What is a mashup?

... and what makes them possible?

Owen Stephens
British Library

Using these slides

These slides were developed by Owen Stephens (owen@ostephens.com) on behalf of the British Library.

Unless otherwise stated, all images, audio or video content are separate works with their own licence. Such licences are stated where known

This work is licensed under a Creative Commons Attribution 4.0 International License http://creativecommons.org/licenses/by/4.0/.

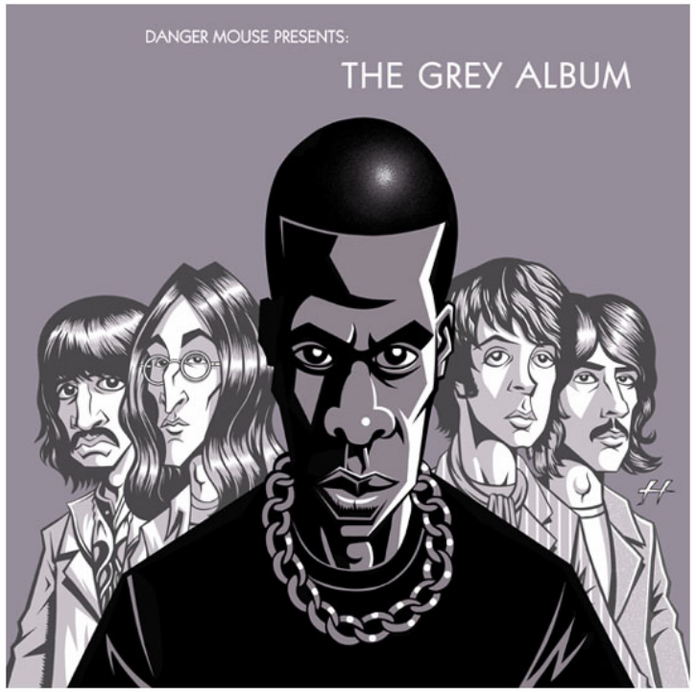
It is suggested when crediting this work, you include the phrase "Developed by Owen Stephens on behalf of the British Library"

The use of the term mashup to describe a ‘fusion of disparate elements’ dates back to at least the 19th Century, but it wasn’t until the late 20th Century that it became common, initially with specific reference to music .

In music at least, the reuse of other’s material in new contexts or new ways has a long history. Some works such as Rachmaninov’s ‘Rhapsody on a theme of Paganini’ (1934) explicitly took an existing theme and set it into a new context, using different instrumentation and playing with the basic structure of the piece to present something new.

A recording of Rachmaninov’s composition licensed as CC-BY-NC-SA is available at https://archive.org/details/RACHMANINOFFRhapsodyOnAThemeByPaganini-Rachmaninoff-NEWTRANSFER

Coal drawing of Niccolò Paganini by Jean Auguste Dominique Ingres (left) http://commons.wikimedia.org/wiki/File:NiccoloPaganini.jpeg (Public Domain)
Photograph of Sergei Rachmaninoff (right) http://commons.wikimedia.org/wiki/File:Sergei_Rachmaninoff_LOC_33968_Cropped.jpg (Public Domain)



Cover image (c) Justin Hampton (originally posted on http://www.dangermousesite.com/)

What is a mashup? “a digital media file containing any or all of text, graphics, audio, video, and animation, which recombines and modifies existing digital works to create a derivative work.”

