United Nations E/CN.3/2015/4

Economic and Social Council

Distr.: General
12 December 2014

Original: English

Statistical Commission
Forty-sixth session
3 – 6 March 2015
Item 3(a) (iii) of the provisional agenda*
Items for discussion and decision: data in support of the post-2015 development agenda: big data

Report of the Global Working Group on Big data for official statistics

Note by the Secretary-General

The Commission has before it the report of the United Nations Global Working Group on Big Data for Official Statistics. The report presents the highlights of the international conference on Big Data for Official Statistics, the outcome of the first meeting of the Global Working Group, and the results of a survey on the use of Big Data for official statistics. The report elaborates on the terms of reference and the proposed programme of work of the global working group, covering the following areas: training, skills and capacity building; linking Big Data and the sustainable development goals; advocacy and communication; access and partnership; and cross-cutting issues, such as classifications and frameworks; and covering also the exploration of specific Big Data sources for official statistics, namely mobile phone data, social media data and satellite imageries.

Points for discussion are contained in the last section of the report.

* E/CN.3/2015/1.

I. Introduction

1. Because of wide and constant use of telecommunication and other devices driven by innovations in technology, digital information is continuously generated, such as data from GPS devices, from automated teller machines, from scanning devices, from sensors, from mobile phones, from satellites or from social media. These high volume, velocity and variety of data, which require new tools and methods to capture, manage, and process them in an efficient way, are commonly referred to as Big Data, and are potentially useful for official statistics. However, many statistical offices will still need to make the business case for this investment in innovation by showing why and how Big Data are useful and relevant. The International Conference on Big Data for Official Statistics – discussed in more detail below – was encouraging in this respect, as it showed good examples of many ongoing Big Data projects.

2. The potential of Big Data sources resides in their timely (sometimes real-time) availability of large amounts of data, which are usually generated at minimal cost. Traditional data sources, such as household and business surveys, take time to conduct and are often costly in the data production. Big Data could supplement, reduce or replace such data collections. However, the statistical community is conscious of the fact that in order to be able to take advantage of these innovative data sources, including its application for SDG monitoring and reporting, it needs to adequately address issues pertaining to methodology, quality, technology, data access, legislation, privacy, management and finance, and present the adequate cost-benefit analysis before introducing Big Data into official statistics.

3. Within this context, the 45th session of the Statistical Commission in 2014 recognized that Big Data constitute a source of information that cannot be ignored and needs to be evaluated on its merits. To achieve this, the Commission supported[1] the proposal to create a global working group (GWG) on Big Data for Official Statistics, which would make an inventory of ongoing activities and examples on the use of Big Data, address concerns related to methodology, human resources, quality and confidentiality, and develop guidelines to classify various types of Big Data sources.

4. The Commission stressed that the terms of reference and the mandate of the group should be based on strategic considerations, in particular with links to the Post-2015 Development Agenda, the Data Revolution initiative and the Fundamental Principles of Official Statistics. The group should also complement and build upon work carried out by regional commissions and other international statistical agencies, and the programme of work should pay special attention to the circumstances of developing countries, such as the legal frameworks or some disadvantages in available IT infrastructure.

5. Following this decision of the Commission, the GWG was created in May 2014 and met in person for the first time on 31 October 2014 in Beijing, China, immediately after the international conference on the use of Big Data for Official Statistics co-organized by the United Nations Statistics Division (UNSD) and the National Bureau of Statistics of China (NBS China). Moreover, UNSD and the United Nations Economic Commission for Europe (UNECE) jointly conducted a survey on Big Data projects and their organizational setting, as input and preparation to these meetings.

[1] See Decision 45/110

6. Section II of this report provides an overview of the international conference followed by a summary of the first GWG meeting in section III. Section IV provides the results of a survey on Big Data projects and their organizational context, and section V presents the points for discussion.

