

Introduction to linked data

Linked data principles



Lets start with a question

- Are we still building new houses on flood plains?
 - Recent UK concern over builders building new homes that were liable to flood
- How might we go about answering that question?
- Look at the available data
 - Land Registry publish data about house sales
 - Environment Agency publish data about flood risks
 - Ordnance Survey publish data about postcodes

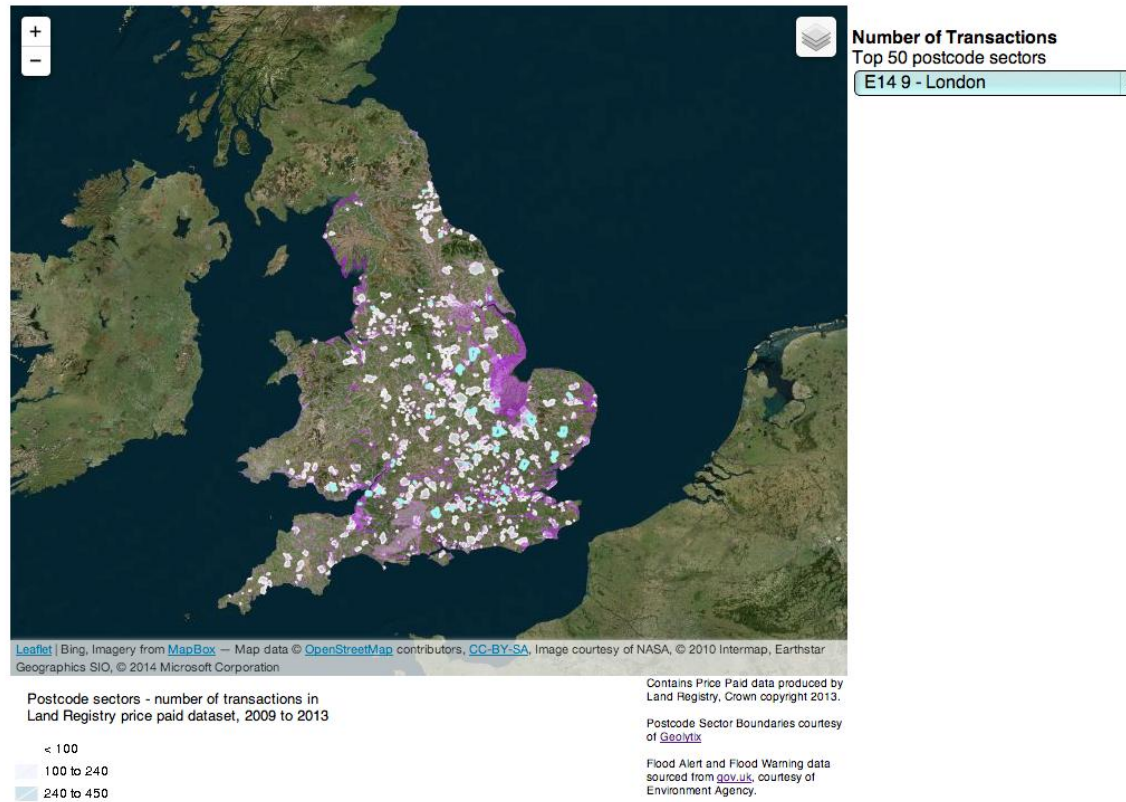
Steve Peters built this app

https://openviz.files.wordpress.com/2014/03/faw_home.png

Flood alert and warning areas and
New build property transaction hotspots, 2009 to 2013

An open data demonstrator powered
OPEN DATA COMMUNITIES
by Open Access to Local Data

April 2014



Data Reuse

- None of the data publishers envisaged this use of the data
 - Land Registry – data is used by the housing market
 - Environment Agency - warning the public of floods
 - Ordnance Survey – publishing key reference data
- Reusable data is valuable
- Reusable data is a rich economic resource
- Inside the enterprise and in public
- That value is reduced by "information friction"

