Start publishing public sector data online

Gavin Wray
14th July 2010
Keep it simple

- Make data easy to find
- Make data easy to licence
- Make data easy to re-use

[From data.gov.uk]
Where’s this coming from?

‘…we want Government information to be accessible and useful for the widest possible spectrum of people.

That is why the Prime Minister asked Sir Tim Berners-Lee to advise on how Government can best use the internet to make non-personal public data as widely available as possible.’

Stephen Timms, Minister for Digital Britain, 26 Oct 2009 [link]
Where’s this coming from?

Putting the Frontline First: Smarter Government

- Establishing public data principles
- Data available via data.gov.uk
- Free for re-use and promoting transparency
- Formation of Local Public Data Panel
- ‘Raw data now’

Gordon Brown’s speech, 7 Dec 2009 [link]

Photo: Gordon Brown, Liam Byrne and Sir Tim Berners-Lee, 17 Nov 2009
Crown copyright [view on flickr]
Where’s this coming from?

data.gov.uk beta launched January 2010
What are expectations of publishers?

‘Start with the “low-hanging fruit”. Whatever else, the raw data should be made available as soon as possible.’

Sir Tim Berners-Lee, Putting government data online, 30 June 2009 [link]
What are expectations of publishers?

‘Start with the “low-hanging fruit”. Whatever else, the raw data should be made available as soon as possible.’

Sir Tim Berners-Lee, Putting government data online, 30 June 2009 [link]

‘...ensure that any data published is made available in an **open format** so that it can be **re-used by third parties**. From July 2010, government departments and agencies should ensure that any information published includes the underlying data in an open standardised format.’

Prime Minister’s letter to Secretary of State, 1 June 2010 [link]
What are expectations of consumers?

‘It’s our data, we have a right to it’

Expectations have increased about access to data, through:

- Freedom of Information Act
- Office of Public Sector Information
- Advisory Panel on Public Sector Information
- Sites such as TheyWorkForYou.com and data.gov

Chris Taggart, OpenlyLocal.com
‘Why open up local government information’ [link]
What are expectations of consumers?

‘It’s our data, we have a right to it’

Expectations have increased about access to data, through:

- Freedom of Information Act
- Office of Public Sector Information
- Advisory Panel on Public Sector Information
- Sites such as TheyWorkForYou.com and data.gov

Chris Taggart, OpenlyLocal.com
‘Why open up local government information’ [link]
Momentum in 2010

Expectations of availability and presentation of data increasing
Momentum in 2010

Expectations of availability and presentation of data increasing

- Jan: Local Public Data Panel formed
Momentum in 2010

Expectations of availability and presentation of data increasing

- Jan: Local Public Data Panel formed
- March: Guardian Datastore - open platform launched
Momentum in 2010

Expectations of availability and presentation of data increasing

- Jan: Local Public Data Panel formed
- March: Guardian Datastore - open platform launched
- April: Ordnance Survey launch OS OpenData
Momentum in 2010

Expectations of availability and presentation of data increasing

- Jan: Local Public Data Panel formed
- March: Guardian Datastore - open platform launched
- April: Ordnance Survey launch OS OpenData
- April: West Midlands workshop on opening public sector data
Momentum in 2010

Expectations of availability and presentation of data increasing

- Jan: Local Public Data Panel formed
- March: Guardian Datastore - open platform launched
- April: Ordnance Survey launch OS OpenData
- April: West Midlands workshop on opening public sector data

- May: Tackling worklessness: good practice in data sharing (I&DeA)
Momentum in 2010

Expectations of availability and presentation of data increasing

- Jan: Local Public Data Panel formed
- March: Guardian Datastore - open platform launched
- April: Ordnance Survey launch OS OpenData
- April: West Midlands workshop on opening public sector data
- May: Tackling worklessness: good practice in data sharing (I&DeA)

- June: Local authorities must publish data on all spending over £500
Momentum in 2010

Expectations of availability and presentation of data increasing

- Jan: Local Public Data Panel formed
- March: Guardian Datastore - open platform launched
- April: Ordnance Survey launch OS OpenData
- April: West Midlands workshop on opening public sector data
- May: Tackling worklessness: good practice in data sharing (I&DeA)
- June: Local authorities must publish data on all spending over £500

- June: Prime Minister’s letter - opening government data to public
Momentum in 2010

Expectations of availability and presentation of data increasing

- Jan: Local Public Data Panel formed
- March: Guardian Datastore - open platform launched
- April: Ordnance Survey launch OS OpenData
- April: West Midlands workshop on opening public sector data
- May: Tackling worklessness: good practice in data sharing (I&DeA)
- June: Local authorities must publish data on all spending over £500
- June: Prime Minister’s letter - opening government data to public

- June: New government’s Public Sector Transparency Board set out ‘draft public data principles’
Where to start

‘Get your data out - for others to consume. Sounds simple.