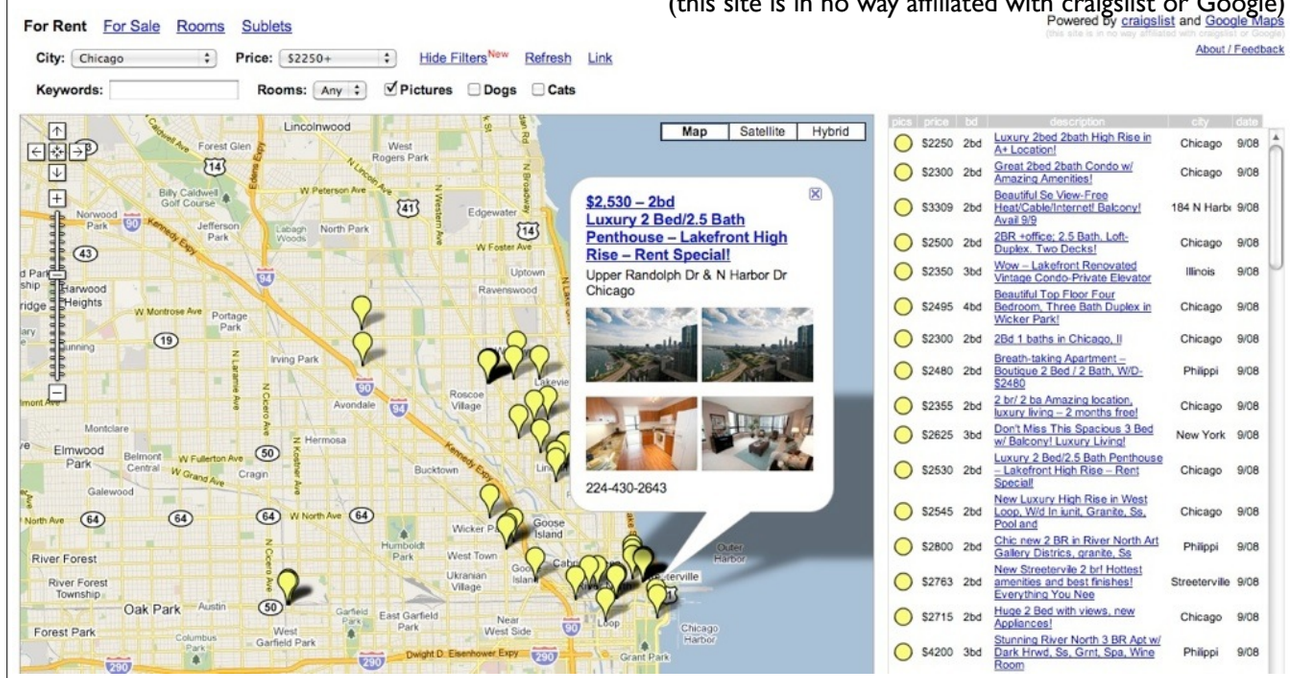
For me - creativity meets technology. While I think it is arguable that technology is not an essential part of this, I believe it is the ease with which digital content can be repurposed latter not essential but enables?

Of particular note is the ‘Grey Album’ by DJ Dangermouse, which is made up entirely of samples from the Beatles’ ‘White Album’ and JayZ’s ‘Black Album’. The ‘Grey Album’ attracted controversy and publicity as it sampled the source material without permission, and EMI subsequently requested those hosting the album to remove it due to the inclusion of material by the Beatles[2].

Mashups beyond music

http://www.youtube.com/watch?v=2T5_0AGdFic

As tools, and content in appropriate formats became more common, the number of mashups being produced increased substantially both in their musical form, and also across other media. Video mashups particularly became popular and easy to distribute or access via YouTube.



http://www.housingmaps.com/

In my view the factors that have encouraged mashups of cultural material are the availability of appropriate tools and source material in easy to manipulate formats. These factors apply equally to other forms of information, and in recent years 'data mashups' combining data from online resources to produce a new view of existing information have become increasingly common. Particularly mashups that combined data with some geographic aspect with online maps started expand rapidly from 2005 when Google Maps was first made available[4]. An early example of this is http://www.housingmaps.com/ which combined information from Craig's List (a classified advertising site) with Google Maps offer a map-based interface to find properties to buy or rent.

Google Maps has been the basis for large numbers of 'mashups' - perhaps reflect later in the day why this is?

Visualising Twitter

http://vimeo.com/4587178

Twitter/flight visualisation

And finally web mashups. Making use of the distributed and open nature of the web to bring it together information and media from different parts of the web to build applications that were greater than the sum of their parts. Perhaps the most common type of mashup we've seen are map mashups

Image: More Mixed Messages, CC-BY-SA, http://www.flickr.com/photos/withassociates

What does ‘open’ mean?

Back to 5 stars of LOD

Not a simple binary open/closed... but rather a continuum
Amount of friction...
Things that create friction:
Explicit restrictions
Uncertainty about possible restrictions
Unusual/Unfamiliar interfaces (Z39.50 anyone?)
Lack of information on data and where it is available

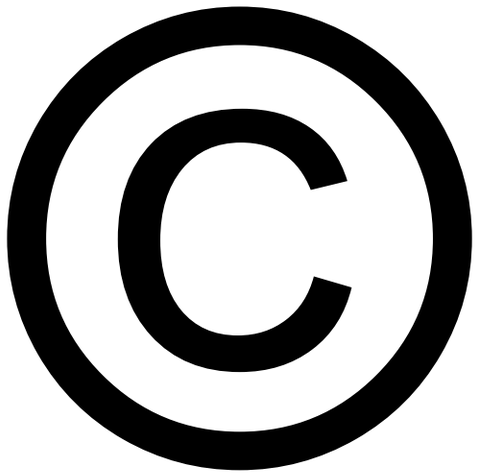
Formats - Paul Walk argues we need a ’richer understanding of openness’ which encompasses not just permissive licensing but, more broadly, the ease with which data can be used, taking into consideration aspects such as format and access mechanisms

Image: Banana Skin, CC-BY, http://www.flickr.com/photos/25414047@N00/5656079620/

Friction can be good. What is not good is thoughtlessly or accidentally introducing friction.

