



Open Data in Practice

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

David Tarrant · @davetaz

Session 2

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.

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

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

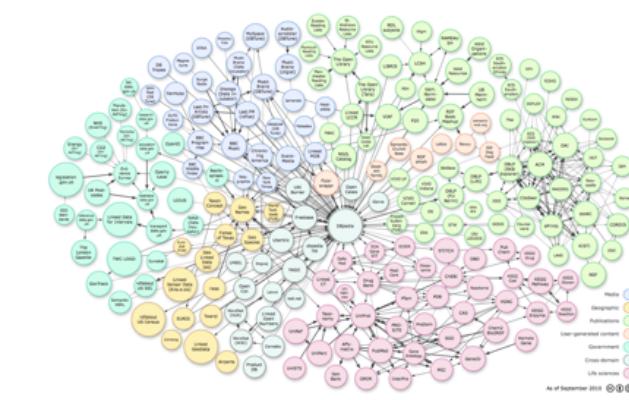


Approaches to publishing data

ON the web



IN the web

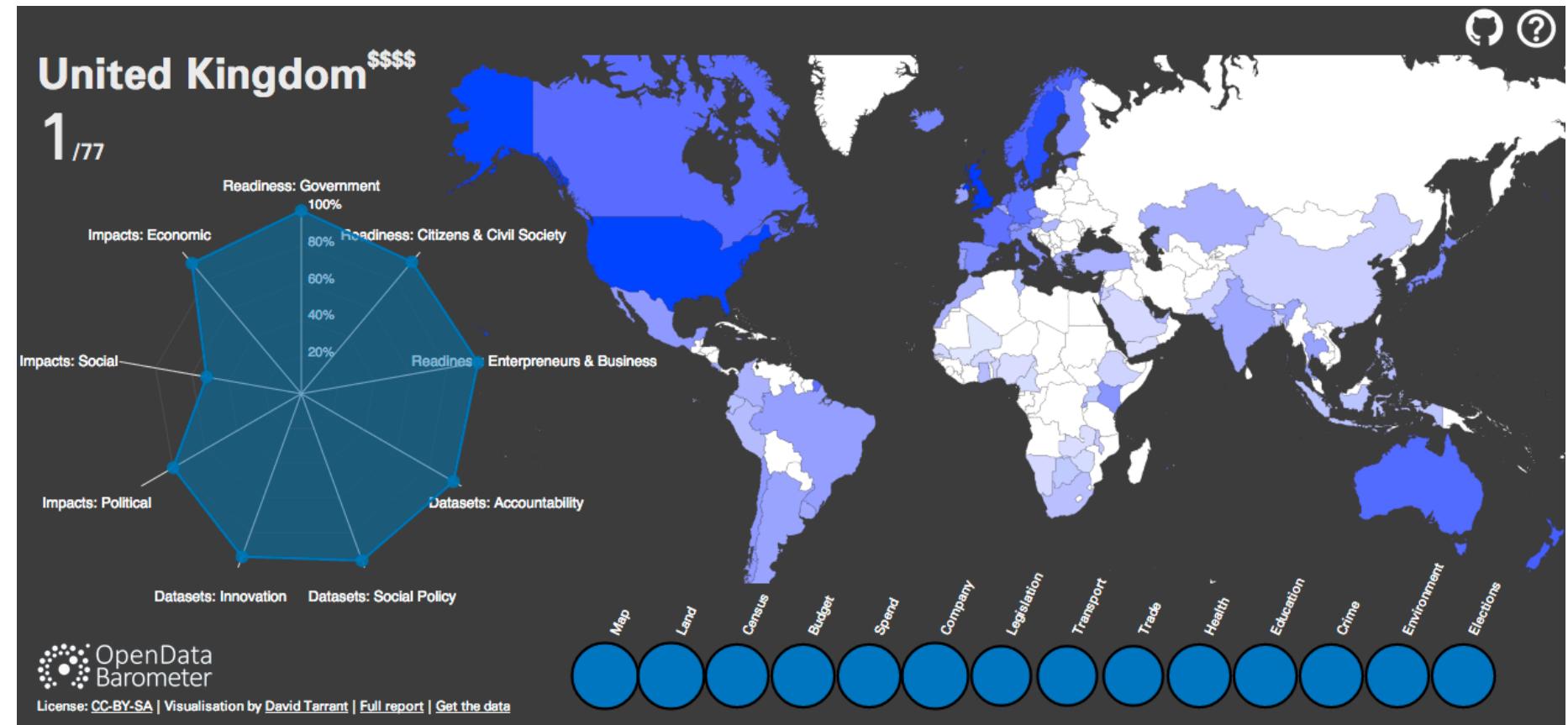


Finding data on the web **(of documents)**

- Government data
- Private sector data
- Google advanced
- Aggregators and portals
- Scraping



Government data



data.gov.XX

DATA.GOV.UK Beta
Opening up Government

Home Data Apps Interact

Datasets Map Search Data Requests Publishers Public Roles & Salaries Spend Reports Site Analytics Reports

Search for data... or conduct map based search

SHOW ONLY... Published datasets (15186)

Accueil Comprendre l'Open Data Thèmes Producteurs Jeux de données Applications Partenaires A propos

THEMES

- Agriculture
- Éducation
- Diplomatie
- Infrastructures
- Eau et Environnement
- TIC
- Santé
- Collectivités territoriales
- Sécurité Publique
- Tourisme et Culture

Open Data Burkina Faso statistiques

- 78 jeux de données
- 28 organisations
- 11 groupes

DATOS.GOB.MX BETA
OPEN TO PARTICIPATE

Data Stories Advances

/ datasets

Filter by your suburb Clear

Buscar conjuntos de datos...

