

## **GSS Transparency Sub Group**

### **Summary Report**

#### **Background**

1. The GSS Transparency Sub Group (TSG) was established in January 2011, at the request of the National Statistician, Jil Matheson, to increase the coherence and visibility of the GSS's response to the UK Government's "open data" agenda.
2. Reporting to the GSS Presentation and Dissemination Committee (PDC), the sub-group's remit was to develop a programme of work that aligned the dissemination of statistical data with the draft public data principles established by the Cabinet Office and ensured a consistent, coherent and positive response across the GSS that conformed to the Code of Practice for Official Statistics. Membership was drawn from Heads of Profession across the GSS and included a representative from Cabinet Office as well as the GESR.
3. The TSG met every two to three months throughout 2011 and early 2012 with a view to completing its work by the end of June 2012. Papers and minutes were made available across the GSS via StatNet. This summary report describes progress made by the group during the period of its existence.

#### **Achievements**

4. The Group concentrated its efforts over the period on:
  - Understanding the current position in the GSS;
  - Establishing a Policy Statement and Guidance for the GSS;
  - Creating an E-Learning Tool for GSS staff;
  - Initiating Centres of Expertise to provide advice for GSS staff;
  - Developing and Promoting GSS Case Studies;
  - Raising Awareness and Promoting Open Data; and
  - Contributing to the wider UK Government policy debate.

#### ***Understanding the current position in the GSS***

5. In order to understand the practical issues around open data technologies, the TSG received presentations on the “five star-rating system” for data formats (from the National Archives) and on the issues around converting data to specific formats such as .rdf and SDMX. This helped to ensure its work was grounded on a firm basis of understanding.

6. At the request of TSG, the ONS Methodology Advisory Service was commissioned to undertake a “stock-take” for the GSS to assess the extent to which data was already made available, preferred data formats and to identify specific concerns from within the GSS. This “stock-take” concluded that the GSS was already making 55% of its data available in so-called “three star” data formats and that some departments were already trialling four and five star data formats or had proposals that would introduce these formats. However, there were concerns about whether the skills existed, certainly consistently across the GSS, to deliver “open data” quickly. There were also reservations about the level of user demand for data in open formats.

7. A copy of the “stock-take” is available at:  
<http://www.knowledgenetwork.gsi.gov.uk/statnet/statnet.nsf/6f111bd2854bbd0280256afc00761507/b47844f4ca3bcd77802578c6003cd84a?OpenDocument>

### ***Establishing a Policy Statement and Guidance for the GSS***

8. The TSG identified early on that a positive statement of intent by the GSS to make its data widely available for reuse was important and should be set out in a policy statement that provided guidance for the GSS, setting out the basic compatibility between open data and the principles of the Code of Practice for Official Statistics, and committing the GSS to publishing its plans for those datasets that it would release over the coming years. This document also reaffirmed the GSS commitment to privacy and confidentiality.

9. Heads of Profession for Statistics were able to input into the construction of the policy statement and guidance which was then published under the National Statistician’s name. A copy is available at: [ ].

### ***Creating an E-Learning Tool for GSS staff***

10. The group noted that members of the GSS would be involved with open data initiatives at different times and to differing extents, depending on their wider departmental engagement.

11. Various TSG members undertook pilots in their own Departments to assess the skills demands that putting data in open formats would make and found that they were not excessive. There would be a need for staff to become more adept at using software data programmes and in designing and presenting data in open formats,

but that these were not difficult - or time-consuming - skills to acquire. It agreed to adopt two complementary approaches to up-skilling GSS practitioners: E-learning and Centres of Expertise.

12. In order to ensure a consistent reference point for GSS members, Ministry of Defence officials agreed to develop a prototype e-learning package that provided an introduction to transparency and open data, explaining some of the basic concepts and tools, and setting out the key messages of the GSS response in relation to privacy and data quality. This learning package has been through development since early 2012 and is now approaching finalisation, has already been available via StatNet as part of testing, and we are currently seeking a permanent host site.

13. A copy of the e-learning can be found on StatNet here: *[Please note the link will be established once the permanent hosting site is confirmed]*

### ***Initiating Centres of Expertise to provide advice for GSS staff***

14. The TSG agreed that an e-learning tool on its own was not sufficient support though for GSS staff and that formal written guidance was not necessarily the best approach in a rapidly changing technological environment. Some staff might require detailed statistical disclosure control advice or have practical technical queries.