Frictionless means losing control (over the data) – that's ok if that's what you want. Can combine approaches by making your interface the easiest to use

Licensing



Licenses, OBD
Things that create friction:
Explicit restrictions
Uncertainty about possible restrictions

Problems with anything beyond PDDL and CC0 - does anyone own the metadata? Conflicting legal advice and decisions

Do you have database rights?

Remember the first * of LOD?

The Problem of the Yellow Milkmaid

"people simply didn't believe the postcards in our museum shop were showing the original painting"

http://pro.europeana.eu/documents/858566/2cbf1f78-e036-4088-af25-94684ff90dc5

Examples of 'open', and friction (some friction is OK?)

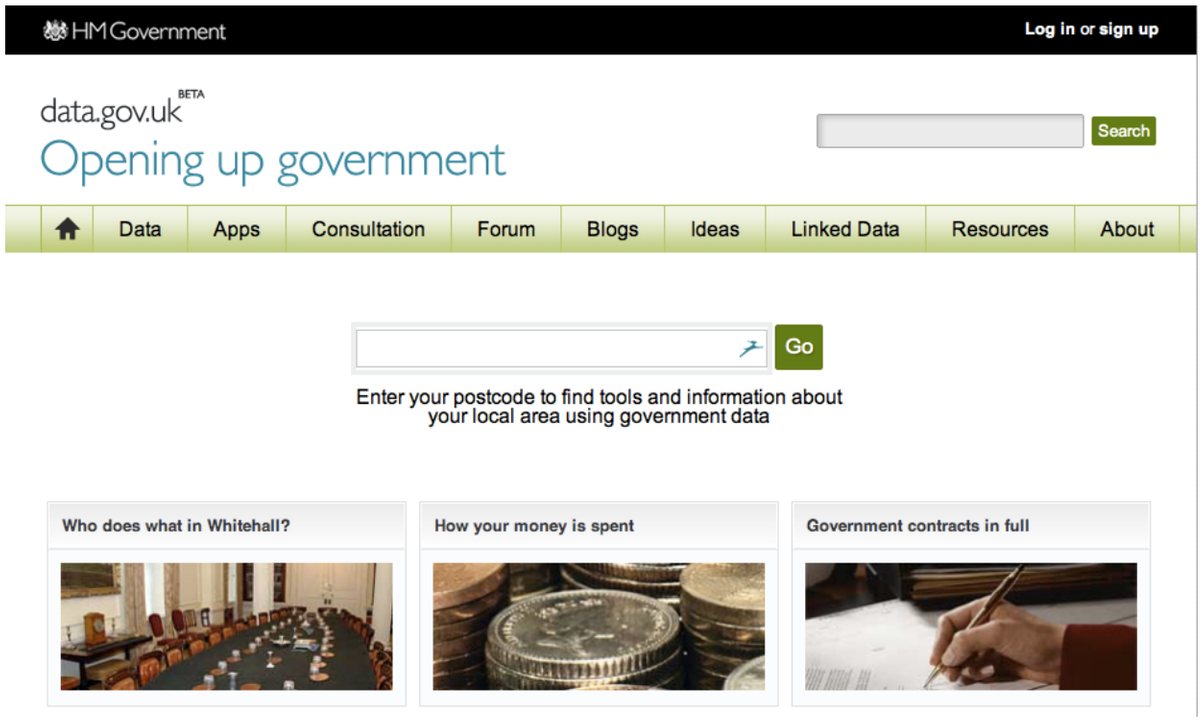
- Europeana
- BNB - originally released with an "NC" license. Moved quickly to CC0
- Problems with NC
- OCLC - compromise
- DPLA - asserts no copyright exists over metadata. But asserts to the extent it exists CC0 applies - compatibility with Europeana

'The Milkmaid', one of Johannes Vermeer's most famous pieces, depicts a scene of a woman quietly pouring milk into a bowl. During a survey the Rijksmuseum discovered that there were over 10,000 copies of the image on the internet—mostly poor, yellowish reproductions1. As a result of all of these low-quality copies on the web, according to the Rijksmuseum, "people simply didn't believe the postcards in our museum shop were showing the original painting. This was the trigger for us to put high-resolution images of the original work with open metadata on the web ourselves. Opening up our data is our best defence against the 'yellow Milkmaid'."

