

# Towards Open Data at DCLG

GSS Seminar: Transparency, Open Data and the  
Code of Practice, 3 April 2012

**David Fry**, Deputy Director, Strategic Statistics Division  
Analysis and Innovation Directorate, DCLG

## DCLG's main statistical themes

- We Collect and produce statistics for significant govt priorities
  - > National Accounts
  - > Planning Reform
  - > Developing Housing markets
- Provides a long-term evidence-base that is:
  - > Unique and authoritative
  - > Robust, and trusted by users
  - > Consistent and comparable (over time and across geographies)



### What?

to be an **exemplar** of transparency and open data.

Routinely publish **data that is easy to find, understand and re-use**

Adopt systems, tools and processes which ensure that data is released in a **reliable, trusted, efficient and cost-effective** manner

### Why?

Demonstrate we are **outward-facing, connected** with communities, and highly **influential** within Whitehall and the local sector

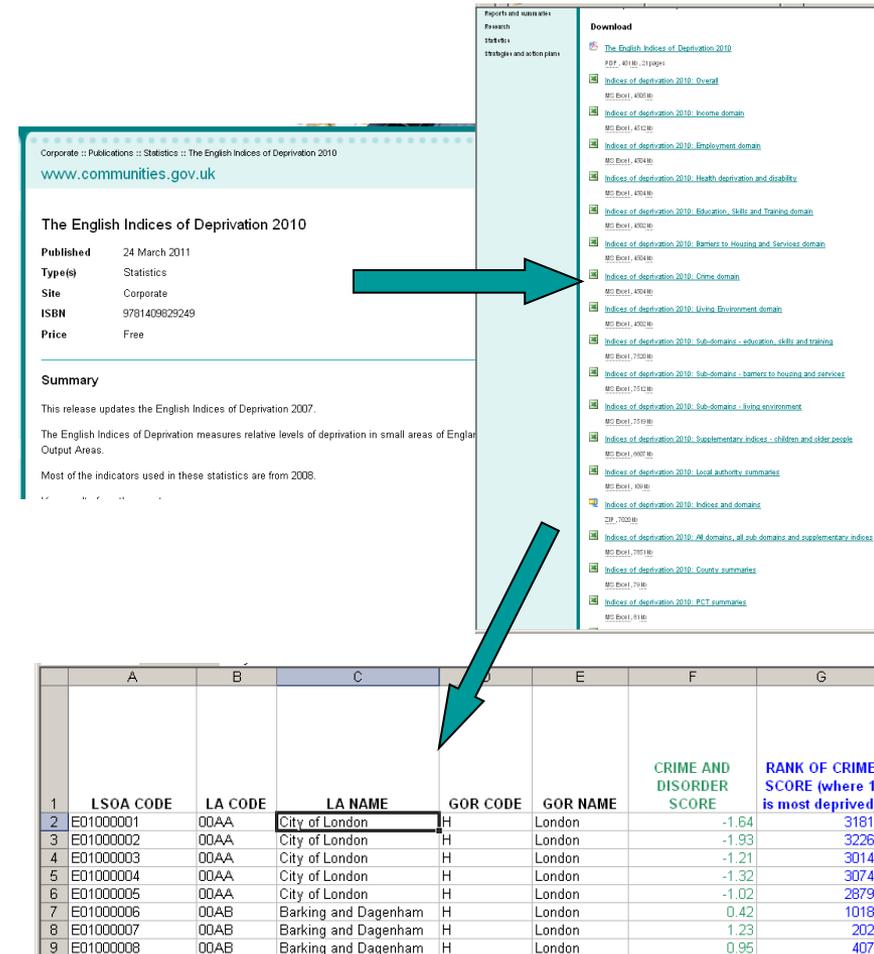
Ensure our **data is actively used alongside other external data sources**, by a **broad and diverse audience**.

- Citizens can engage more effectively with what we do, and hold us to account.
- DCLG, our ALBs and our partners exploit opportunities for efficiencies and savings. <sup>3</sup>

# Transparency and open data at DCLG: The opportunity for open data (1)

Supporting more efficient, cost-effective re-use

- From a user perspective dissemination via spreadsheets is:
  - > Inefficient
  - > Unreliable
  - > Unaffordable
- We believe Open and Linked Data could enable:
  - > More efficient, robust data publishing
  - > smarter/cheaper/faster data re-use (for DCLG itself, and our data user community).



