

Report on Departmental Open Data Commitments and adherence to Public Data Principles for the period between July and September 2012

12th December 2012

Launched in June 2012, the Open Data White Paper committed the Cabinet Office to provide a quarterly Written Ministerial Statement on progress against the Public Data Principles as set out by the Public Sector Transparency Board. A Written Ministerial Statement was laid in Parliament today which summarised progress made in this area. Further details are set out below.

Alongside the publication of the Open Data White Paper, departments published Open Data Strategies which set out their individual programme of work towards embedding the Public Data Principles as business as usual as well as commitments on the publication of a number of datasets. Progress made in these strategies as well as the departmental commitments set out in the Prime Minister's letters of May 2010 and July 2011 on transparency and open data are also reported on as part of the evidence of progress.

This is the first of the Written Ministerial Statements on Transparency and Open Data and covers the reporting period between 1st July 2012 and 30th September 2012.

Progress made on adherence to the Public Data Principles

The Public Data Principles cover aspects of the release of data such as format, access, availability and re-use. The following provides an update on how Government has applied the principles:

Public Data Principles: Format

1. The Principles call for data to be released in open formats; that they are re-usable; machine-readable; standardised to allow comparisons to be made; and linkable. The principles use the Sir Tim Berners-Lee Five Star Deployment Scheme to rate datasets for their level of openness. The current minimum standard for the publication of government datasets through data.gov.uk is 'Three Stars', i.e. the data should be published in a non-proprietary format.
2. At present, 52% of datasets published by departments and their arm length bodies are of 'Three Stars' and above quality with the remaining datasets at 'Two Stars' or below. This can be attributed to the fact that there are many legacy datasets published. With the new direction to publish all datasets in the 'Three Star' format as minimum, we expect this figure to increase as departments move towards publishing in non-proprietary formats. The majority of these legacy datasets are in a 'Two Star' format and although not considered open by the Five Star Deployment Scheme, they are considered by some users to be in a relatively accessible format because they are not yet conversant in non-proprietary formats. However, it should be noted that due to the resources it would potentially take to convert these legacy datasets into a non-proprietary format, it is unlikely that they will be updated.
3. There are a number of datasets on data.gov.uk that have not been counted in this exercise as they are National Statistics and are therefore governed by their own standards. These datasets will remain available on data.gov.uk and the Cabinet Office is currently working with the Office of National Statistics on a solution to cater to users' needs.

4. Improvements to data.gov.uk over the summer include a new tool that allows each dataset to be automatically rated according to the Five Star Deployment Scheme. The tool also provides the ability to more accurately identify broken links so that continuity during departmental website migration remains.
5. Most datasets relating to corporate transparency are a minimum 'Three Stars' and since March 2011, all organogram publications are published to the level of 'Five Star' linkable data. There are a number of departments that additionally supply some corporate transparency data in 'Excel' format which is a proprietary format and therefore considered 'Two Star'. Some users consider 'Excel' as being more accessible because they are not yet conversant in non-proprietary formats, hence departments also catering for these users' needs. The Cabinet Office is working with departments to resolve the situation where corporate transparency data is occasionally only published in an 'Excel' format.

Public Data Principles: Access

1. Extensive work has been undertaken on data.gov.uk to ensure that all web-links provided allow for direct access to the dataset without the need for further registration or searching, although this exempts datasets in relation to National Statistics. Discussions are underway that will look to collect the necessary metadata and direct web-links from this important sector.
2. In departmental Open Data Strategies, it was reaffirmed that all datasets should be released under the Open Government Licence which allows for free access and re-use of dataset. It is apparent that there are still some datasets that are being released under different licences and that these are datasets that are held by third parties that are not necessarily government organisations but they are performing a public task. Through Transparency Sector Panels, departments are working with these third parties to either release the datasets under the Open Government Licence or revise the terms of the current licence to allow for greater access and re-use.
3. An example of this is within the Transport sector where a substantial amount of data is historically held by third parties. The Transport Sector Panel worked with these third party organisations to release valuable datasets under a more open licence than before albeit not the Open Government Licence. This new licensing arrangement has allowed for free access to train and bus timetable data.
4. The independently chaired Data Strategy Board was established in July and is tasked with advising Ministers on prioritising the public datasets with the greatest potential for growth. Its remit extends across the public sector, including the Public Data Group of Trading Funds, and it reports jointly to BIS and Cabinet Office Ministers.
5. The Data Strategy Board and Public Data Group are now both operational. In June 2012, Stephan Shakespeare was appointed as Data Strategy Board Chair and Claudia Arney as Public Data Group Chair. Heather Savory was also appointed as Chair of the Open Data User Group.
6. Work on data inventories has marginally progressed between July and September with the Cabinet Office working with central departments to look into how they can store data as well as provide access to the detail behind the datasets and how their infrastructure impacts on the central collection of data in the form of an inventory. At present, a small number of central departments are looking towards creating a flexible approach to collating inventories that can be used by all departments despite the difference in

