STATISTICS DENMARK

UN Global Working Group on Big Data for Official Statistics

4th UN Conference on Big Data, Bogota, 8-10 November 2017

Concept Note

Background and motivation

The UN Global Working Group on Big Data for Official Statistics (GWG) was created in March 2014 with the purpose to address the competition and potential threat of indicators based on large amounts of real time "Big Data" released by commercial providers. This could make the traditional, not-so-flexible and not-so-timely official statistics obsolete.

The GWG was asked to investigate the benefits and challenges of the use of Big Data for official statistics in general and for the monitoring of progress on the Sustainable Development Goals in particular. The GWG needed to address issues of methodology, quality, technology, access, legislation, privacy, data management and finance, and provide cost-benefit analyses of the use of Big Data in the production of official statistics.

The GWG responded to these requests, among others, by making a Big Data project inventory, which shows both the benefits and challenges of the use of Big Data, and having Task Teams work on access, privacy, partnerships, communication, training and skills, and specifically on methodology, technology, access and training in the areas of mobile phone, satellite imagery and social media data.

Future direction

After three years of work, some refection on the future direction is needed. Big Data turned out not to be an imminent threat to official statistics. Big Data can be treated as additional data sources with some peculiar characteristics, and the statistical community gained some experience in using Big Data in the production of official statistics. Access to proprietary data (including Big Data) was moved from a research item to an item on the work programme of the revision of the national quality assurance framework. Experience was gained on building partnerships with private sector and academia, and the work on satellite, mobile phone and social media data is expected to soon lead to regular capacity building programmes for official statistics.

Whereas the GWG has made good progress, some questions and challenges remain. Official statistics needs to make further progress in delivering timelier, more frequent and more granular data; the private sector remains a formidable competitor in that respect, and the statistical community needs to forge further engagement with those new partners to combine the strengths on both sides.

This means that additional steps need to be taken now by the GWG to ensure that results from the work of the existing Task Teams become usable for the statistical community; that Big Data, administrative data and with traditional statistical data sources are really treated together via a multisource approach; that the challenge is <u>not</u> to overcome problems of using Big Data in itself, but to achieve trusted data collaboration; and that in that respect the GWG needs a collaborative Global Platform for the overall organization of its work.

In short, the strength of the community of official statistics is trusted data, and the real need of the community is close collaboration with private sector, academia, research community and civil society.

Conference theme: *Data collaboratives and trusted data*

Data collaboratives are a new challenge and new opportunity for the community of official statistics – in relation to Big Data, to the SDGs, to the sharing of data, services, technologies, and know how.

- How can statistical offices, technology companies and data owners collaborate in a mutually beneficial way in a changing world, in which data are seen as the most important source for creating wealth and development for all?

- What are the experiences and lessons learned from existing data collaboratives in relation to coverage, inclusion (and exclusion) of partners, activities, management and financing?

- How can we share micro-data and other sensitive data in a federated cloud environment given regulatory frameworks for data privacy and statistical laws protection confidentiality?

- How could we effectively and collaboratively use modern tools and services, such as data lakes, integrated geo-spatial data and statistics, or open source elastic stack while adapting job profiles and skills sets in the statistical office?

These and similar questions will be addressed at the 4th UN Conference on Big Data in Bogota, Colombia from November 8th to 10th – organised in collaboration of the UN Global Working Group, DANE and the Colombian Ministry for ICT.