Corporate :: Publications :: Statistics :: The English Indices of Deprivation 2010  
[www.communities.gov.uk](http://www.communities.gov.uk)

**The English Indices of Deprivation 2010**

**Published** 24 March 2011  
**Type(s)** Statistics  
**Site** Corporate  
**ISBN** 9781409829249  
**Price** Free

**Summary**

This release updates the English Indices of Deprivation 2007. The English Indices of Deprivation measures relative levels of deprivation in small areas of England Output Areas. Most of the indicators used in these statistics are from 2008.

**Download**

- The English Indices of Deprivation 2010 PDF, 4018, 21pages
- Indices of deprivation 2010 - Overall
- Indices of deprivation 2010 - Income domain
- Indices of deprivation 2010 - Employment domain
- Indices of deprivation 2010 - Health deprivation and disability
- Indices of deprivation 2010 - Education, Skills and Training domain
- Indices of deprivation 2010 - Barriers to Housing and Services domain
- Indices of deprivation 2010 - Crime domain
- Indices of deprivation 2010 - Living Environment domain
- Indices of deprivation 2010 - Sub-domains - education, skills and training
- Indices of deprivation 2010 - Sub-domains - barriers to housing and services
- Indices of deprivation 2010 - Sub-domains - living environment
- Indices of deprivation 2010 - Supplementary indices - children and older people
- Indices of deprivation 2010 - Local authority summaries
- Indices of deprivation 2010 - Indices and domains
- Indices of deprivation 2010 - All domains, all sub-domains and supplementary indices
- Indices of deprivation 2010 - County summaries
- Indices of deprivation 2010 - PCT summaries

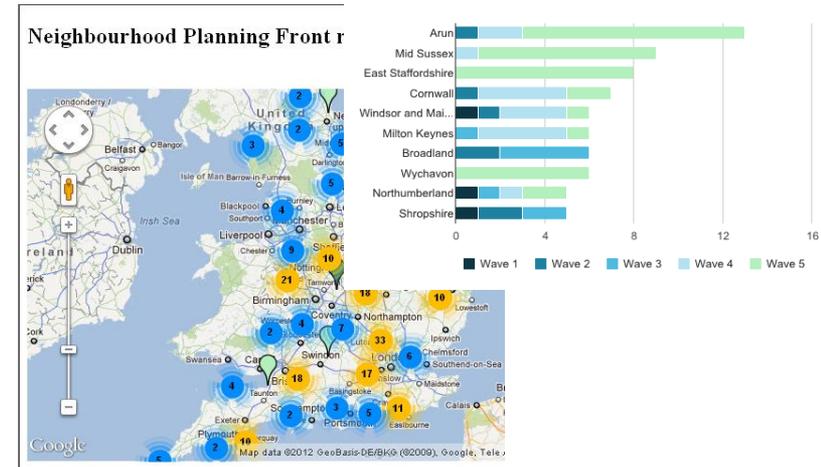
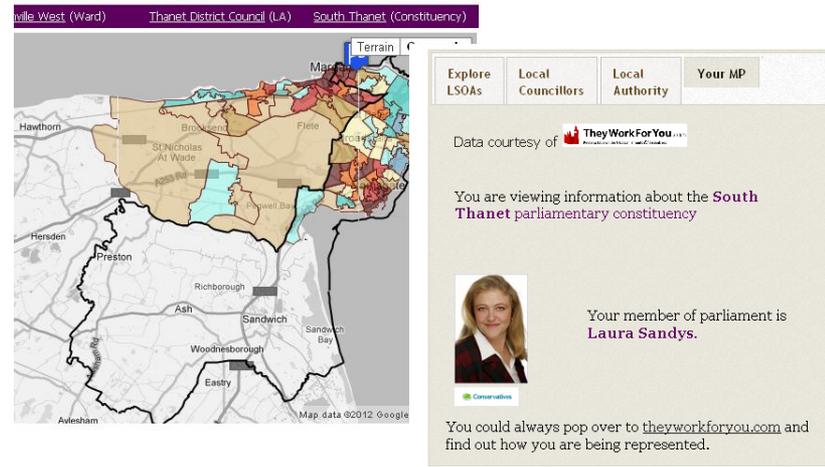
	A	B	C	D	E	F	G
	LSOA CODE	LA CODE	LA NAME	GOR CODE	GOR NAME	CRIME AND DISORDER SCORE	RANK OF CRIME SCORE (where 1 is most deprived)
1							
2	E01000001	00AA	City of London	H	London	-1.64	31816
3	E01000002	00AA	City of London	H	London	-1.93	32269
4	E01000003	00AA	City of London	H	London	-1.21	30147
5	E01000004	00AA	City of London	H	London	-1.32	30744
6	E01000005	00AA	City of London	H	London	-1.02	28792
7	E01000006	00AB	Barking and Dagenham	H	London	0.42	10184
8	E01000007	00AB	Barking and Dagenham	H	London	1.23	2026
9	E01000008	00AB	Barking and Dagenham	H	London	0.95	4074