circumstances whilst ensuring that there is a minimum standard of information available in the inventories. The draft inventories will be published on data.gov.uk as a temporary resource for users.

Public Data Principles: Availability

1. Data.gov.uk was designed as a single portal for government data and underwent a major update in June 2012 to allow for easier access to datasets and better functionality for users. Since then, data.gov.uk team has continued to improve through feedback from users and departmental administrators.
2. Recent improvements have included the tool to determine the rating of a dataset; improved Application Programming Interface that allow for easier and direct access feeds for users; and the provision of status reports of common departmental datasets.
3. Metadata provided by publishers on data.gov.uk has in the past often failed to meet a minimum standard and this has resulted in poor access and inability to re-use datasets. High quality metadata allows for more refined and useful search results for users. Improvements to the processes and understanding for adding and editing metadata on data.gov.uk, including mandatory input for certain metadata fields has helped to improve searches and the 'findability' of datasets on the website.
4. New reporting tools introduced to data.gov.uk allow for users to now look at the availability of corporate transparency data (specifically spend and Government Procurement Card data) and now provides an 'at a glance' overview of each departmental release with additional functionality allowing for users to see the timeliness of the data publication.
5. Work currently underway by the Department for Communities and Local Government will see an unprecedented amount of their datasets released as 'Five Star' linkable data. Though the procurement of a 'triple store' for the department – an essential requirement to allow storing and use of 'Five Star' data – the department will allow developers to make even more and better use out of the data published and in turn offer better information to interested communities.

Public Data Principles: Re-use

1. The increased re-use of datasets is one of the key desired outcomes of the Transparency and Open Data agenda. Despite the large number of datasets released by departments, the anticipated re-use of the data has been low up to this point. It is evident that more work needs to be carried out in a number of areas, for example, raising more awareness of the opportunities possible through the availability of free and re-usable data to the public, making more datasets free to re-use and helping foster the skills necessary to re-use those datasets.
2. The formation of the Open Data User Group has given the agenda a direct link to the external stakeholders that are generally seen as the key users and re-users of datasets, either for commercial purposes or for improving public services in general. The Open Data User Group is currently examining which public sector datasets could be prioritised for release as open data to the benefit of the UK. This data is primarily from the member organisations of the Public Data Group - currently the Met Office, Ordnance Survey, Land Registry and Companies House - but the Open Data User Group is free to advise on the release of data from other public sector organisations.

3. Since the re-launch of data.gov.uk, the website has become an active hub for the open data community to discuss issues, including the re-use of datasets. The improved online community space and feedback service has afforded faster and more direct contact with the relevant officials in the Cabinet Office and departments.
4. Transparency Sector Panels are required to include membership of non-government data users and their role is to provide challenge to departments on not only what data is released but also the format and frequency of the data, which in turn allows for the wider user community the ability to re-use the data to its full potential. Further work in this area is planned for later this year.

Completion of transparency and open data commitments undertaken by departments

As part of the July to September reporting period, departments were tasked to complete a report on the following:

1. Commitments set out in the Prime Ministers letters to departments on transparency and open data in May 2010 and July 2011;
2. Commitments set out in Open Data Strategies;
3. An overall 'openness' rating of data published on data.gov.uk up until 30 September 2012; and
4. Whether they had established a Transparency Sector Panel and whether it included a privacy expert.