Information Friction

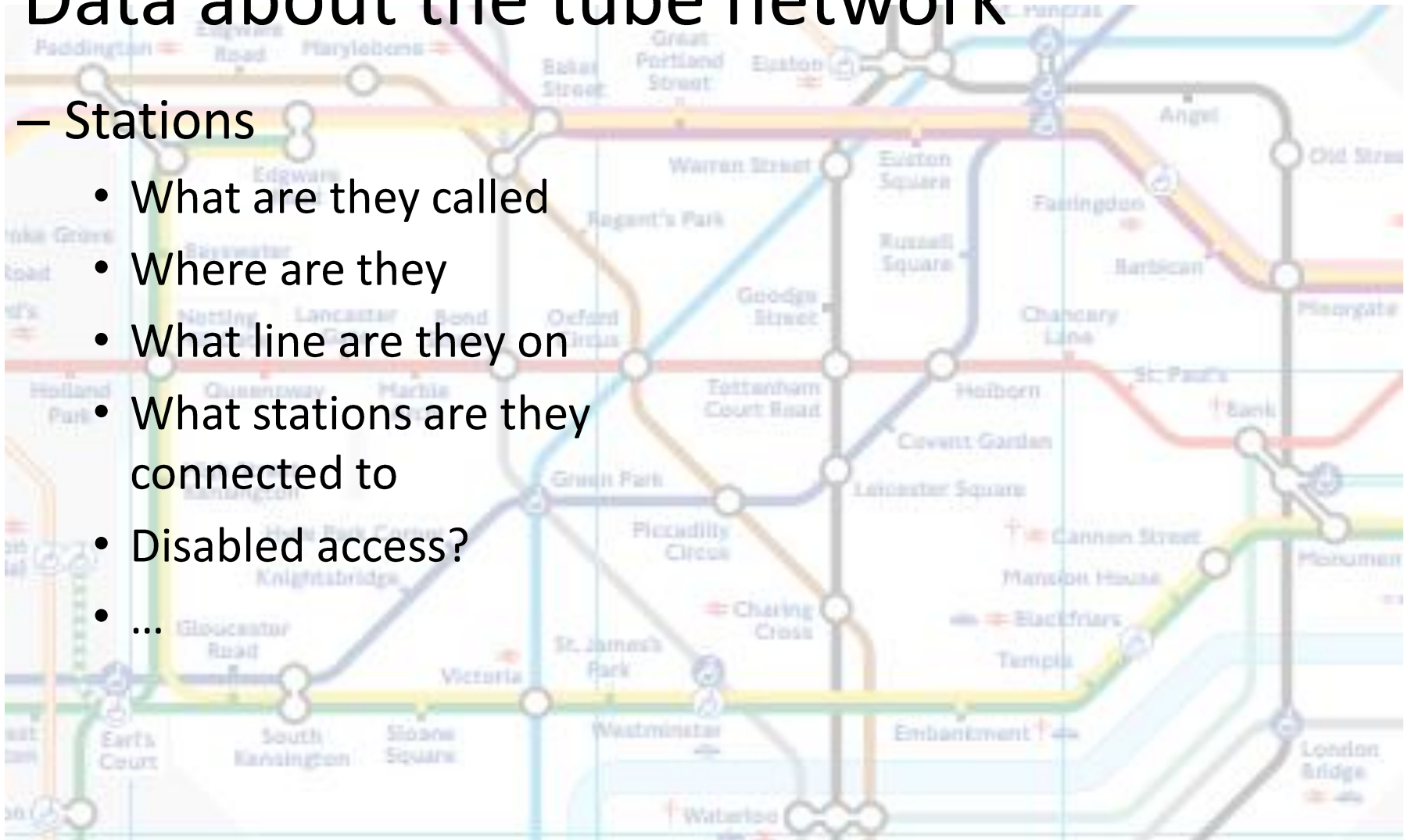
- Anything that gets in the way of using data to answer a question
- Examples of friction
 - finding data (knowing it exists)
 - accessing data
 - knowing what the data actually means
 - integrating data from different sources
 - the data structures don't join up
 - they don't use the same identifiers
 - ...