# Transparency and open data at DCLG: The opportunity for open data (2)

Enabling effective, appropriate re-use alongside other sources

- New policies will involve more data, at more granular levels:
  - > Neighbourhood Planning
  - > Neighbourhood Community Budgets
- Statistics will increasingly be used directly alongside other sources:
  - > E.g. Neighbourhood Planning = DCLG stats + Local Authority Planning Register + community-led plans.
- We cannot (and should not) predict the full range of potential uses (or abuses!)
- Open and Linked Data provides standards and mechanisms to support/encourage “responsible re-use”.

Top 10 Local Authorities - by number of Neighbourhood Planning Frontiers

Explore LSOAs | Local Councillors | Local Authority | Your MP

Data courtesy of **TheyWorkForYou**

You are viewing information about the **South Thanet** parliamentary constituency

Your member of parliament is **Laura Sandys**, Conservative

You could always pop over to [theyworkforyou.com](http://theyworkforyou.com) and find out how you are being represented.

Ground-breaking and innovative... but data not fully open

**Step 1: Select your Place(s):**

Place selection criteria:  Standard  Advanced Post LGR April 2009

Select a Place

Search Places

UK » England

North

York

**Selected Place(s)**

1. East Midlands
2. London
3. South East

Check Available Statistics

**Step 2: Select your Statistic(s):**

Select a Statistic

Search Statistics

All » By Theme » House prices and

**Selected Statistic(s)**

Table Map Place

Group by: Statistic

**Lower Quartile House Prices (Land Registry)**

	Q2 2007	Q3 2007	Q4 2007
East Midlands	115000	115000	115000
London	205000	215000	215000
South East	164995	170000	170000

**Ratio Lower Quartile House Prices to Lower Quartile earnings**

	1999	2000	2001
East Midlands	3.38	3.52	3.72
London	4.93	5.58	6.3
South East	4.79	5.36	5.83



### Other DCLG sources

Organogram      Business plan      Budgets & spending

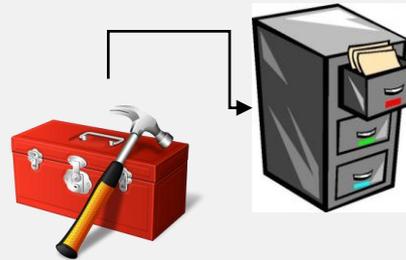
  
We work for you

**Shelter**  
Housing databank

**Devon**  
County Council  
Local Information Systems



- Ready for use now by our early adopters incl. (Shelter, and Local Authorities).
- Work from April to “show and tell” opportunities; capture benefits; build the case for further investment



Apps and linking to other sources

Open APIs (find and retrieve data in open formats)

DCLG open data filing cabinet

Data loading and conversion process and tools

DCLG inputs



- we built a trial service, using one key data set - the Indices of Deprivation
- Developed for free, by a LinkedData start-up company
- But we needed to do more to:
  - > Understand and demonstrate benefits & applications
  - > Investigate moving away from spreadsheets
- So Ran a low value (£40k) mini competitive procurement



**Open Data Communities**

Browse the Data    Execute a SPARQL query

**Datasets**

Each dataset contains values for a single indicator and a single year, with one value for each of the 32484 LSOAs in England. In addition there are supporting datasets, such as the list of all LSOAs and their locations. Each dataset has a page that explains more about the data it contains and how it's organised.