Image reproduced from "Europeana Whitepaper No. 2: The Problem of the Yellow Milkmaid", http://pro.europeana.eu/documents/858566/2cbf1f78-e036-4088-af25-94684ff90dc5, CC-BY-SA

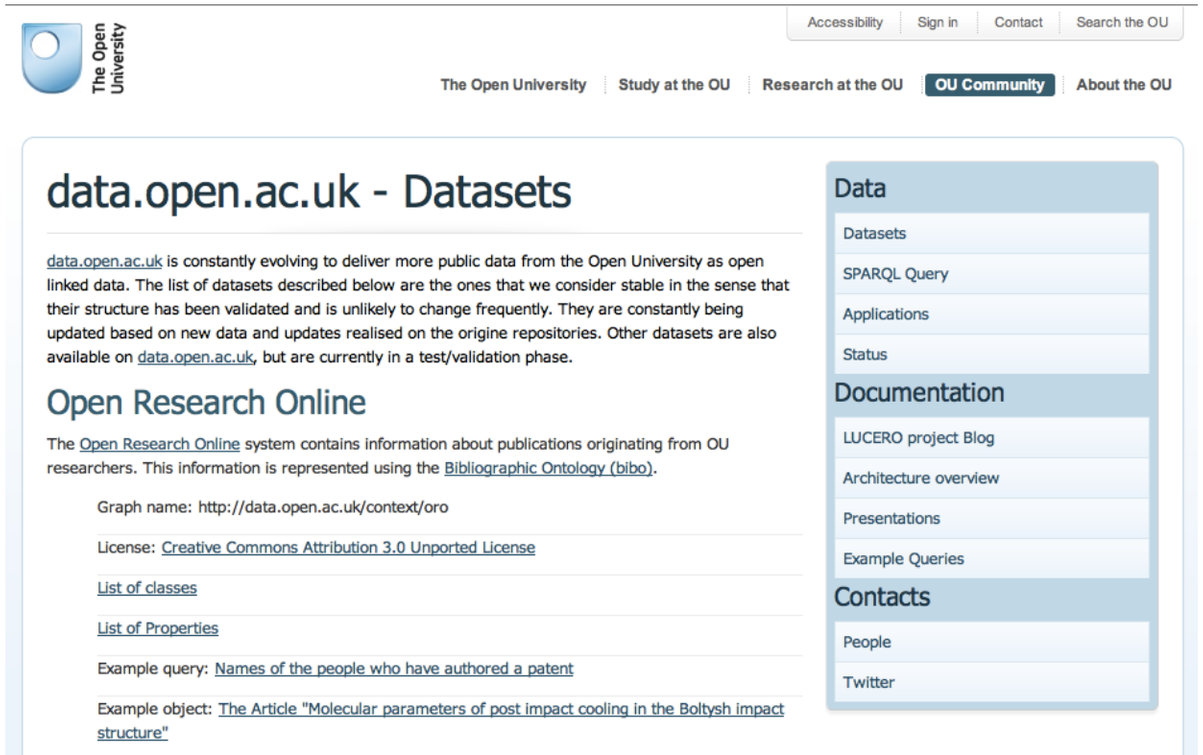
http://www.data.gov/

US Government produced catalogue of available data at data.gov