Just take the spreadsheet (or similar file) that you use to track information, post it on your web site and link to it from a description posted in an accompanying web page.’

Richard Wallis, Talis
The data publishing three-step [link]
Describe and link your sources

Make data easy to find  •  Make data easy to licence  •  Make data easy to re-use

Describe and link to your sources
Describe and link your sources

Make data easy to find \rightarrow \text{Make data easy to licence} \rightarrow \text{Make data easy to re-use}

Describe and link to your sources

Source: LFS 2008

Creator West Midlands Regional Observatory
Rights Creative Commons Attribution-ShareAlike 2.0 UK: England & Wales
Describe and link your sources

Make data easy to find  →  Make data easy to licence  →  Make data easy to re-use

Describe and link to your sources

Source: LFS 2008

https://www.nomisweb.co.uk
Describe and link your sources

‘...when ONS data or National Statistics emanating from another source (eg other government departments or devolved administrations) are included in a WMRO publication, in addition to citing the source of the data, an acknowledgement that the Crown holds the copyright is also required.’

Office for National Statistics Regional Statistician, West Midlands
Describe and link your sources

‘If a “national statistic” is cited in a report, an appropriate reference/note to accompany it would be:

Source: National Statistics website www.statistics.gov.uk

Crown copyright material is reproduced with the permission of the Controller of HMSO

Office for National Statistics Regional Statistician, West Midlands
Describe and link your sources

‘Where national statistics have been transformed in some way or bespoke data have been obtained (eg from an ONS dataservice team, using Supercross within WMRO, or from the Regional Statisticians) and incorporated into tables, charts or maps produced in-house, an appropriate reference/note would be:

Source: Office for National Statistics, Annual Business Inquiry
Crown copyright material is reproduced with the permission of the Controller of HMSO
Table prepared by West Midlands Regional Observatory ’

Office for National Statistics Regional Statistician, West Midlands
Submit dataset to the Resource Catalogue on wmro.org

Regional Resource Catalogue

Title: Dataset for places vulnerable to the recession

Type: Dataset

Date Created: 2009-05

Description: This dataset highlights the regions in the West Midlands that are most vulnerable from the recession.

Geographic Coverage: West Midlands Local Authority District and wards

Time Period: 2009
Why licence public data?

‘...will potential consumers of the data be confident on their ability to use and/or reuse it?’

Richard Wallis, Talis
The data publishing three-step [link]
Why licence public data?

‘Copyright is a BIG problem - we would like to release the data under an open license (e.g. OpenDatabase or Click-Use) but can’t, and could be forced to remove data by “rights holders”’

Chris Taggart, OpenlyLocal.com
‘Why open up local government information’ [link]
Why licence public data?

Ideal, clear unambiguous examples give you confidence:

- Open Data Commons Attribution License (ODC-By)
  [http://www.opendatacommons.org/licenses/by/1.0/](http://www.opendatacommons.org/licenses/by/1.0/)
Why licence public data?

Make data easy to find ➔ Make data easy to licence ➔ Make data easy to re-use

Ideal, clear unambiguous examples give you confidence:

- Open Data Commons Attribution License (ODC-By)
  http://www.opendatacommons.org/licenses/by/1.0/

- Click-Use Public Sector Information Licence
Why licence public data?

Make data easy to find ➔ Make data easy to licence ➔ Make data easy to re-use

Ideal, clear unambiguous examples give you confidence:

- Open Data Commons Attribution License (ODC-By)
  [http://www.opendatacommons.org/licenses/by/1.0/](http://www.opendatacommons.org/licenses/by/1.0/)

- Click-Use Public Sector Information Licence

Typical unclear example - I don’t know if I can use this data:

- ‘Crown copyright’
  (No link to terms)
Examples in Observatory work

‘Adopt an existing standard of intellectual property rights. Management team have agreed adoption of Creative Commons.’

Investors in Excellence improvement plan
West Midlands Regional Observatory
Examples in Observatory work

Make data easy to find  ➔  Make data easy to licence  ➔  Make data easy to re-use

Reports

| Publisher                  | West Midlands Regional Observatory  
|                           | Level 3, Millennium Point  
|                           | Curzon Street  
|                           | Birmingham B4 7XG  
| Telephone:                 | 0121 202 3250  
| Fax:                       | 0121 202 3240  
| Email:                     | enquiries@wmro.org  
| Web site:                  | www.wmro.org

| Rights                     | Creative Commons Attribution-Share Alike 2.0 UK: England & Wales
|                           | http://creativecommons.org/licenses/by-sa/2.0/uk/

| Document contact           | Andy Phillips  
|                           | Head of Skills Research  
|                           | West Midlands Regional Observatory

Creator: West Midlands Regional Observatory
Rights: Creative Commons Attribution-ShareAlike 2.0 UK: England & Wales
Examples in Observatory work