**Supporting datasets**

- Lower layer Super Output Areas
- UK Administrative Geography
- UK Postcodes

**2005 datasets**

- Estimated mid-year population 2005

**2007 datasets**

- Index of Multiple Deprivation Ranking, 2007
- Index of Multiple Deprivation Scores, 2007

**2010 datasets**

- Index of Multiple Deprivation Ranking, 2010
- Index of Multiple Deprivation Scores, 2010
- Index of Multiple Deprivation, Barriers to Housing and Services Domain, Ranking 2010
- Index of Multiple Deprivation, Barriers to Housing and Services Domain, Scores 2010
- Index of Multiple Deprivation, Crime Domain, Ranking 2010
- Index of Multiple Deprivation, Crime Domain, Scores 2010
- Index of Multiple Deprivation, Education Skills and Training Deprivation Domain, Ranking 2010
- Index of Multiple Deprivation, Education Skills and Training Deprivation Domain, Scores 2010
- Index of Multiple Deprivation, Employment Domain, Ranking 2010
- Index of Multiple Deprivation, Employment Domain, Scores 2010
- Index of Multiple Deprivation, Health Deprivation and Disability Domain, Ranking 2010
- Index of Multiple Deprivation, Health Deprivation and Disability Domain, Scores 2010
- Index of Multiple Deprivation, Income Domain, Ranking 2010
- Index of Multiple Deprivation, Income Domain, Scores 2010
- Index of Multiple Deprivation, Living Environment Deprivation Domain, Ranking 2010

Data organised in separate, but linkable datasets

Datasets include each IMD domain, plus sets for standard geographic identifiers (LAs; postcodes; parliamentary constituencies)

Each set is accessible, over the web, via a unique identifier (or URI).

### Other formats

This resource is also available as:

[RDF/XML](#)

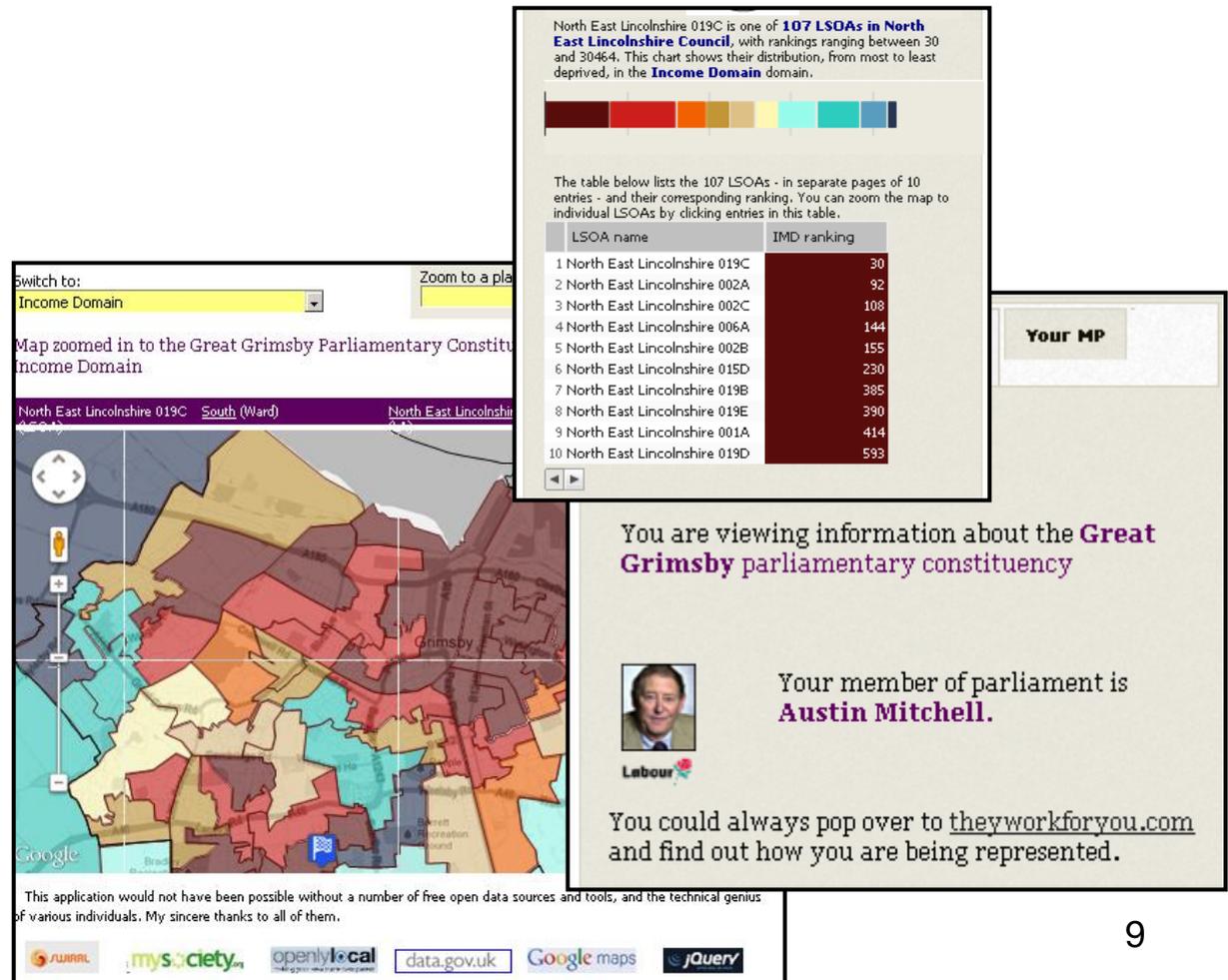
[JSON](#)