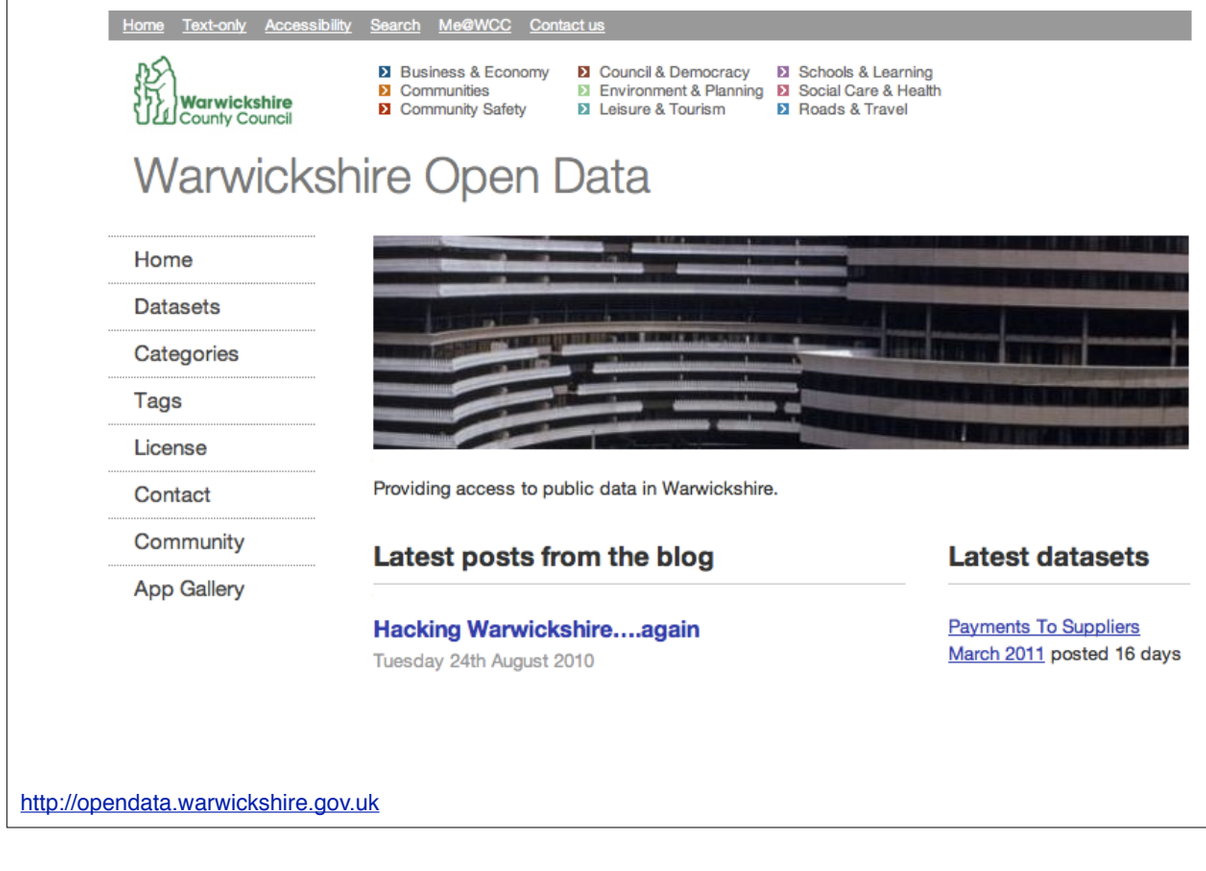
http://data.gov.uk/

UK Government data.gov.uk



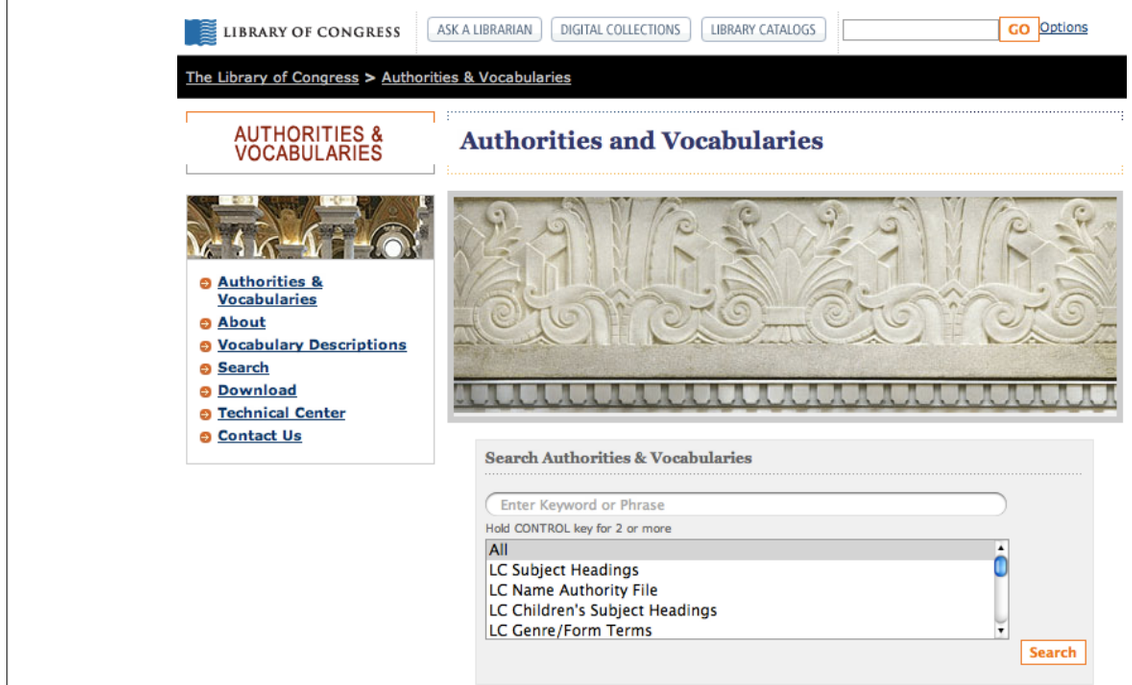
http://data.open.ac.uk

Now other organisations doing similar – Universities like data.open.ac.uk (and others)