London Transport

Data about the tube network

– Stations

- What are they called
- Where are they
- What line are they on
- What stations are they connected to
- Disabled access?
- ...



London Tourist Information Data about 'points of interest'

– Point of interest

- name
- What type
- Description
- Location
- Current exhibitions
- Opening times
- How to get there
- Disabled access
- ...

Inside the National Gallery, London

5 Minute Exercise

- Find a partner to work with
- For each of London Transport and London Tourist information
 - Sketch a single page spreadsheet with some data they might publish
 - data that app developers might find useful
 - single page so can share as CSV file
- You could choose to:
 - work on one each
 - work independently
 - see how well the data would join up

Solutions

How was that?

? Issues / Problems / Observations

? do the two CSV files connect?

Two I prepared Earlier

| Station | Line | Dir | Next | Line | Dir | Next | ... | Lat/long | Disabled |
|---------------|---------|------|----------------------|---------|------|-------------|-----|----------|----------|
| Oxford Circus | Central | East | Tottenham Court Road | Central | West | Bond Street | ... | ... | ... |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |

| POI | Type | Description (English) | Description (French) | ... | Opening Times | Map Ref | Tube | .. |
|------------------|-------------|--------------------------|--------------------------|-----|--|---------|---------------|-----|
| National Gallery | Art Gallery | The national gallery ... | La Galerie Nationale ... | ... | Daily 10am – 6pm, Friday 10am – 9pm Closed ... | ... | Charing Cross | ... |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |

Issues / Problems / Observations

? Repeated variable number of columns for connections, descriptions etc

? Opening times a bit tricky – need a human to read those

? Irregular data – current exhibitions makes sense for a gallery – but for a park?

? Different map coordinate systems in different datasets

? Not sure if the data would join up

epimorphics
Image courtesy of Ryan Wick - <http://www.flickr.com/photos/ryanwick/>
depends on spelling of station names

Introduction to Linked Data

epimorphics.com

▶ Classroom – learn RDF and SPARQL by doing

- ▶ London Museums
- ▶ London Tube Stations

- ▶ What is linked data?
- ▶ How does linked data work?
- ▶ Using linked data
- ▶ Querying linked Data
- ▶ Getting started

Resources

- ▶ Resources and links

Culture

Linked Data API

csv | html | json | rdf | text | ttl | xml


Search Results

Show Search Form

Art in Perpetuity Trust

long -0.0208

lat 51.4769



nearest underground station distance 327

comment Art in Perpetuity Trust, also known as APT, is a London-based charity focusing on developing visual arts. It comprises 37 studios for visual artists and a gallery in Deptford.

type

- Museum
- Museum
- Place
- TouristAttraction


nearest underground station Deptford Bridge

same as Art_in_Perpetuity_Trust

British Museum

long -0.126931

lat 51.519459



nearest underground station distance 431

comment The British Museum is a museum in London dedicated to human history and culture. Its permanent collection, numbering some eight million works, is among the largest and most comprehensive in existence and originates from all continents, illustrating and documenting the story of human culture from its beginnings to the present. The British Museum was established in 1753, largely based on the collections of the physician and scientist Sir Hans Sloane.

type

- Museum
- Museum
- Place

Map



On This Page

- > Art in Perpetuity Trust
- > British Museum
- > De Morgan Centre
- > Dulwich Picture Gallery
- > Foundling Museum
- > Geffrye Museum
- > Institute of Contemporary Arts
- > National Maritime Museum
- > October Gallery
- > Parasol unit foundation for contemporary art
- > next >

Sort by

- label
- long
- lat
- nearest underground station distance
- comment
- type
- nearest underground station
- same as