Responses include progress of arms length bodies associated with the department. Details of the returns are set out below.

Completion of commitments set out in the Prime Minister's first letter of May 2010

1. Of the seventeen central departments, nine reported as 'Met', with the publication format and frequency. Of the eight departments that reported being not up to date, four reported a serious delay. A majority of the departments which have reported a delay have stated this is because their arms length bodies have not fulfilled their commitments. Other explanations for the delays are due to issues with legacy infrastructures that were never designed to produce the data now required by the new commitments. Additionally, some have sometimes struggled to publish datasets as per the commitments due to limited resources.
2. Cabinet Office proposes the following actions to resolve these issues:
 - a. To work with departments to enhance engagement of arms length bodies through workshops and good practice awareness sharing;
 - b. Through the implementation of the Government ICT Strategy, ensure new ICT systems are developed with transparency functionality as standard; and
 - c. Work with and provide departments with the means, in the form of good practice sharing and stronger guidance, to embed transparency and open data as part of their corporate process to try to reduce the resources required to fulfil the commitments.

Completion of commitments set out in the Prime Minister's second letter of May 2010

1. The second round of commitments centred on specific sectors as some key departments were tasked to publish datasets that were the most beneficial to users. These sectors were education, health, transport, crime and justice and business.

2. Of those applicable, eleven departments reported the compliance to these commitments as 'Met' and five departments reported a delay. One department reported that it was unlikely that it would meet its commitment as it was unable to release the level of transactional detail requested due to security issues. In light of this, they are supplying the data regularly but with a lesser level of granularity than other departments.
3. Cabinet Office proposes the following actions to resolve this issue:
 - a. As above, work with departments to produce stronger guidance and good practice examples to assist embedding transparency within the corporate process.

Completion of commitments set out in departmental Open Data Strategies

1. Thirteen departments reported as 'Met' or being 'On Track' for the commitments set out in their Open Data Strategies. Only four departments reported a delay with these delays being caused by a number of issues, including the resolution of privacy concerns surrounding the potential publication of datasets, machinery of government changes, resolution of data quality issues and agreement across bodies collectively publishing the data.
2. Cabinet Office proposes the following actions to learn from this issue:
 - a. Ensure the capturing of all issues surrounding decisions taken regarding privacy, data quality and business processes are used as good practice to assist in the potential future publication of similar datasets.

Completion of commitment to nominate privacy experts in Transparency Sector Panels

1. Of the seven Transparency Sector Panels already established (including the Local Government Data User Group), five have appointed privacy experts. There are sectors that have not established an external panel, in some cases because it is not deemed necessary to have one, but have appointed a privacy expert as part of their internal transparency group.
2. Cabinet Office proposes the following actions to resolve this issue:
 - a. Work with relevant departments to establish privacy experts in the remaining panels.

Table of Departmental Returns

Department	PM Letter May 2010	PM Letter July 2011	Open Data Strategy commitments	Openness Scoring
BIS	Delayed	Delayed	On Track	55%
CO	Delayed	Delayed	On Track	49%
DCLG	Met	Met	On Track	5%
DECC	Serious Delay	Serious Delay	On Track	35%
DEFRA	Serious Delay	Serious Delay	Met	44%
DCMS	Delayed	Delayed	Met	60%
DfID	Met	Met	On Track	77%
DfE	Met	Met	Met	88%
DfT	Met	Met	Delayed	51%
DH	Met	Met	Met	40%
DWP	Met	Met	On Track	40%
FCO	Serious Delay	Unlikely to be met in full	On Track	60%
HMRC	Met	Met	On Track	18%
HMT	Serious Delay	Met	Met	95%
HO	Met	Met	Serious Delay	56%
MoD	Delayed	Met	Delayed	47%
MoJ	Met	Met	Delayed	66%

Key for Departmental Returns Table: Datasets release

Met	All commitments have been met
On Track	All commitments are on schedule to be met by times specified
Delayed	Some (less than 20%) commitments are unlikely to be met within a month of the deadline
Serious Delay	More than 20% of the commitments are unlikely to be met within a month of the deadline, or some commitments are unlikely to be met within one quarter of the deadline
Unlikely to be met in full	Some commitments are unlikely to be met at all

Calculation for openness scoring category

As defined in the Open Data White Paper and subject to online analytical tool on data.gov.uk

External Sector Panel and Privacy Expert appointed category

Note: Not all departments are expected to have an external Sector Panel and therefore a nominated Privacy Expert.