127 datasets found Sort by: Relevance

RATING SOCIAL PROGRAMS

Database containing the evaluation of the results of social programs from the federal government subject to the annual assessment.

data.gov.sg

Singapore Government Integrity • Service • Excellence Contact Us | Sitemap | Feedback Search this site

Home Data Sharing Principles Data Catalogue App Showcase For Developers News & Events

First-stop to Discover Government Data

Search Data Catalogue

By Theme By Government Agency

Business & Economy Education Energy & Environment Finance

Government

The screenshot shows the INEGI (Instituto Nacional de Estadística y Geografía) website. The main navigation bar includes links for Statistics, Geography, Home, Contact, INEGI Mobile, and social media follow buttons. The page title is "Statistics by topic - Google Chrome". The URL in the address bar is www3.inegi.org.mx/sistemas/sisept/Default.aspx?t=mdemo148&s=est&c=29192. The main content area displays a table of population data for various Mexican states and the Federal District. A dropdown menu is overlaid on the table, offering export options: "Excel 5.0 (.Xls)" (selected), "Word (.Doc)", and "Format:". At the bottom right of the page, there is a large blue banner with the text "OPEN DATA".

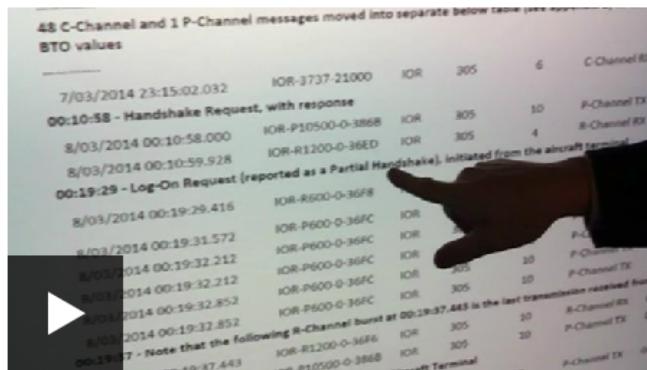
| | Federal District | 474 860 | 541 516 | 720 753 | 906 063 | 1229576 | 1757530 | 3050442 | 4870876 | 6874165 | 8831079 | 8235744 | 8489007 | 8605239 | 87 |
|-----------------|------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|----|
| Durango | 296 979 | 370 294 | 483 175 | 336 766 | 404 364 | 483 829 | 629 874 | 760 836 | 939 208 | 1182320 | 1349378 | 1431748 | 1448661 | 15 | |
| Oaxaca | 657 182 | 846 055 | 1040396 | 970 003 | 1084545 | 1152754 | 1421513 | 1727200 | 2013424 | 2369076 | 3019560 | 3228895 | 3438765 | 35 | |
| Puebla | 992 426 | 1021133 | 1101600 | 1024955 | 1150425 | 1294620 | 1625830 | 1973837 | 2508226 | 3347685 | 4126101 | 4624365 | 5076686 | 53 | |
| Querétaro | 232 305 | 232 389 | 244 663 | 220 231 | 234 058 | 244 737 | 286 238 | 355 045 | 485 523 | 739 605 | 1051235 | 1250476 | 1404306 | 15 | |
| Quintana Roo | NA | NA | 9109 | 10,966 | 10,620 | 18,752 | 26,967 | 50,169 | 88,150 | 225 985 | 493 277 | 703 536 | 874 963 | 11 | |
| San Luis Potosí | 571 420 | 575 432 | 627 800 | 445 681 | 579 831 | 678 779 | 856 066 | 1048297 | 1281996 | 1673893 | 2003187 | 2200763 | 2299360 | 24 | |
| Sinaloa | 261 050 | 296 701 | 323 642 | 341 265 | 395 618 | 492 821 | 635 681 | 838 404 | 1266528 | 1849879 | 2204054 | 2425675 | 2536844 | 26 | |
| Sonora | 192 721 | 221 682 | 265 383 | 275 127 | 316 271 | 364 176 | 510 607 | 783 3 | | | | | | | |
| Tabasco | 134 956 | 159 834 | 187 574 | 210 437 | 224 023 | 285 630 | 362 716 | 496 3 | | | | | | | |
| Tamaulipas | 209 106 | 218 948 | 249 641 | 286 904 | 344 039 | 458 832 | 718 167 | 10241 | | | | | | | |
| Tlaxcala | 168 358 | 172 315 | 184 171 | 178 570 | 205 458 | 224 063 | 284 551 | 346 6 | | | | | | | |



Government / Private



Flight MH370: Malaysia releases raw satellite data



The BBC's Richard Westcott visited Inmarsat's headquarters to find out what the data tells us about MH370's fate

The Malaysian government has released the raw data used to determine that the missing Malaysia Airlines flight MH370 crashed into the southern Indian Ocean.