View

training.epimorphics.com/transport/london-underground/station/deptford_bridge

Apps admin LR EPO ops Search Results Home - epimorph

Transport

Linked Data API

Deptford Bridge [Show Search Form](#)

http://training.epimorphics.com/transport/london-underground/station/deptford_bridge

long -0.0216

lat 51.474

incoming connection

Connection from Elverson Road to Deptford Bridge along the Docklands Light Railway

from stop Elverson Road

on line Docklands Light Railway

Connection from Greenwich to Deptford Bridge along the Docklands Light Railway

from stop Greenwich

on line Docklands Light Railway

nearby to

Art in Perpetuity Trust

long -0.0208

lat 51.4769

nearest underground station distance 327

comment Art in Perpetuity Trust, also known as APT, is a London-based charity focusing on developing visual arts. It comprises 37 studios for visual artists and a gallery in Deptford.

type Museum
Museum
Place
TouristAttraction

nearest underground station deptford_bridge

same as Art in Perpetuity Trust

csv | html | json | rdf | text | ttl | xml

Map

View

> all > basic > description > outgoing > station

- ★ long
- ★ lat
- ★ incoming connection > label
- ★ incoming connection > from stop
- ★ incoming connection > on line
- ★ nearby to > long
- ★ nearby to > lat
- ★ nearby to > nearest underground station distance
- ★ nearby to > comment
- ★ nearby to > label
- ★ nearby to > type
- ★ nearby to > nearest underground station
- ★ nearby to > same as
- ★ outgoing connection > label
- ★ outgoing connection > on line
- ★ outgoing connection > to stop

Observations

- All the interesting things have a web address
 - <http://training.epimorphics.com/transport/london-underground/station/paddington>
- Get data back from those web addresses as
 - A web page (html)
 - Or RDF (web standard data format)
 - Or JSON, XML, CSV, Text
- That data uses web addresses to link to data about other things
 - POI data links to Tube data
- Those links let you go hopping from one dataset to another
- Creating a Web of data

Linked Data ‘Manifesto’...

<http://www.w3.org/DesignIssues/LinkedData.html>

1. “Use URIs as names for things”
 - [ie. use web addresses to name all manner of things.]
2. “Use HTTP URIs so that people can look up those names.”
3. When someone looks up a URI, provide useful information, using standards (RDF)
 - [ie. put information behind the name]
4. “Include links to other URIs. so that they can discover more things.”



Linked Data's Strengths

- data integration
- reuse of data
- structural flexibility
- extensibility
- accessibility

http://explore.data.parliament.uk/

Under the hood you will find

The screenshot shows a web browser window with the URL `lda.data.parliament.uk/researchbriefings?_view=Research+Papers&_pageSize=10&_page=0`. The browser's address bar and tabs are visible. The website header includes the logo for `data.parliament.uk` and a navigation menu with options like `xml`, `ttl`, `text`, `rdf`, `json`, `html`, and `csv`.

The main content area is titled "List of API endpoints" and features a "Search Results" section. The search results for ID `419675` are displayed in a table-like format:

| | |
|--------------------|---|
| 419675 | |
| date | 02/10/2015 17:35:24 |
| abstract | A new National Living Wage (NLW) was announced by the Chancellor of the Exchequer at the 2015 Summer Budget. The NLW applies to workers aged 25 and over and will be introduced at an initial rate of £7.20 per hour from April 2016. It effectively takes the place of the National Minimum Wage for workers aged 25 and over. The Government intends the NLW to reach 60% of median earnings by 2020 and has asked the Low Pay Commission, which currently recommends the level of the National Minimum Wage, to set out how this may be achieved. The Office for Budget Responsibility estimates this would mean the NLW will be around £9.35 per hour in 2020, based on its forecasts for average earnings. The economic impacts of the NLW are highly uncertain and largely depend on how employers choose to deal with the extra costs imposed by having to pay staff a higher wage. This note looks at the likely numbers of people affected and other economic impacts of the change. |
| description | The Government announced a new National Living Wage at the 2015 Summer Budget. This note gives a brief overview of the possible impact, including numbers of employees affected. |
| identifier | CBP-7319 |
| section | [17113] |

On the right side of the page, there are two panels: "On This Page" showing a list of results (1 to 10 of 6113) and "Sort by" with various sorting options like `date`, `abstract`, `description`, `identifier`, `section > pref label`, `title`, `topic > pref label`, `type`, `publisher > pref label`, and `sub type > pref label`.

http://www.streetlist.co.uk

http://www.streetlist.co.uk/streets/Sole-Street-Wigan-WN1-3YE#.VPRM581tPUY

Worsley Hall, Marsh Green and Kitt Green.