External Sector Panel	The department has a Sector Panel with external data users as members ('external data users' are defined as persons working outside government or the public sector)	The department does not have a Sector Panel or the panel does not have external data users as members
Privacy Expert appointed	The Sector Panel has a member who is an expert on privacy and/or data protection	The Sector Panel does not have a member who is an expert on privacy and/or data protection

Accompanying departmental commentary

BIS	Sector Panel: BIS has a number of boards: The Data Strategy Board, The Social Mobility Transparency Board (including several privacy experts) and The Research Sector Transparency Board.
CO	The Prime Minister's Letters: Publication of expenditure data was delayed awaiting sign-off. Data now published and processes have been clarified. Openness score reflects a quantity of legacy data. Data added this year meets the 'Three Star' standard.
DCLG	<p>The Prime Minister's Letters: DCLG has published additional material not outlined in the letters including publication of all DCLG staff earning over SCS minimum, which is further than other department who currently only publish salaries at Director-level and above. DCLG has also published transparency data at a £500 threshold from April 2010, and at a £250 threshold from May 2012.</p> <p>Open Data Strategy Commitments: Of the 33 action points, only one is not currently on schedule by the times specified. This relates to the proposed publication of a staff organogram reflecting the DCLG SCS team, which is currently work in progress.</p> <p>Openness Score: Our openness score of 5% is based on the bulk of our outputs, which are 'Two Star' spreadsheets. We have around 7,000 spreadsheet files on the DCLG website, and the majority have no equivalent 'Three Star' output. DCLG is aiming to routinely publishing 'Five Star' data, across all of our outputs. The department also wishes to broaden the debate beyond formats, to also consider Application Programming Interfaces: which are key to making data fully accessible and useful.</p> <p>Privacy Expert appointment: There is currently no privacy expert on the DCLG Local Public Data Panel, chaired by Professor Nigel Shadbolt from Southampton University. However, steps are currently being taken to ensure a privacy expert is appointed by the end of the year.</p>
DECC	The Prime Minister's Letters: DECC is rated as 'Serious delay' for both Prime Ministers letters commitments. Serious delays in arms length bodies and action already underway to improve performance.
DEFRA	<p>Serious delays relating to the commitments set out in both Prime Minister's letters and relate to the challenge of compiling and providing meaningful descriptions of expenditure transactions for DEFRA's large number of arms length bodies. The newly created tool for calculating the 'Openness score' will provide valuable insights enabling DEFRA to improve the quality of information on data.gov.uk.</p> <p>Section Panel: Plans are in place to establish a Sector Panel including a privacy expert in early 2013.</p>
DCMS	DCMS will ensure that outstanding items from PM May 2010 and July 2011 letters are made fully available very shortly. The Department has not set up an external panel with a privacy expert because it holds relatively little data.