The data was first released to relatives of passengers, who have been asking for greater transparency, before copies were also provided to media.

The document released on Tuesday comprises 47 pages of data, plus notes, from British firm Inmarsat.

| Time | Channel Name | Ocean Region | GES ID (octal) | Channel Unit ID | Channel Type | SU Type | Burst Frequency Offset (Hz) BFO | Burst Timing Offset (microseconds) BTO |
|----------------------|-------------------|--------------|----------------|-----------------|--------------|-----------------------------------|---------------------------------|--|
| 03/2014 16:00:13.406 | IOR-R1200-0-3603 | IOR | 305 | 8 | R-Channel RX | 0x15 - Log-on/Log-off Acknowledge | | |
| 03/2014 16:00:13.905 | IOR-P10500-0-3859 | IOR | 305 | 10 | P-Channel TX | 0x15 - Log-on/Log-off Acknowledge | 103 | 14820 |
| 03/2014 16:00:17.430 | IOR-R1200-0-3603 | IOR | 305 | 8 | R-Channel RX | Eleven Octet User Data | | |
| 03/2014 16:00:17.906 | IOR-R1200-0-3603 | IOR | 305 | 8 | R-Channel RX | Eleven Octet User Data | 103 | 14740 |
| 03/2014 16:00:18.404 | IOR-R1200-0-3603 | IOR | 305 | 8 | R-Channel RX | Eight Octet User Data | 103 | 14780 |
| 03/2014 16:00:18.905 | IOR-P10500-0-3859 | IOR | 305 | 10 | P-Channel TX | 0x62 - Acknowledge User Data | 103 | 14820 |
| 03/2014 16:00:20.906 | IOR-P10500-0-3859 | IOR | 305 | 10 | P-Channel TX | 0x71 - User Data (ISU) - RLS | | |
| 03/2014 16:00:20.906 | IOR-P10500-0-3859 | IOR | 305 | 10 | P-Channel TX | Subsequent Signalling Unit | | |
| 03/2014 16:00:22.906 | IOR-R1200-0-3603 | IOR | 305 | 10 | P-Channel TX | Subsequent Signalling Unit | | |
| 03/2014 16:00:23.407 | IOR-R1200-0-3603 | IOR | 305 | 8 | T-Channel RX | Subsequent Signalling Unit | | |
| 03/2014 16:00:23.905 | IOR-P10500-0-3859 | IOR | 305 | 8 | T-Channel RX | Subsequent Signalling Unit | 86 | 9820 |
| 03/2014 16:00:27.741 | IOR-T1200-0-3607 | IOR | 305 | 8 | T-Channel RX | Subsequent Signalling Unit | 88 | 9820 |
| 03/2014 16:00:27.901 | IOR-T1200-0-3607 | IOR | 305 | 8 | T-Channel RX | Subsequent Signalling Unit | 88 | 9820 |
| 03/2014 16:00:28.061 | IOR-T1200-0-3607 | IOR | 305 | 8 | T-Channel RX | Subsequent Signalling Unit | 88 | 9820 |
| 03/2014 16:00:28.221 | IOR-T1200-0-3607 | IOR | 305 | 8 | T-Channel RX | Subsequent Signalling Unit | 88 | 9820 |
| 03/2014 16:00:28.405 | IOR-T1200-0-3607 | IOR | 305 | 8 | T-Channel RX | Subsequent Signalling Unit | 88 | 9820 |
| 03/2014 16:00:28.541 | IOR-T1200-0-3607 | IOR | 305 | 8 | T-Channel RX | Subsequent Signalling Unit | 88 | 9820 |

OPEN DATA ?

What we know



Suppliers

The screenshot shows the BP website's 'Statistical Review downloads' page. At the top, there's a navigation bar with links for 'About BP', 'Products and Services', 'Sustainability', 'Investors', 'Press', 'Careers', and 'Gulf of Mexico restoration'. Below the navigation is a breadcrumb trail: Home > About BP > Energy economics > Statistical Review of World Energy 2013 > Downloads. The main content area is titled 'Statistical Review downloads' and contains a sub-section titled 'Statistical Review 2013'. It describes the section for downloading the printed edition, supporting PowerPoint Excel workbook, and launch presentation speech and slides. Below this, there are download links for 'Statistical Review of World Energy 2013' (pdf, 9.6MB), 'Statistical Review 2013 workbook' (xlsx, 1.5MB) [highlighted with a red box], 'Statistical Review 2013 speech slide pack' (pptx, 2.9MB), and 'Statistical Review 2013 Speech' (pdf, 663.0KB). A large red 'X' is overlaid on the right side of the page.



BP: You may not frame this site nor link to a page other than the home page without our express permission.
To find this Google “**bp statistical review**”



Suppliers (ish)

The screenshot shows the Sky website's 'Bigger Picture' section. The left sidebar lists various data categories. The main content area features a heading 'Environmental impact data' with a sub-section 'Our progress so far'. It includes a photograph of a film crew and a summary of Sky's environmental targets and progress. A table titled 'Summary performance against environment targets' is shown, with a large red 'X' overlaid on the right side.

