

Technical Areas and RTD Challenges

Jeni Tennison

Linked Data advisor for data.gov.uk



Getting started

catalogues & CSVs

Photo from dfulmer <http://www.flickr.com/photos/annarbor/4350629792/>

- publishing Excel, CSVs, machine-readable data, pointers to APIs
- registry site to collect them, enable people to search
- start to find there are problems...

24														
25	Unweighted base – household crime	10,905	10,059	16,310	14,900	32,720	36,395	44,973	47,610	47,027	46,765	46,252	44,610	
26														
27	Theft from the person	434	438	680	621	604	690	584	576	574	581	725	525	-23
28	Snatch theft from person	86	79	80	83	74	88	92	71	72	80	103	64	-20
29	Stealth theft from person	348	359	600	538	529	602	492	504	502	501	622	461	-23
30														
31	Other theft of personal property	1,586	1,739	2,069	1,935	1,407	1,344	1,154	1,196	1,141	987	1,096	1,036	-50
32														
33														
34	All violence	2,074	2,556	4,176	3,593	2,728	2,714	2,320	2,349	2,471	2,200	2,114	2,087	-50
35	Wounding	508	624	914	804	648	709	577	547	578	477	466	501	-45
36	Assault with minor injury	609	784	1,356	1,198	709	623	629	572	571	492	533	428	-68
37	Assault without injury	793	966	1,567	1,257	1,015	1,079	860	918	1,002	917	844	823	-48
38	Robbery	164	182	339	334	356	303	255	311	320	315	272	335	-1
39														
40	Violence with injury	1,194	1,441	2,408	2,184	1,497	1,441	1,300	1,227	1,270	1,063	1,116	1,065	-56
41	Violence without injury	881	1,115	1,768	1,409	1,231	1,273	1,020	1,121	1,201	1,137	998	1,021	-42
42														
43	Domestic violence	292	534 ⁷	989	814	626	506	401	357	407	343	293	290	-71
44	Acquaintance	774	1,043 ⁷	1,816	1,642	862	949	828	817	845	776	691	679	-63
45	Stranger	844	797 ⁷	1,004	784	883	956	836	863	894	766	852	783	-22
46	Mugging (robbery + snatch theft)	250	259 ⁷	419	417	430	391	347	382	392	394	375	398	-5
47														
48	Unweighted base – personal crime	10,905	10,059	16,337	14,937	32,787	36,450	45,069	47,729	47,138	46,903	46,220	44,559	
49														
50	All acquisitive crime ⁸	6,418	10,009	12,148	10,587	7,642	7,394	6,129	6,047	6,040	5,540	5,977	5,427	-55
51	Household acquisitive crime	4,234	7,651	9,060	7,697	5,275	5,057	4,136	3,965	4,005	3,657	3,883	3,531	-61
52	Personal acquisitive crime	2,184	2,358	3,088	2,891	2,367	2,337	1,993	2,082	2,035	1,883	2,094	1,895	-39
53														
54	ALL HOUSEHOLD CRIME	6,947	10,410	12,426	10,562	7,879	7,592	6,645	6,632	6,923	6,282	6,583	5,939	-52
55	ALL PERSONAL CRIME	4,094	4,733	6,925	6,149	4,739	4,748	4,058	4,120	4,186	3,768	3,936	3,648	-47
56														
57	ALL BCS CRIME ⁹	11,041	15,142	19,351	16,712	12,618	12,341	10,703	10,752	11,109	10,050	10,518	9,587	-50
58	Unweighted base – personal crime	10,905	10,059	16,337	14,937	32,787	36,450	45,069	47,729	47,138	46,903	46,220	44,559	
59	1. For an explanation of year-labels see 'Conventions used in figures and tables' at the start of this volume.													
60	2. The numbers are derived by multiplying incidence rates by the population estimates for England and Wales, that is: for household crimes, by 23,525,137 households and for personal crimes, by 44,647,810 adults.													
61	3. Prior to 2001/02, BCS estimates relate to crimes experienced in a given calendar year. From 2001/02 onwards the estimates relate to crimes experienced in the last 12 months based on interviews in the given financial year.													
62	4. Estimates of the total number of households in England and Wales in 2004/05, 2005/06, 2006/07, 2007/08 and 2008/09 have been revised. Estimates of the number of household crimes for these years will differ from those previously published.													
63	5. BCS estimates from interviews in 2008/09 have been revised based on revised LFS microdata and may vary slightly from previously published estimates. See Section 8 of the User Guide for more information.													
64	6. Percentage changes for crimes such as snatch theft, robbery and domestic violence should be treated with caution because the number of victims interviewed is low (around 200 in 2009/10).													
65	7. The 1991 estimates for domestic, acquaintance and stranger violence and mugging were calculated based on the estimate for all violence. Estimates for these individual categories could not be calculated using the 1991 data as the population estimates for these categories were not collected for that year.													
66	8. It is not possible to calculate whether a change in all acquisitive crime is statistically significant. Changes in both all personal acquisitive crime and all household acquisitive crime in the same direction indicate that there is a change.													
67	9. Statistical significance for change in all BCS crime cannot be calculated in the same way as for other BCS figures (a method based on an approximation has been developed). For more information see Section 8 of the User Guide.													
68	10. See Section 5 of the User Guide for more information about the crime types included in this table.													
69	11. Figures for BCS years not presented in this table are included in an extended version of the table, available online at http://www.homeoffice.gov.uk/rds/crimeew0910.html													