Moore Street East



Recent House Sales in Sole Street

7 Sole Street

Sale Price: £53,000

Sale Date: 11/04/2013

6 Sole Street

Sale Price: £91,000

Sale Date: 01/10/2007

Source: Land Registry Price Paid Data accessed via API



Population Statistics

ONS Area Code: E01006387

ONS Area Name: Wigan 009D

Average Age : 41.36

Population: 1319.00

Male Population: 650.00 (49.28%)



Nearest Libraries

Ince Library

Smithy Green, Wigan, WN22AT

Distance: 0.70 Miles

[Visit Website](#)

Wigan Childrens Library

The Wiend Centre, Wigan, WN11PF

Distance: 0.77 Miles

[Visit Website](#)

Wigan Library

College Avenue, Wigan, WN11NN

Distance: 0.79 Miles

[Visit Website](#)



Nearest Tourist Attractions

Worsley Hall, Marsh Green and Kitt Green.

[Visit Greater-manchester Police Website](#)

[Email Greater-manchester Police](#)

Telephone: 0161 856 7124

Policing Links:

[Neighbourhood Community Website](#)

Recent Crime in the Area

| | |
|-----------------------|----|
| Public-order | 28 |
| Violent-crime | 80 |
| Criminal-damage-arson | 31 |
| Drugs | 11 |
| Other-theft | 33 |
| Shoplifting | 29 |
| Burglary | 34 |
| Theft-from-the-person | 14 |
| Vehicle-crime | 11 |
| Bicycle-theft | 4 |
| Other-crime | 4 |
| Robbery | 3 |
| Possession-of-weapons | 2 |

Source: [data.police.uk](#) accessed via API

Map of Crimes Committed Near Sole Street



http://landregistry.data.gov.uk

HM Government

Linked Data API

csv | html | json | rdf | text | ttl | xml

Search Results Show Search Form

2013-11-57320

| | | |
|------------------|--|-------|
| transaction id | CFE63CDD-DDB0-42F3-AACD-00CA35248E46 | 🔍 |
| new build | <input checked="" type="checkbox"/> | 🔍 |
| publish date | 06/11/2013 | < 🔍 > |
| transaction date | 25/01/2002 | < 🔍 > |
| price paid | 73000 | < 🔍 > |
| type | TransactionRecord | 🔍 |
| estate type | Freehold | 🔍 |
| has transaction | CFE63CDD-DDB0-42F3-AACD-00CA35248E46 | |
| property address | [2921c6b3c0e6b16ac8e0c57dff12102784728ce5] | |
| county | KENT | 🔍 |
| district | DOVER | 🔍 |
| locality | DOVER | 🔍 |
| paon | 33 | 🔍 |
| postcode | CT17 9L | 🔍 |

On This Page ?

- > 2013-11-57320
- > 2013-11-362814
- > 2013-11-361432
- > 2013-11-603801
- > 2013-11-942795
- > 2013-11-85804
- > 2013-11-734696
- > 2013-11-783386
- > 2013-11-16852
- > 2013-11-647540
- > next >

Sort by ?

- transaction id
- new build
- publish date
- transaction date
- price paid
- type
- estate type
- has transaction
- property address > county
- property address > district
- property address > locality
- property address > paon

https://permid.org/

Open PermID
BETA THOMSON REUTERS

About Contact Terms & Conditions Privacy Policy Register Login

Home Entity Search Record Matching Tagging API Entity Download What's new? How to use Open PermID

Open PermID

Put Data in Context

Record Matching

Match your own entity lists to Open PermIDs

Match

Record Matching API

Entity Search

Search for entity by name, ticker or ric

3,460,500 Organizations 240,000 Equity Instruments 1,170,000 Equity Quotes

Entity Search API

Tagging

Use Thomson Reuters Open Calais to tag your unstructured content with Open PermID

Tag

Open Calais Tagging API

UK Environment Agency

<http://environment.data.gov.uk/id/bathing-water>

DATA.GOV.UK Beta **Bathing water linked data** → go to ... page data format view

Locations of results

Powered by OpenStreetMap

© Crown copyright and database rights 2015 Ordnance Survey.