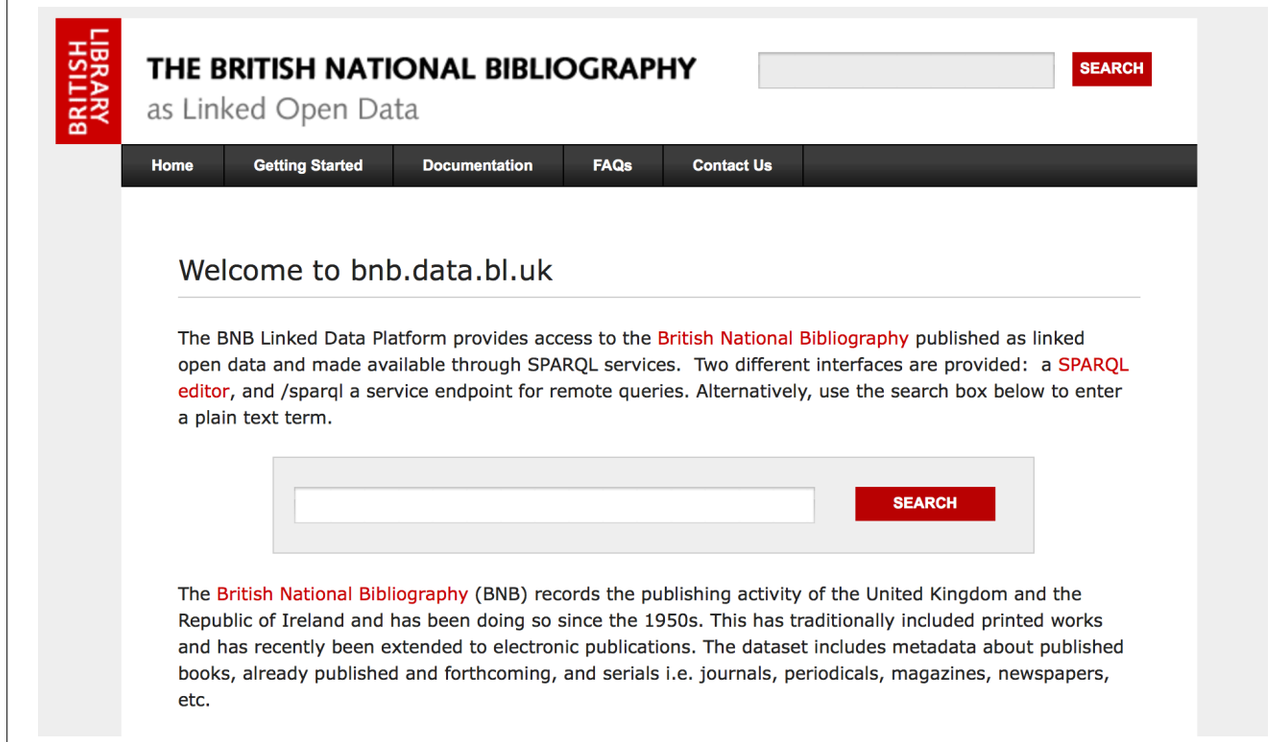
http://opendata.warwickshire.gov.uk

Local Government - eg Warwickshire



http://id.loc.gov

http://id.loc.gov



http://bnb.data.bl.uk

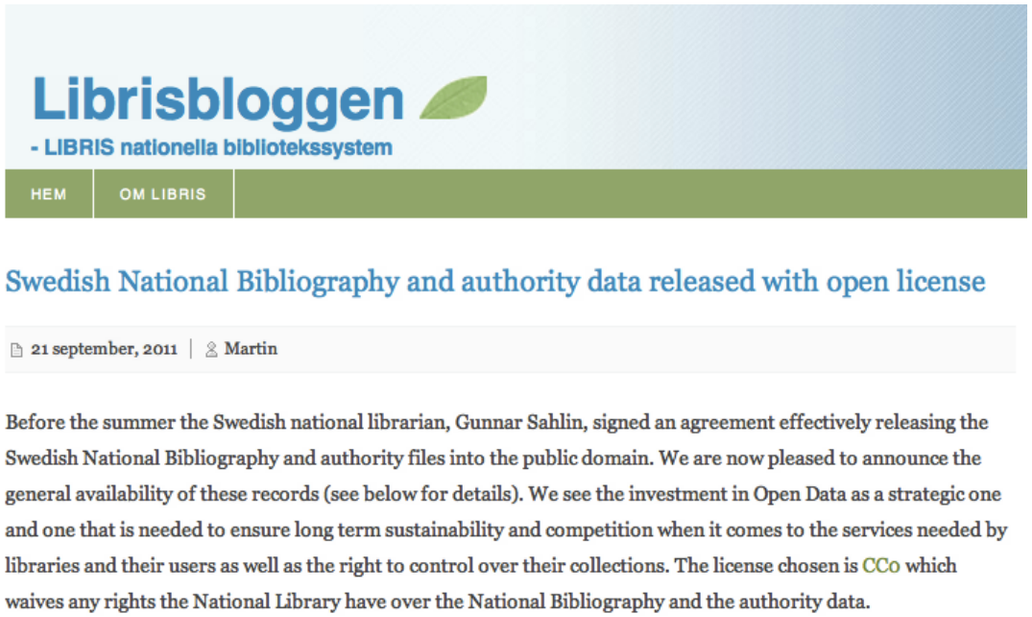
http://bnb.data.bl.uk

http://openmetadata.lib.harvard.edu/

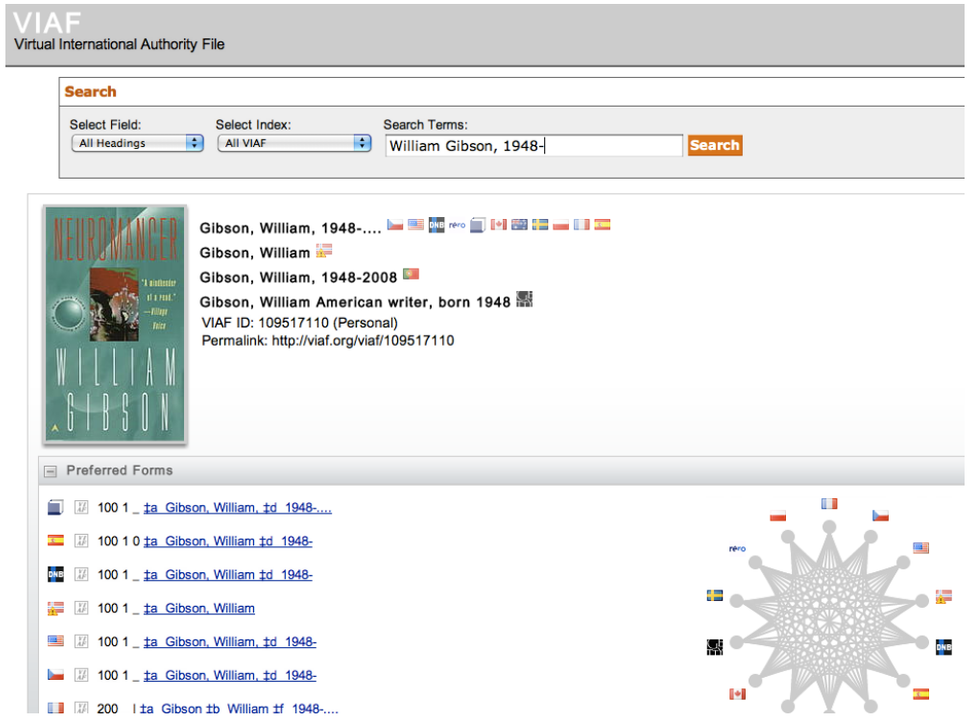
http://data.bnf.fr/

http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkeddata.html