Make data easy to find ➔ Make data easy to licence ➔ Make data easy to re-use

Datasets

<table>
<thead>
<tr>
<th>Publisher</th>
<th>West Midlands Regional Observatory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Various. The source data for each chart is stated in each worksheet</td>
</tr>
<tr>
<td>Rights</td>
<td>Crown copyright material is reproduced with the permission of the Controller of HMSO. This dataset is prepared by the West Midlands Regional Observatory. The rights to the data rest with each source quoted above each table. <a href="http://www.opsi.gov.uk/advice/crown-copyright/copyright-guidance/copyright-and-publishing">http://www.opsi.gov.uk/advice/crown-copyright/copyright-guidance/copyright-and-publishing</a></td>
</tr>
<tr>
<td>Rights</td>
<td>Creative Commons Attribution-Share Alike 2.0 UK: England &amp; Wales <a href="http://creativecommons.org/licenses/by-sa/2.0/uk/">http://creativecommons.org/licenses/by-sa/2.0/uk/</a></td>
</tr>
</tbody>
</table>

Creator: West Midlands Regional Observatory
Rights: Creative Commons Attribution-ShareAlike 2.0 UK: England & Wales
Make data easy to re-use

Get data available in basic formats that can be used by people and machines
Make data easy to re-use

Get data available in basic formats that can be used by people and machines

- CSV, XLS, XML, KML, JSON
Make data easy to re-use

Get data available in basic formats that can be used by people and machines

- CSV, XLS, XML, KML, JSON
- Google Spreadsheets
Make data easy to re-use

Get data available in basic formats that can be used by people and machines

- CSV, XLS, XML, KML, JSON
- Google Spreadsheets

Tabular data in Word or PDF files are useless

Creator West Midlands Regional Observatory
Rights Creative Commons Attribution-ShareAlike 2.0 UK: England & Wales
Use consistent identifiers

- ‘Is it open but opaque? The terms, codes, identifiers etc. you use may be meaningless, or worse still ambiguous, to those outside your organisation, or even your department.’

Richard Wallis, Talis
The data publishing three-step [link]
Use consistent identifiers

- ‘Is it open but opaque? The terms, codes, identifiers etc. you use may be meaningless, or worse still ambiguous, to those outside your organisation, or even your department.

  - Could your data be made more consistent with other data you, or similar organisations, already publish?’

Richard Wallis, Talis
The data publishing three-step [link]
Use consistent identifiers

Make data easy to find → Make data easy to licence → Make data easy to re-use

ONS Standard Names and Codes (SNAC) ids to describe wards
Use consistent identifiers

Birmingham City Council: Billesley ward

Size: 586.558 hectares
Os Snuc Ref: DCN11W
Os Gss Code: E05001181
Ordnance Survey Ref: 700000000000039
Fix My Street Reports: http://fixmystreet.com/reports/DCN11W

Councillors for this ward
- Updated Len Gregory (Conservative)
- Updated Colin Hughes (Conservative)

Neighbourhood Police Team
- Acting Sergeant James Tunnicliffe
- PCSO Arlene Milwood
- PCSO Adam Woodbridge
- PC Michael Doyle

About this Ward
Experimental
10,860 dwellings, 25,874 people live in this
Use consistent identifiers

Make data easy to find  ➔  Make data easy to licence  ➔  Make data easy to re-use
Use consistent identifiers

Make data easy to find  ➔  Make data easy to licence  ➔  Make data easy to re-use
Use consistent identifiers

Are there other consistent classifications you know of?

- Business sectors
- Other levels of geography
- ...?
Links

- Much of this presentation is adapted from Chris Taggart’s presentation on opening up local government data.
- Also indebted to Richard Wallis’ post The Data Publishing Three-Step
- Prime Minister’s letter to Secretary of State regarding transparency, 1 Jun 2010: http://www.cabinetoffice.gov.uk/newsroom/statements/transparency/pm-letter.aspx
- http://www.theyworkforyou.com/
- http://data.gov.uk
- http://www.guardian.co.uk/data-store
- Tackling worklessness: good practice in data sharing: http://www.idea.gov.uk/idk/aio/19221282
- West Midlands Local Public Data workshop April 2010: http://localdata.citizenshipfoundation.org.uk
- http://www.opendatacommons.org/licenses/by/1.0/
- http://delicious.com/gavinwray/opendata
- http://openlylocal.com
- http://www.w3.org/DesignIssues/GovData.html
Contact

- Gavin Wray
  Email gavin.wray@wmro.org
  Direct tel: 0121 202 3252

- Communications Team
  West Midlands Regional Observatory
  Level 3, Millennium Point
  Curzon Street
  Birmingham B4 7XG

- Web: www.wmro.org
  Blog: http://wmro.wordpress.com
  Twitter: @thewmro