Bigger Picture

- Approach
- Our contribution
- Responsible business
- Inspiring action
- News, views & videos
- How we are doing
- How we report
- Data commentary
- Financial data
- Creative industries data
- UK economy data
- Customers data
- People data
- Business partners data
- Environmental impact data**
- Community data
- Sport data

Environmental impact data

At Sky we understand to build a sustainable business we need to minimise our environmental impact from our operations, create more sustainable products and services and work with our business partners and customers to inspire them to take action to protect the environment too.

Our progress so far

Last year we reported good progress towards our ten environment targets that were set in 2009 so we set new, more ambitious targets to 2020. This year we have continued to reduce our emissions when normalised against turnover. We have achieved this by investing in energy efficiency measures, working with our people to reduce the energy they use day to day and through our on-site renewable energy sources. Detailed commentary about this data is provided below.

| | Target | 2009/10 | 2010/11 | 2011/12 | 2012/13 |
|--|--------|---------|---------|---------|---------|
| Reduction in gross CO ₂ e emissions relative to revenue(%) ¹ | -50 | -8 | -21 | -29 | -33 |
| Energy obtained from owned or controlled renewable energy at Sky-owned sites(%) ² | 20 | - | - | 2 | - |
| Increase in fleet fuel efficiency(%) ³ | -15 | - | - | - | -5 |
| Reduction in CO ₂ e | | | | | |



http://corporate.sky.com/the_bigger_picture/how_we_are_doing/environmental_impact_data

Suppliers (ish)

Product Environment Reports



Date introduced
September 10, 2013

Environmental Status Report

iPhone 5s is designed with the following features to reduce environmental impact:

- Arsenic-free display glass
- Mercury-free LED-backlit display
- Brominated flame retardant-free
- PVC-free

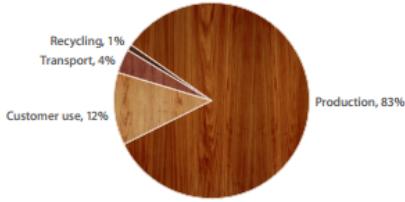
Apple and the Environment

Apple believes that improving the environmental performance of our business starts with our products. The careful environmental management of our products throughout their life cycles includes controlling the quantity and types of materials used in their manufacture, improving their energy efficiency, and designing them for better recyclability. The information below details the environmental performance of iPhone 5s as it relates to climate change, energy efficiency, material efficiency, and restricted substances.¹

Climate Change

Greenhouse gas emissions have an impact on the planet's balance of land, ocean, and air temperatures. Most of Apple's corporate greenhouse gas emissions come from the production, transport, use, and recycling of its products. Apple seeks to minimize greenhouse gas emissions by setting stringent design-related goals for material and energy efficiency. The chart below provides the estimated greenhouse gas emissions for iPhone 5s over its life cycle.²

Greenhouse Gas Emissions for iPhone 5s



| Stage | Percentage |
|--------------|------------|
| Production | 83% |
| Customer use | 12% |
| Transport | 4% |
| Recycling | 1% |

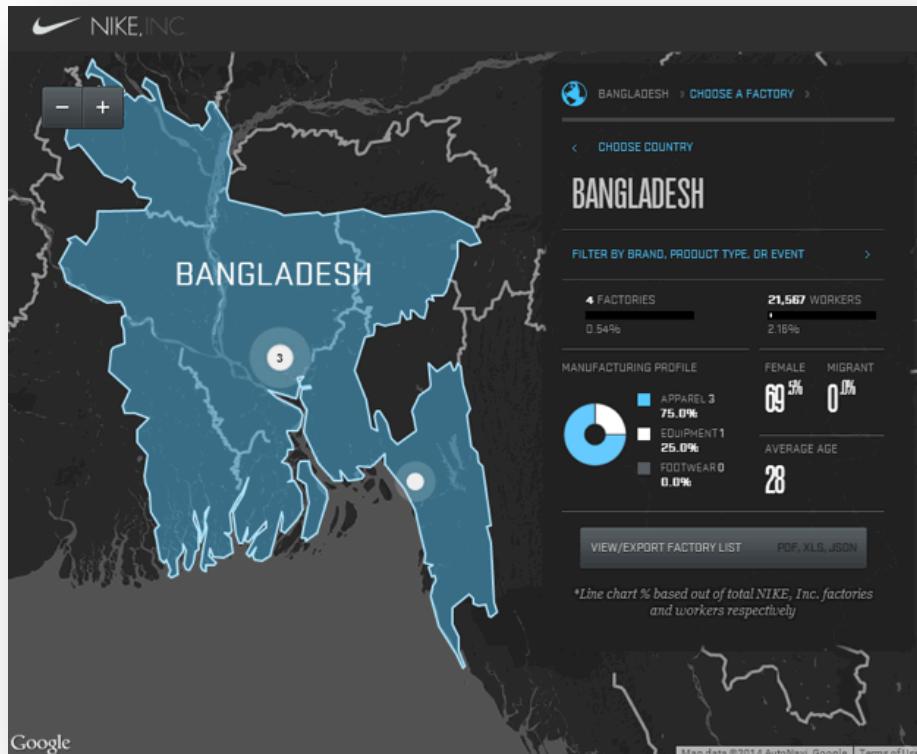
Total greenhouse gas emissions: 75 kg CO₂