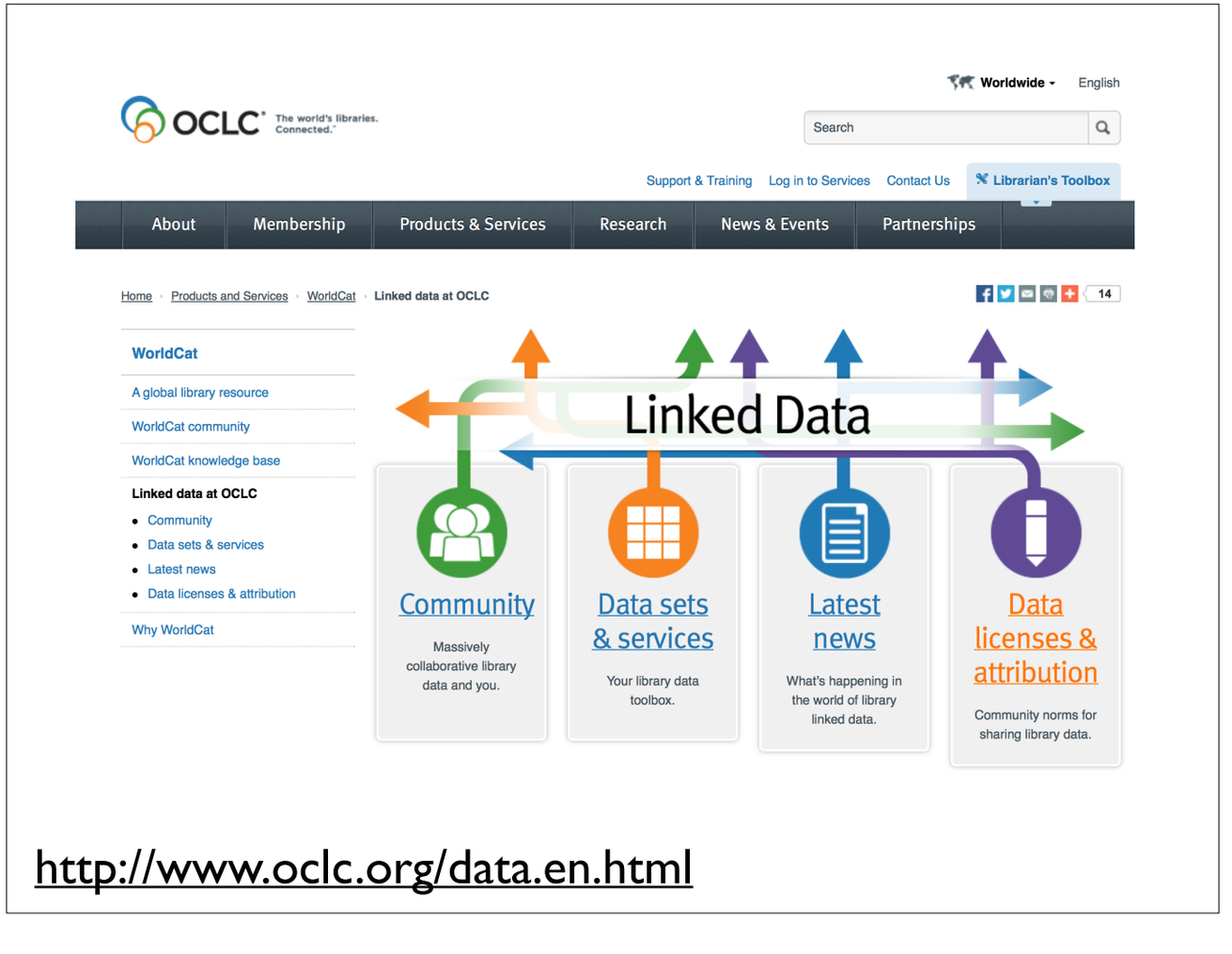
http://datos.bne.es



http://librisbloggen.kb.se/2011/09/21/swedish-national-bibliography-and-authority-data-released-with-open-license/



http://viaf.org/viaf/data/



http://www.oclc.org/data.en.html

http://www.amazon.co.uk

Not necessarily about open licensing - but being clear about T&C

Amazon, Twitter, Guardian - not open, but still used. Not always the only factor. Amazon have T&C that dictate ‘purpose’ and mandate linking to Amazon - deliberate friction that has not prevented massive uptake of the API

What is a mashup?

- a ‘fusion of disparate elements’
- an act of creation, not a technology
- enabled by making resources available through widely adopted mechanisms; in formats that enable use/reuse; and with permissive licensing