DfID	DFID are regarded as the leaders in Aid Transparency as evidenced by the Aid Transparency Index 2012 (http://www.publishwhatyoufund.org/index/2012-index/). Although no formal sector panel was established in this reporting period, we regularly included the user community in open data decisions. We intend to set up a sector panel in Q1 2013. A privacy expert is part of the core transparency team.
DfE	A newly constituted Education Transparency Board with external members will meet in early 2013.
DfT	-
DH	Openness score: The 40% figure here is derived from all DH data available on data.gov, excluding national statistics (1734 datasets) and includes data from a large number of NHS organisations where we have no direct control over format of data releases. The figure for data released under our recent Open Data commitments (PM letters and Open Data Strategy) is 72% Privacy Expert appointment: The current panel has some knowledge but a privacy expert will be appointed as part of review of membership, to reflect changes in health and care system, for the next meeting (planned for 22 January 2013).
DWP	DWP has established the Welfare Sector Transparency Board and the External Stakeholder Group (http://www.dwp.gov.uk/about-dwp/what-we-do/transparency/) to help develop Stat Xplore (the planned new publicly available tabulation tool), shape management information for Universal Credit and Fit Note data and make improvements to the Wealth and Assets Survey data.
FCO	PM Letter May 2010: The FCO is working through a significant backlog of contracts (including ICT contracts) and tenders for publication, ensuring that sensitive information is not inadvertently released. With the help of additional resources, the backlog has been reduced, but it is unlikely to be eliminated in the next quarter. The FCO is making considerable efforts to eliminate a backlog of monthly data on spend over £25,000. In the last month we have posted three months of data on data.gov.uk. We expect to eliminate the backlog in the next quarter. We are also working with our arms length bodies to ensure that they eliminate any backlogs. PM Letter July 2011: We provide an 'expense type' for every transaction. We are unable to provide a more detailed description of the scope and purpose of every transaction, for reasons of security and cost. We do not use unique reference indicators supplied by DBIS or HMRC because many of our suppliers are foreign companies or entities that do not have such indicators. Because of the nature of most of its work, the FCO does not have an external Sector Panel, though it has a Transparency Steering Group which includes a wide range of internal stakeholders. It has recruited a privacy expert to join the Steering Group.
HMRC	HMRC has met the publication requirements from the first and second Prime Minister letters. The department has developed terms of reference for the Tax Transparency Sector Board, drafted a list of stakeholders to invite to the panel and identified a privacy expert.
HMT	HM Treasury has implemented the reporting of all of the required relevant transparency datasets outlined in both PM letters however, 3 out of the 13 (23%) datasets due to be published by the Treasury were delayed by around one and four months including Workforce Management Data, GPC spend and <25K spend. These datasets will be back on track for the next quarter reporting round. Open Data Strategy: All commitments to produce new sets of data have been met except for the one to provide a consolidated statement on the use of EU funds in the UK and this is still on target for 2014. Openness scoring: Assessed all but one or two Treasury datasets can be considered as being available in 'open' formats.
HO	The Home Office has a serious delay in publishing three of the eight datasets

	<p>commitments detailed in our Open Data Strategy:</p> <ol style="list-style-type: none"> 1) Beat meeting engagement is now to be published in ONS Focus on Perceptions publication in November 2012; 2) Costs of serious organised crime: Delay due to complexity of the data and need to ensure this is robust; and 3) Disruption of organised crime: Delay due to need to reach agreement with all law enforcement partners in defining what counts as disruption and resolution of data quality issues. Work is in hand to publish this data in 2012/13.
MoD	<p>Some MoD non-departmental public bodies and Trading Funds are late (last updated 2011) in providing updated organogram and staff salary data against PM Letter May 2012 commitments. MoD central has met these commitments. Overall the delayed data sets represent less than 1% of the total coverage of MoD. GPC spend data (PM letter July 2011) is current to Aug 2012. All NDPBs, Trading Funds and MoD Top Level Budget areas have been asked to provide new data for publication in December 2012. The one dataset in the MoD Open Data Strategy that was delayed (due Q2 2012) has now been published. MoD judges it has published 31 of its datasets at 'Three Stars' quality or better. Of the 106 MoD datasets available on data.gov.uk, 41 relate to official or national statistics and are published in a set, non-open format. At Cabinet Office direction MoD has calculated its openness score as 31 datasets out of 65. The MoD Open Data Strategy stated that effort would be focused on achieving rather than exceeding 'Three Star' quality by 2014. There is currently no requirement for a MoD Sector Panel, but MoD will participate in sector panels created by other government departments.</p>
MoJ	<p>The Ministry of Justice has been delayed in meeting two of the commitments from our Open Data Strategy:</p> <ol style="list-style-type: none"> 1) To explore the feasibility of creating a 'MyPage' service containing data on individual users' interactions with MoJ services - this work is expected to be completed by the end of November 2012; and 2) To test the Open Justice microsite with members of the public - this work is expected to be completed within the next quarter. <p>All other commitments have been met or are on track.</p>