<http://www.apple.com/uk/environment/reports/>



Suppliers



<http://manufacturingmap.nikeinc.com/#>

You agree not to change or delete any ownership notices from materials downloaded or printed from the Platform. You agree not to modify, copy, translate, broadcast, perform, display, distribute, frame, reproduce, republish, download, display, post, transmit or sell any Intellectual Property or Content appearing on the Platform



Google advanced

Google site:gov filetype:xls

Web Images Maps Shopping More ▾ Search tools

About 4,150,000 results (0.22 seconds)

[XLS] [Code List or Concept \(Acronym\)](#) ↗
www.acquisition.gov/short_codelistsTS.xls Share
File Format: Microsoft Excel - View as HTML
A, B, C, D, F, G, H, 1, Code List or Concept (Acronym), Definition, Authoritative or Maintenance Agency, Current Version, Update Frequency and/or ...

[XLS] [Approps - Foreign Assistance.gov](#) ↗
www.foreignassistance.gov/Full_ForeignAssistanceData.xls
File Format: Microsoft Excel
A, B, C, D, E, F, G, H, 1, Planning Data, 2, 3, 4, Fiscal Year, Fiscal Year Type Account Name, Agency Name, Operating Unit, Category, Sector, Amount ...

[XLS] [TSB Monthly Cash Flow Projection](#) ↗
www.dia.iowa.gov/tsb/cashflow.xls

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

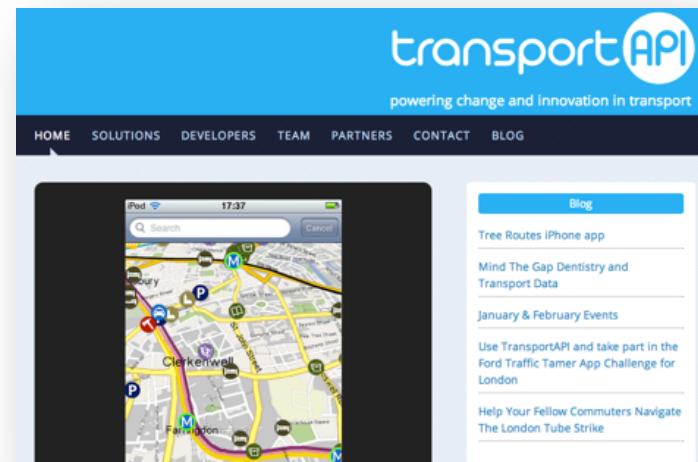


Aggregators and portals

Collect together data from across the web into one place.



enigma.io

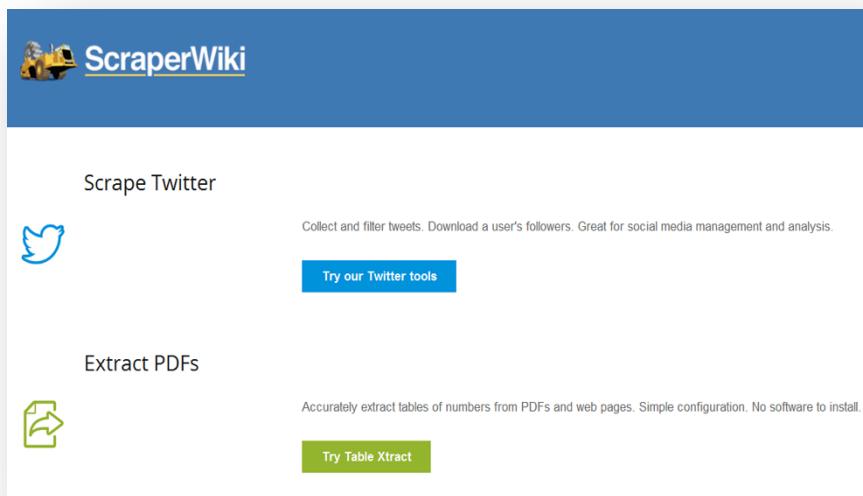


transportAPI



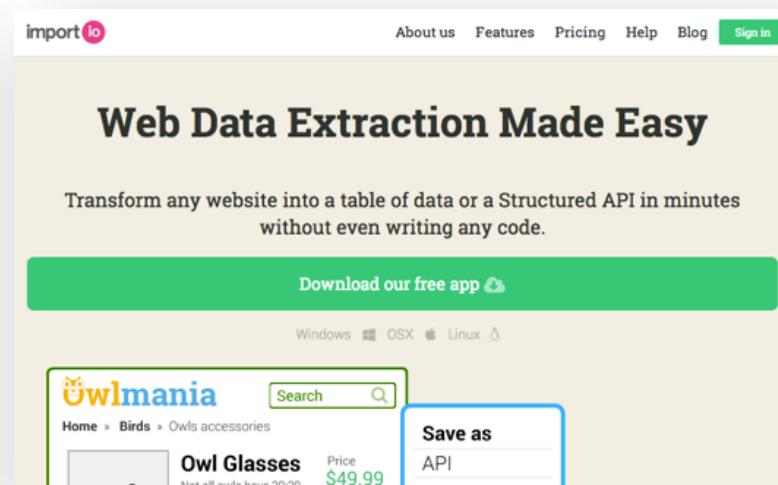
Scraping

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



The ScraperWiki homepage features a blue header with the logo 'ScraperWiki' and a yellow icon of a bulldozer. Below the header, there are two main sections: 'Scrape Twitter' and 'Extract PDFs'. The 'Scrape Twitter' section includes a Twitter icon, a brief description, and a 'Try our Twitter tools' button. The 'Extract PDFs' section includes a PDF icon, a brief description, and a 'Try Table Xtract' button.