Spittal

<http://environment.data.gov.uk/id/bathing-water/ukc2102-03600>
a bathing water , coastal bathing water

| | |
|--------------------------------------|-----------------------------|
| ebwid notation | ukc2102-03600 |
| pref label | Spittal |
| water quality impacted by heavy rain | true |
| country | England |
| district | E06000048 Northumberland |
| envelope | Map bounds for Spittal |
| latest bathing water profile | 2015:1 |

European Environment Agency

http://www.eea.europa.eu/portal_relations/linked-open-data-sparql#c1=all&c3=8&b_start=0

European Environment Agency 

Search Europe's environment... [Advanced search](#) [A-Z Glossary](#) [Search](#)

Topics Data and maps Indicators Publications Media About EEA The EEA is an agency of the European Union 

You are here: Home / Possible content relations / Linked Open Data (Sparql)

Linked Open Data (Sparql)

Content-Type

Search on title






[Search](#)

all items in current results

[1] 2 3 4 5 6 7 ... 84 [next 8 items »](#)

Title starts with

All A B C D E F G H I J K L M N O P Q
R S T U V W X Y Z



This site uses cookies

We use cookies to record some preference settings and to analyse how visitors use our web site. Cookies do not contain any personal information about you. If you wish, see how to delete/disable cookies in your web browser but then our site may not work correctly. See also our privacy policy.

Korean Intellectual Property Office

<http://lod.kipo.kr/data/sparql>

The screenshot shows the SPARQL Endpoint interface. At the top, there is a navigation bar with the logo and text '특허청 | 산업재산권 LOD' on the left, and links for '서비스 소개', '데이터 브라우저', '데이터 검색', 'SPARQL Endpoint', and '어플리케이션' on the right. The main heading is 'SPARQL Endpoint'. Below this, there is a section titled '데이터셋 선택' with a 'SPARQL 생성기' button on the right. The interface includes two dropdown menus for selecting a dataset and a query, a 'Result format' dropdown set to 'HTML', and a 'Reasoning On / off' toggle switch currently set to 'ON'.

Getty Vocabularies

<http://vocab.getty.edu/>

The screenshot shows a web browser window with the address bar displaying vocab.getty.edu. The page header includes the site logo, navigation links for 'SPARQL' and 'Queries', a search bar with a dropdown menu set to 'Any', and a 'Brief' dropdown. The main content area is titled 'The Getty Vocabularies' and features a welcome message: 'Welcome to the SPARQL endpoint vocab.getty.edu'. Below this, a paragraph explains that the vocabularies (AAT, TGN, ULAN, and CONA) provide structured terminology for art and material culture. A section titled 'NOW Available:' lists three vocabularies: 'The Art & Architecture Thesaurus (AAT)', 'The Getty Thesaurus of Geographic Names (TGN)', and 'The Union List of Artist Names (ULAN)', each with a brief description. An 'Ontology Update Note' at the bottom mentions changes to the ontology and mapping. On the right side, there is a 'Tweets' section with two tweets from @seeksanusername and one from @genaortega. Below the tweets is a 'Community Support Forum' section with a 'NEW TOPIC' button, a refresh icon, a 'Mark all as read' button, and a 'Help' link. At the bottom of the forum section, there is a 'Join group to post' button.