[Turtle](#)

[N-triples](#)

All data is available for direct downloading, in a choice of open, re-usable formats

- Brings together DCLG IMD and external datasets
- IMD retrieved from trial “LinkedData” store
- Also queries external sources, over the web, in real-time, incl:
  - “TheyWorkForYou” - data about local MPs
  - “openlylocal.com - data about Local Authorities and councillors
- Developed using free, open tools



North East Lincolnshire 019C is one of **107 LSOAs in North East Lincolnshire Council**, with rankings ranging between 30 and 30464. This chart shows their distribution, from most to least deprived, in the **Income Domain** domain.

The table below lists the 107 LSOAs - in separate pages of 10 entries - and their corresponding ranking. You can zoom the map to individual LSOAs by clicking entries in this table.

LSOA name	IMD ranking
1 North East Lincolnshire 019C	30
2 North East Lincolnshire 002A	92
3 North East Lincolnshire 002C	108
4 North East Lincolnshire 006A	144
5 North East Lincolnshire 002B	155
6 North East Lincolnshire 015D	230
7 North East Lincolnshire 019B	385
8 North East Lincolnshire 019E	390
9 North East Lincolnshire 001A	414
10 North East Lincolnshire 019D	593

You are viewing information about the **Great Grimsby** parliamentary constituency

Your member of parliament is **Austin Mitchell**.

You could always pop over to [theyworkforyou.com](http://theyworkforyou.com) and find out how you are being represented.

This application would not have been possible without a number of free open data sources and tools, and the technical genius of various individuals. My sincere thanks to all of them.

Logos: AURAL, mysociety, openlylocal, data.gov.uk, Google maps, jQuery

- Provides evidence base for DCLG locality management team.
- Brings together DCLG and external datasets
- Queries external sources, over the web, in real-time
  - E.g. NOMIS sources on unemployment, mid-year population estimates
- Developed using free, open tools

**Locality Management demonstrator**

Switch to a different region  
Cheshire, Greater Manchester and Greater Merseyside

Switch to a different local authority: Bolton

---

**Facts and figures**

**Cheshire, Greater Manchester and Greater Merseyside**

19 councils in this region: Bolton

**Bolton (Metropolitan District)**

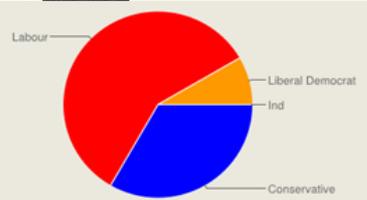
Include other councils in the locality region

Local democracy: Local election results

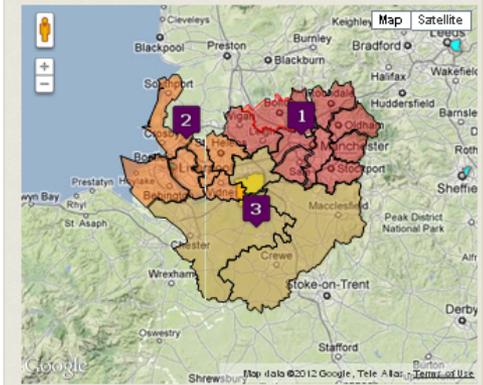
Bolton - percentage share of council seats  
Results of local election on 06-May-1911  
Proportion of seats contested: Third  
Majority party: **Lab**(Gain)

Area name	Proportion elected	Winning party	Gain or hold	Con	Lab	LibDem	Ind	other
Bolton	Third	Lab	Gain	20	35	5	0	0

Source: local election results [gov.uk/election-results](https://www.gov.uk/election-results)