scraperwiki.com



The import.io homepage has a light beige background. At the top, it displays the 'import.io' logo and a navigation bar with links to 'About us', 'Features', 'Pricing', 'Help', 'Blog', and 'Sign in'. The main heading is 'Web Data Extraction Made Easy', followed by a subtext: 'Transform any website into a table of data or a Structured API in minutes without even writing any code.' A large green button says 'Download our free app' with a download icon. Below this, there's a screenshot of a web page from 'Owlmania' showing an 'Owl Glasses' item with a price of '\$49.99'. A 'Save as API' button is overlaid on the screenshot. At the bottom, there are links for 'Windows', 'OSX', and 'Linux'.

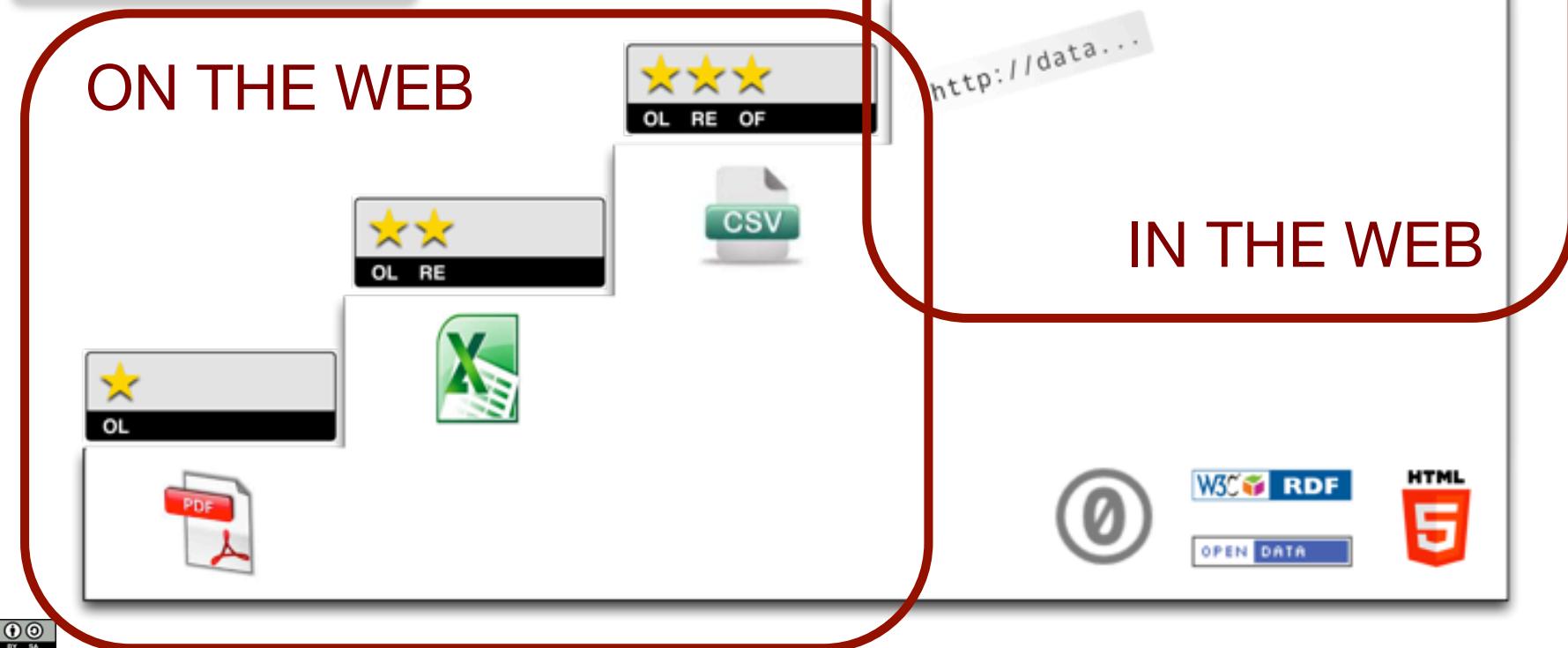
import.io



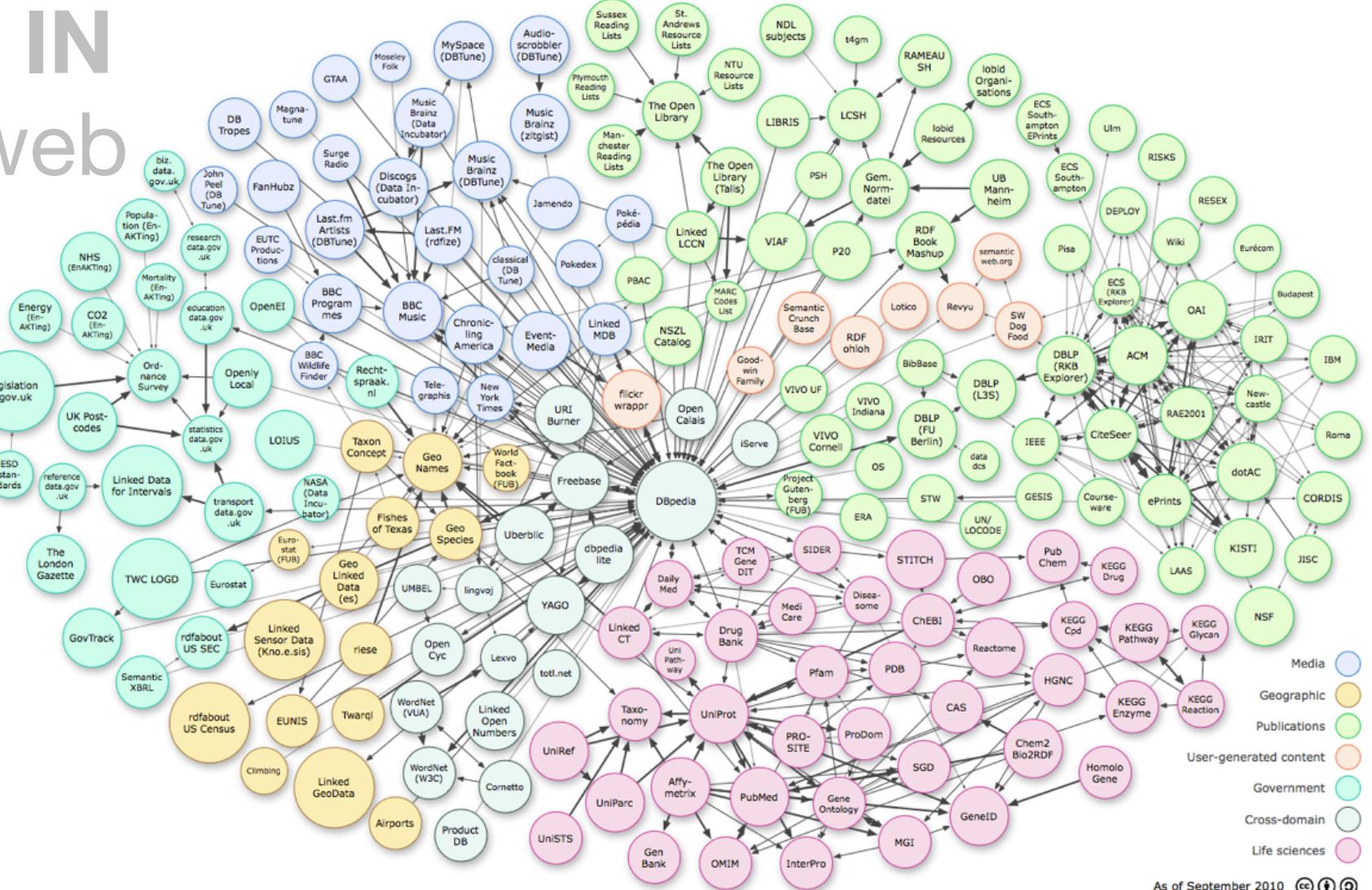
5-Stars



ON THE WEB



Data IN the web



As of September 2010 

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 **IN THE WEB**
4. Do some content negotiation
5. Spot the API
6. Scrape (or search google again)



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

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

GOV.UK

Search

Home > Business and self-employed > Imports and exports

Trade Tariff

Search the tariff name or code

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 |

UK Trade Tariff



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

one