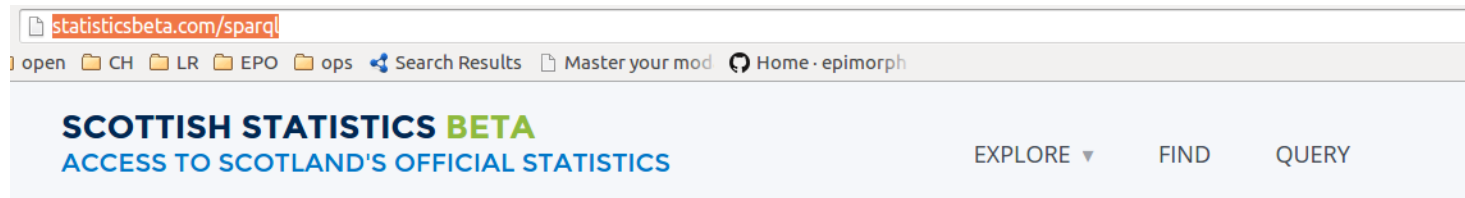
US Library of Congress

http://id.loc.gov/

The screenshot shows the id.loc.gov website. At the top, there is a navigation bar with the Library of Congress logo and links for 'ASK A LIBRARIAN', 'DIGITAL COLLECTIONS', and 'LIBRARY CATALOGS'. A search bar is also present. Below this, a breadcrumb trail reads 'The Library of Congress > Linked Data Service'. The main content area is titled 'LC Linked Data Service Authorities and Vocabularies'. On the left, there is a sidebar with a menu of links: 'Linked Data Service', 'About', 'Main Dataset Descriptions', 'Preservation Dataset Descriptions', 'Small Datasets Descriptions', 'Search', 'Download', 'Technical Center', and 'Contact Us'. The main content area features a search box with the placeholder text 'Enter Keyword or Phrase' and a dropdown menu showing 'All', 'LC Subject Headings', 'LC Name Authority File', 'LC Classification', and 'LC Children's Subject Headings'. To the right of the search box, there is a note: '**Please Note: LC Classification entries are not included in general search results. You must explicitly select LC Classification in order to search the scheme. This is temporary while the impact of adding LCC to the current system is better understood.' Below the search box are 'Search' and 'Reset' buttons. At the bottom, there is a section titled 'Available Datasets' with a list of links: 'LC Subject Headings', 'LC Name Authority File', 'LC Classification', 'LC Demographic Group Terms', 'MARC Relators', 'MARC Countries', and 'Identifiers'. A note at the bottom right says 'Schemes and smaller codelists'.

Scottish Statistics Office

<http://statisticsbeta.com/sparql>



statisticsbeta.com/sparql

open CH LR EPO ops Search Results Master your mod Home · epimorph

SCOTTISH STATISTICS BETA
ACCESS TO SCOTLAND'S OFFICIAL STATISTICS

EXPLORE ▾ FIND QUERY

SPARQL 1.1 endpoint

ENTER QUERY

```
1 PREFIX dcterms: <http://purl.org/dc/terms/>
2 PREFIX owl: <http://www.w3.org/2002/07/owl#>
3 PREFIX qb: <http://purl.org/linked-data/cube#>
4 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
5 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
6 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
7 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
8
9 SELECT *
10 WHERE {
11   ?s ?p ?o
12 }
```

Run Query

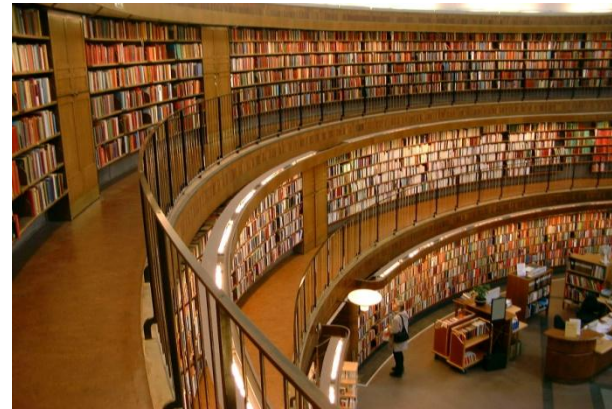
Two views of data

- Data is owned by an application



- optimised for the application
- tends only to be of use to a specific application

- Data is a collection of 'facts'



- context free
- reusable
- can be combined with other 'facts' from other sources

Questions?



www.epimorphics.com