15. To address this, the Group established three Centres of Expertise on specific issues, virtual communities consisting of members from across the GSS who could offer practical advice based on experience. These Centres of Expertise are now operational and cover:

- Technical issues - advice on linking data, creating linked files, software etc;
- Legal and policy issues - advice about maintaining privacy, undertaking privacy impact assessments, legal implications of different choices; and
- Privacy testing - testing that data is securely anonymised etc.

The names of the experts have been made available to GSS staff on StatNet.

### ***Developing and Promoting GSS Case Studies***

16. Early on, it was agreed that it was important to lead by example. The TSG looked for opportunities to put open data principles into practice as a means of identifying the pitfalls and developing skills. During 2011 and 2012 the following case studies were developed:

- Consumer Price Index (ONS) - Underlying CPI data is now routinely released alongside the statistics that they underpin. This data has been disclosure

controlled, and there is anecdotal evidence to suggest that the private sector is using it to develop new economic forecasting models;

- Court Sentencing Data (MoJ) - underlying data is now routinely made available in open, reusable formats. The data is based on information that is a matter of public record, but has been anonymised;
- Reoffending Data (MoJ) - underlying data is now made available in open, reusable formats, but has been disclosure controlled. To test whether it was possible to identify an individual offender through the mosaic effect, MoJ invited students to attempt to unpick the data comparing it with public sources of information. The results of this exercise were important in helping decisions about how to present the data and what variables to include in the public release;
- Tabulation Tool (DWP) - developed in-house tools that enable data to be analysed and reused in new ways, including the development of mapping tools to help in visualisation and mapping, are being applied to data drawn from administrative sources;
- Index of Multiple Deprivation (DCLG) - outsourced the development of tools to allow IMD datasets to be accessed directly through the internet. This data has its own UL+RL, and this means that it can be easily linked to external datasets to provide scope for richer analysis.

### ***Raising Awareness and Promoting Open Data***

17. The Group wanted to ensure that basic messages were promulgated at the working level to GSS staff at all levels, in a consistent way. This was achieved by briefing Heads of Profession and through a TSG-organised GSS seminar which explained Open Data and the GSS response. This all-day event was highly successful with excellent feedback from attendees. It was attended by over 100 people and included a short talk on the importance of our data from the Chair of the TSG, a session from the Cabinet Office setting the policy scene, talks showing what the GSS had already achieved, practical case study sessions, a demonstration of the e-learning tool and an overview of statistical disclosure.

18. A key issue that emerged from the day was the importance of Privacy Impact Assessments when deciding what data to make available.

### ***Contributing to the wider UK Government policy debate***

19. Throughout its existence, the TSG Chair, members and secretariat remained in close contact with policy officials at the Cabinet Office. The Chair, and a representative from the National Statistician's Office, were invited to and gave a presentation to the Transparency Board about Open Data and Official Statistics.

20. The group, primarily through its secretariat, has contributed to the development of Open Data policies and has shown the GSS to be committed to the broad direction of the agenda, subject to the guiding principles of the Code of Practice for Official Statistics. In particular it has:

- reached consensus that the existing arrangements for accessing Official Statistics through the internet, supplemented by a link from data.gov.uk to the Publication Hub, will satisfy the Cabinet office objective of a single portal for accessing data;
- reached consensus that underlying data should not be released in open formats before the release of the statistics that they underpin;
- discussed with Cabinet Office the importance of statutory prohibitions on disclosure in determining which data can be made available and influenced thinking about the way forward;
- shown that there are already in place mechanisms that allow access, in controlled conditions, to disclosive statistical data and that statistical data is not as “closed” as some had assumed;
- commented on drafts of policy documents and the “right to data” white paper, including helping with case studies.

## Communications

20. The Group recognised the need to promote and encourage a positive response to the open data agenda from the GSS and provide reassurance that, although it provided some challenges for the future, there were no fundamental conflicts between Open Data and the Code. The chair of the TSG, Guy Goodwin, wrote to HoPs in the summer of 2011 setting out the proposed GSS response to open data and encouraging the GSS to move quickly towards a general principle of making data available in “four-star” open formats, which TSG understood to be a reasonable target during the coming period. Papers and minutes were made available across the GSS via StatNet.