DOCTOR WHO

Home Episodes Clips Galleries Latest News Characters Monsters Fun and Games More

On iPlayer
This programme will be available shortly after broadcast



It's Tomorrow... Get the Latest on the Launch!
What's happening and how to follow the action during tomorrow's big launch in Cardiff.

On TV



The Day of the Doctor
SATURDAY 19:00
BBC THREE
All upcoming
(0 NEW AND 1 REPEAT)

BBC Music and Programmes

2. Look for alternative links



Business Insight - NEWSASIA

NEWS TV WATCH LIVE

Wed, Aug 06 2014

ASIA PACIFIC SINGAPORE WORLD BUSINESS SPORT ENTERTAINMENT TECHNOLOGY HEALTH LIFESTYLE VIDEOS WEATHER MORE ▾

CHANGELIVES LUMINARY AWARDS START-UP

APORE STORIES

Raise of up to 12% for Home Team officers, with sign-on bonuses of up to S\$30,000

al Day Award 2014

SP (2) Ng

(SP)

MEDIACORP

Officers from the Home Team, both past and present were recognised at the Home Team National Day Observance Ceremony on Wednesday.

9 hours ago

Pay rise, special bonus for about 23,000 nurses

10 hours ago

50,000 openings on Jobs Bank for Singaporeans, PRs

1 hour ago

NUS University Town identified as a high-risk dengue cluster

10 hours ago

Scroll down!

LIFESTYLE

VIDEOS

A large black arrow points downwards from the headline "Raise of up to 12% for Home Team officers, with sign-on bonuses of up to S\$30,000" towards the "LIFESTYLE" and "VIDEOS" sections at the bottom of the page.