**Maps**



**Bolton** is part of 1 Local Economic Partnership:

- Greater Manchester

LEPs in Cheshire, Greater Manchester and Greater Merseyside:

- 1 Greater Manchester
- 2 Liverpool City Region

## Data cabinet - example features and functionality

### Find out about a dataset

OPEN DATA COMMUNITIES

Our Data > Theme: Housing > Dataset

Dataset

### Household projections by district, England, 1991-2033

URI  
http://opendatacommunities.org/id/dataset/housing/household-projections

Data API Related

Overview  
This dataset holds estimated numbers of households per local authority district from 1991 to 2008 and projected numbers of households from 2013 to 2033.

Frequency  
Yearly

Geography by Local Authority

Years covered  
1991-2033

Created  
26 Mar 2012

Modified  
26 Mar 2012

Downloads

Further information  
The latest national statistics on household projections to 2033 for England were released in line with the Code of Practice for Official Statistics on 26th November 2010.

### Link, "in context" to dataset guidance

Further information  
The latest national statistics on household projections to 2033 for England were released in line with the Code of Practice for Official Statistics on 26th November 2010.

This Statistical Release presents National Statistics on the projected number of households in England and its regions to 2033. The figures in this release are based on the 2008-based household projections and replace the 2006-based household projections released in March 2009.

Household projections are produced by applying projected household formation rates to the household projections published by the Office for National Statistics. The assumptions for national household and population projections are demographic trend based.

Our attempt to predict the impact that future government policies, changing economic conditions or other factors might have on demographic behaviour. They provide the levels and structures that would result if the assumptions based on previous trends in the population and household formation rates were to be realised in line with the Code of Practice for Official Statistics on 26th November 2010.

This is based on the Excel spreadsheet published [here](#).

The website also provides [further explanation and commentary on this data](#) as well as [guidance and supporting text](#).

A household is defined as: one person living alone, or a group of people living at the same address who share common housekeeping or a living room.

Figures are provided in units of 1000 households.

Household projections to 2008 are based on ONS mid-year population estimates and projected rates of population growth from the 2001 Census. From 2009 onwards, household projections are based on the 2008-based household projections.

### Browse to data for a specific area and time point (akin to spreadsheet cell)

Our Data > Theme: Housing > Dataset: ... > Type: Observation > Resource

Resource

### Number of households (thousands), 2018, Wolverhampton

URI  
http://opendatacommunities.org/id/housing/household-projections/2018/wolverhampton

Predicate	Object
Data set	Household projections by district, England, 1991-2033
Label	Number of households (thousands), 2018, Wolverhampton
Number of households	104
Observation status	Forecast value
Reference area	http://statistics.data.gov.uk/id/local-authority-district/00CW
Reference period	http://reference.data.gov.uk/id/year/2018
Type	Observation
Unit multiplier	Thousands

### Link to guidance for a specific data item

Resource

### Number of households

URI  
http://opendatacommunities.org/def/housing/households

Predicate	Object
Comment	Number of households. A household is defined as one person living alone, or a group of people living at the same address who share common housekeeping or a living room.
Defined by	Vocabulary of terms for DCLG housing data
Label	Number of households
Subproperty of	Observation
Type	Measure property, Property

### Link to and explore related datasets

Household projections by district, England, 1991-2033

URI  
http://opendatacommunities.org/id/dataset/housing/household-projections

Data API Related

Related Datasets

Title	Years	Frequency	Geography
Duty owed, but no accommodation secured This dataset contains numbers of households accepted by the authority as homeless and in priority need, but for which no accommodation has yet been secured.		Quarterly	Local Authority
Estimated total households per local authority district, 2008 This datasets contains the estimated total number of households for each local authority in mid 2008.			

1. Know your users - and engage them in designing your solution
2. Understand where you want to be - focus on benefits
3. Accept it will take time, and several iterations, to get there
4. Take advantage of free help and advice from external experts
  - > The market is still very new
5. Prototypes and trials are valuable
  - > and arguably a necessity in these austere times
6. Creating Linked Data is tricky:
  - > Lack of good “off the shelf” conversion and publishing tools
  - > The complexities of statistical data (e.g. annotations, mix of time periods; mix of local and regional geographies)
  - > Processing spreadsheets - we should probably start by adapting underlying systems and sources
7. Open and Linked Data works, and is important for official statistics!
  - > Get going now, and start realising the potential.