## Next Steps

21. The TSG recognised the importance that the formal closure of the group in June 2012 should not stop momentum across the GSS. In particular, the following will need to be considered over the coming period, resulting from the TSG’s initiatives:

- **Centres of Expertise:** Over time, the composition of the centres of expertise will need to evolve. We **recommend** that the Chair of GSS PDC is responsible for ensuring the future of the Centres, using StatNet to publicise their existence and achievements. Draft Terms of Reference for these Centres are attached as Annex B to this paper.

- **E-learning:** the e-learning package developed by MoD for TSG will be finalised in the coming weeks. Two issues need to be resolved, and we **recommend** that NSO take these forward:
  - identifying an appropriate host for this package. For technical reasons, StatNet is not suitable in its current form. Discussions with Civil Service Learning are underway with a view to seeing if they are prepared to host the package, which is relevant to other groups of civil servants;
  - Responsibility for keeping the package up to date and ensuring the content is refreshed from time to time to ensure it maintains its relevance - it should be reviewed next towards the end of 2013.
  
- **Case Studies:** it is important that the bank of case studies remain relevant and up to date, with new studies being shared to show how to adapt to changing technologies and new situations. The Chair of GSS PDC, through the Centres of Expertise, will be well-placed to identify the potential for new case studies and we **recommend** that the Chair assume this oversight role.
  
- **Promoting Open Data:** the Open Data agenda will continue to develop and to raise new issues for the GSS. To ensure that the GSS remains on the front foot, we **recommend** that GSS PDC should have Open Data as a standing item on its agenda for the next few years. We also **recommend** that Open Data Updates should feature on the agendas of the annual “Statisticians in Government Conference” for the next two years.
  
- **Contribution to the Policy Debate:** links with the Cabinet Office, and the governance arrangements around the Open Data agenda, will continue to develop and it is important that the GSS continues to be closely involved, so as to avoid being sidelined. We **recommend** that the National Statistician’s Office should retain the responsibility for this, nominating individuals from across the GSS to fulfil specific roles as these arise.

James Edmonds

Secretary, GSS Transparency Sub Group

June 2012

## **Annex A: List of Members of the GSS Transparency Sub-Group**

Guy Goodwin (Chair)

David Fry (DCLG)

Iain Bell (MoJ)

Anthony Coyne (HM Treasury)

Ganka Mueller (GESR)

Prabhat Vaze (GES)

Daryl Lloyd (DfT)

Jude Hillary (DfE)

Pam Davies (UK Statistics Authority)

Janet Giles (ONS)

Colin Bowler (ONS)

James Edmonds (Secretary)

Cabinet Office ERG also had an open invitation to join the meeting.

## **Annex B: Centres of Expertise: Terms of Reference**

### **Scope**

Centres of expertise provide advice to practitioners who are actively moving towards transparency releases, sharing knowledge and experience. They support departments in identifying the full range of issues that need to be considered in developing data releases under the open data agenda. Their aim is to ensure that practitioners have the information and advice they need to ensure that their releases of statistical data promote the open data agenda and conform to the Code of Practice for Official Statistics.

### **Technical centre of expertise: Terms of reference**

To provide practitioners across the GSS with specialist knowledge of technical issues (such as: open data star ratings and choice of data format, converting files into open formats, data linking, data security etc), based on experience within Government, to assist them in making decisions on releasing data under the open data initiative.

### **Legal and Policy centre of expertise: Terms of reference**

To provide practitioners across the GSS with specialist knowledge of legal and policy issues, (such as: implications of statutory prohibitions on disclosure and the Data Protection Act; how to conduct privacy impact assessments; and the legal implications of the technical solutions adopted) based on experience within Government, to assist them in making decisions on releasing data under the open data initiative.

### **Privacy testing centre of expertise: Terms of reference**

To provide practitioners across the GSS with specialist knowledge of the development and application of privacy and disclosure control issues (including how to test data to ensure that it is securely anonymised before release; how to prevent disclosure through the jigsaw affect) based on experience within Government, to assist them in making decisions on releasing data under the open data initiative.

### **Members for the Centres of Expertise**

Members of the Centres of Expertise will occasionally present at GSS workshops on the work they have done and should be identified by their Head of Profession. Members will usually be:

- at SSTO/SEO/G7 level or above
- have been confirmed by their Head of Profession that they have
  - o significant knowledge in this area
  - o have undertaken a major piece of work in the area
  - o have some time available to answer phone queries from people across the GSS.