

Feels like a record manager

Simple API and search

Lots of community tools



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

Evolution of CKAN



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**



Open Data Soft

[← Back](#)

9,782 records

Text search

Code Postal

51100 17
30900 15
11100 13
21000 13
62100 13
06200 12

[More](#)

Présence

R 9,339
A 441
N 2

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	INTERMARCHÉ 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 Iecourbe	Avia	Gazole	1.46	1.7	1.72
9	SarI 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	INTERMARCHÉ 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



OpenDataSoft

Open Data Soft

[← Back](#)

9,782 records

Text search

Code Postal

51100	17
30900	15
11100	13
21000	13
62100	13
06200	12

[More](#)

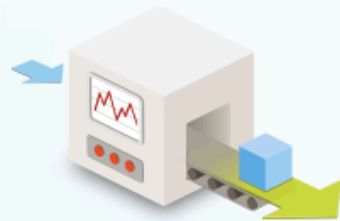
Présence

R	9,339
A	441
N	2

Data as a Service (DaaS)
Hosted enterprise solution
Rich Interface
Query based API (3-Star)



Open Data Soft



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



OpenDataSoft

Open Data Soft

[← Back](#)

9,782 records

Text search

Code Postal

51100	17
30900	15
11100	13
21000	13
62100	13
06200	12

[More](#)

Présence

R	9,339
A	441
N	2










Data as a Service (DaaS)
Closed Source (main product)
Hosted enterprise solution
EU Based
Rich Interface
Query based API (3-Star)





 Search

View Types

-  Datasets
-  Charts
-  Maps
-  Calendars
-  Filtered Views
-  External Datasets
-  Files and Documents
-  Forms
-  APIs

Data as a Service (DaaS)
Hosted enterprise solution
Allows user created content
Full linked API (5-Star)



 <https://opendata.socrata.com/>

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














Data as a Service (DaaS)
Closed Source (main product)
US Based

Clean Interface
Powerful API (SODA)

View Types

-  Datasets
-  Charts
-  Maps
-  Calendars
-  Filtered Views
-  External Datasets
-  Files and Documents
-  Forms
-  APIs





 Manage

 More Views

 Filter

 Visualize

 Export

 Discuss

 Embed

 About



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

1 2  comprehensive knowledge
archive network
<http://www.flickr.com/photos/okfn>

2 3  **OpenDataSoft**

4 5  **Socrata**



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

 **GOV.UK**

Search

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

Trade Tariff

Search the tariff

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

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

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

```
21  "id": 1,
22  "position": 1,
23  "title": "Live animals; animal products",
24  "numeral": "I",
25  "chapter_from": "01",
26  "chapter_to": "05",
27  "section_note_id": 1
28
29  "id": 2,
30  "position": 2,
31  "title": "Vegetable products",
32  "numeral": "II",
33  "chapter_from": "06",
34  "chapter_to": "14",
35  "section_note_id": 6
36
37  "id": 3,
38  "position": 3,
39  "title": "Animal or vegetable fats and oils and their cleavage products; prepared edible fats; animal or vegetable waxes",
40  "numeral": "III",
41  "chapter_from": "15",
42  "chapter_to": "15",
43  "section_note_id": null
44  },
45  {
46  "id": 4,
47  "position": 4,
48  "title": "Prepared foodstuffs; beverages, spirits and vinegar; tobacco and manufactured tobacco substitutes",
49  "numeral": "IV",
50  "chapter_from": "16",
51  "chapter_to": "24",
52  "section_note_id": null
53  },
54  {
55  "id": 5,
56  "position": 5,
57  "title": "Mineral products",
```



Linked data

Amazing but hard to publishing and use.

EEE Building

<http://id.southampton.ac.uk/building/32> ← This is the URI

Detail

Facilities

Services

Energy

Site: Highfield Campus
Construction: 2006
Architect: John McAsian & Partners
Features: Building 32 is non-residential

[View Disability Report for this Building](#)

Occupants

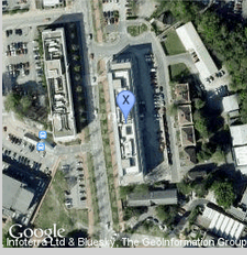
Electronics & Computer Science
Southampton Education School
Agents, Interactions & Complexity
Web & Internet Science
Leadership School Improve & Effectiveness
Lifelong & Work-Related Learning
Mathematics & Science Education
Social Justice & Inclusive Education
Teaching Only Staff
Deanery



©2010 Francois-Xavier Beckers (CC-BY)



ton.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...
Location → "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](#)
outhampton.ac.uk/ns/disabledGoPage → <http://www.disabledgo.com/en/access-guide/building->
Locations:easting → "442544" [xsd:integer](#)
Locations:northing → "115392" [xsd:integer](#)
Organization → University of Southampton
outhampton.ac.uk/ns/ombielName → "Bldg 32 (EEE)"
Feature → Building 32 is non-residential
outhampton.ac.uk/ns/buildingDate → "2006"
outhampton.ac.uk/ns/buildingArchitect → John McAsian & Partners
patial → "POLYGON((-1.3961073411331264 50.93683868764933,-1.3958347895092957 50.9368567227702,-1.3956958407975968
06065737417,-1.3959558923017397 50.93603859197583,-1.3961073411331264 50.93683868764933))"
outhampton.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>, <http://data.southampton.ac.uk/image-archive/buildings/600/32.jpg>, ...show 5 more...
← is event:place of ← RAeS Solent Branch Christmas Special Lecture - The Red Arrows





Formats, Structures and Files

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

One domain 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.





Thank-you