Challenges

complexity & understandability

- Government data is not tabular
 - developers want tabular data they can import into a spreadsheet
 - but there's a reason government publishes in Excel and PDFs
 - annotations, provisos
 - transport and organisational networks
 - semi-structured data such as legislation
- Government data is not easy to understand
 - use of codes
 - definitions of measures
 - experts are the data owners

The London Gazette.

Published by Authority.

From Monday April 4. to Thursday April 7. 1687.

His M A J E S T I E S Gracious Declaration

To all His Loving Subjects for
LIBERTY of CONSCIENCE.

J A M E S R.

W T having pleased Almighty God not only to bring Us

We do likewise Declare that it is Our Royal Will and Pleasure, That from henceforth the Execution of all and all manner of Penal Laws in Matters Ecclesiastical, for not coming to Church, or not Receiving the Sacrament, or for any other Non-conformity to the Religion Established, or for, or by Reason of the Exercise of Religion in any manner whatsoever, be immediately Suspended; And the further Execution of the said Penal Laws and every of them is hereby Suspended.

And to the end that by the Liberty hereby Granted, the Peace and Security of Our Government in the Practice thereof, may not be indangered, We have thought fit, and do hereby straitly Charge and Command all Our Loving

More challenges

trust & distribution

- Government data is important
 - provenance matters
 - versioning matters
 - timeliness matters
- Government data is distributed
 - different public bodies produce data about the same things
 - needs to be easy to integrate
 - talk about things slightly differently
 - may publish similar information in completely different ways



Linked Government Data

data like the web

Photo from fdecomite <http://www.flickr.com/photos/fdecomite/1457493536/>

- shared identifiers for easy integration
 - including codes and column names
 - increasing transparency & reusability
- for schools, hospitals, administrative areas, ...
 - used consistently across datasets
 - providing a hub from which information can be found
- flexible data model to handle complex data
 - distributed extensibility
 - support for heterogenous requirements
- focus on incremental improvement
 - not everyone adopts at the same time
 - supports evolution
 - grow like the web



No silver bullets

remaining issues

Photo from An Nguyen Photography <http://www.flickr.com/photos/annguyenphotography/3267723713/>

- foreign mindset to government
 - tend to like to enforce solutions big bang/top down
- toolset maturity & mindshare
 - working directly with RDF isn't easy (yet)
 - developers (think they) prefer something simpler
- issues of versioning, trust, provenance still loom large

It's not enough for governments to publish data
They need to publish data **well**