2. Look for alternative links



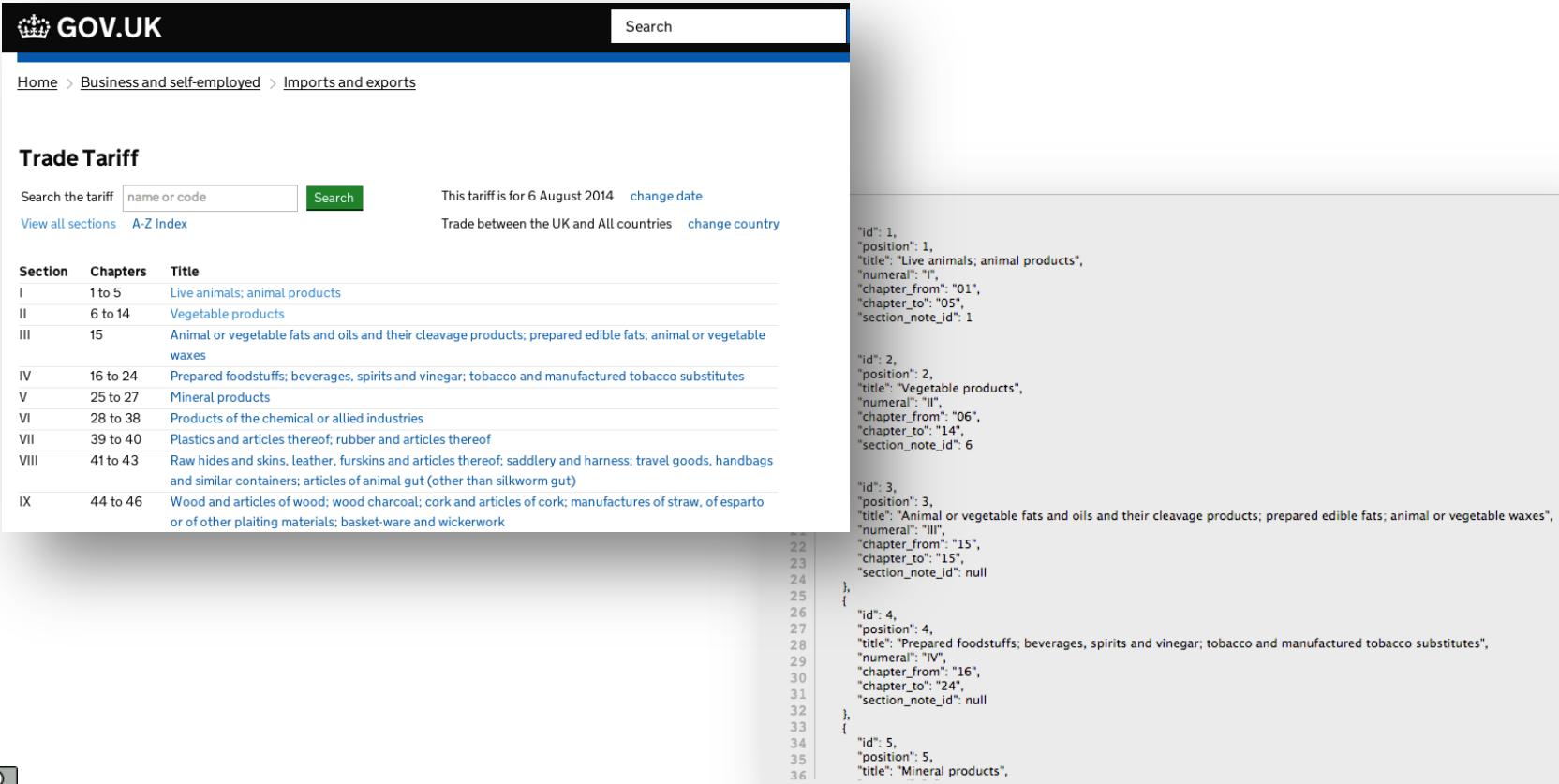
| | | | |
|---|---|---|---|
|  CHANNEL NEWSASIA MediaCorp News Group. © 2014 MediaCorp Pte Ltd. All Rights Reserved. Terms and Conditions Privacy Policy About MediaCorp Pte Ltd | NEWS Asia Pacific Singapore World Business Sport Entertainment Technology Health Lifestyle Videos Photos Special Reports Archives | 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



The developers secret



GOV.UK

Search

Home > Business and self-employed > Imports and exports

Trade Tariff

Search the tariff 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": 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": 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": 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}]]

CC BY SA


```

Linked data

Amazing but hard to publishing 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](#)

Site: Highfield Campus
Construction: 2006
Architect: John McAslan & 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)



Salisbury Road
Nightingale
University Road
Annex
High Voltage Laboratory
Underwater Tank
Life Sciences
Lattes
R.J. Mitchell

[a Building](#)

[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/oresteate/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-](#)
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](#), [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](#)



Finding data on the web

(
Techniques 3-5 are not
covered in this session. Please
ask your trainer for more
information if there is time.

1. Look for JSON (json, .csv etc)
2. Look for XML (feeds etc)
3. Look for embedded data
4. Do some content negotiation
5. Spot the API
6. Scrape (or search google again)



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.

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

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