II. International Conference on Big Data for Official Statistics, 28-30 October 2014, Beijing, China

7. UNSD and NBS China jointly organized a three-day international conference on Big Data for Official Statistics from 28 to 30 October 2014 in Beijing. The conference was attended by about 120 participants from more than 40 countries, from various regional and international organizations, as well as from the research, academic, and the private sector communities. The conference focused on three clusters of Big Data sources namely (i) mobile phone, GPS and other tracking devices, (ii) satellite imaginary and other geo-spatial information, and (iii) twitter and other social media. In relation to each of these sources, the conference discussed their typology, their specific challenges and the partnerships necessary to exploit them. The final sessions of the conference discussed common benefits and challenges of Big Data sources, ways of introducing innovation and how to make the business case for Big Data.

8. The considerable number of projects presented at the conference showed the innovative spirit in the statistical community. Applications of various Big Data sources to a wide-range of statistical domains, demonstrate that Big Data really has the potential to improve official statistics. These innovative tools can add to the portfolio of a statistical system complementing existing applications or

providing more flexible, short-term solutions to highly relevant policy questions. The different projects also demonstrated how some of the challenges related to methodology, access, privacy and skills can be addressed. The next paragraphs present some of the highlights of the conference.

9. Mobile phone records were explained in their technical detail, including the differences between active and passive mobile phone location data. Several interesting statistical applications and national experiences were shown, such as for tourism and day time mobility statistics and for estimation of population census data, poverty mapping and tracking of mobility patterns in case of disease outbreaks. Given the widespread use of mobile devices, including in the developing world, mobile phone data have great potential to provide real-time, low-cost information on pertinent development issues. The key challenges remain the protection of confidentiality, access to mobile phone data and public trust in its use.

10. Satellite imagery has great potential, especially for agriculture statistics, to provide more frequent and timelier data, at a very disaggregated level; however, the estimation methods, for instance for estimating crop yields, are still being tested. The ongoing work of ABS on satellite imagery data could lead to complementing and partly replacing the existing surveys for measuring agricultural crop production. Satellite images are available at a frequency of once every two weeks and could reduce the frequency (and associated costs) of surveys. Similar applications of satellite imagery for official statistics are being explored in Colombia, China, Mexico and other countries, including for testing and experimentation of ecosystem accounting.

11. Twitter, Facebook and other social media sources are arguably the largest datasets on human behaviour and the statistical community is already exploring these for data-driven social science applications, for instance, on health related issues. The example of the Netherlands showed a very promising example of obtaining consumer sentiment estimates from Facebook and Twitter data, which may potentially reduce the need for surveys and provide early estimates, since the social media estimates can be produced with a higher frequency and reduced costs. In Italy and China, web scraping tools are used for estimating job vacancy rates, which can support current labour statistics by providing improved monthly predictions and finer territorial estimates. The statistical community is making strong progress in this area, even if there are many methodological challenges.

12. The commonalities across data sources were also discussed, in particular in terms of methodology and quality concerns, but also regarding privacy, partnership and IT. It was recognized that each of the different Big Data sources encounters all challenges in one way or another, and that there is an immediate need to get to know those in all detail. Once the challenges of the Big Data sources are better understood, the statistical community can leverage their commonalities. For instance, templates of umbrella agreements for access to data with globally operating companies could be established irrespective of the particular data source. The statistical community should collaborate rather than compete with the private sector, in order to advance the potential of Big Data for official statistics. At the same time, it should remain impartial and independent, and invest in communicating the advantages of exploiting the wealth of available digital data to the benefit of the society. Building public trust will be the key to success.

13. Finally, the current and future roles of the community of official statisticians were discussed. For example, official statisticians may be asked more frequently to validate information provided by the market. It was emphasized that the statistical community should remain acutely aware of its mission, namely to provide high quality and impartial information on relevant topics to the society at large and to policy makers in particular. Innovation is crucial and the statistical community should embrace it, but the traditional core business will remain and is still needed. For instance, results from Big Data applications will need to be benchmarked, and traditional surveys will be needed to provide those benchmarks.

14. The statistical community should identify and build business cases for Big Data. Efforts should be made to link the use of Big Data to the Post-2015 development agenda and its sustainable development goals, as these will pose huge demands on the statistical systems for years to come. The many proposed targets of the Post-2015 development agenda can certainly not all be measured by indicators based on Big Data. However, Big Data is expected to play an important role given their timeliness and geo-spatial detail. One way forward is to explore proxy indicators based on Big Data that could provide more frequent information than surveys. In this manner, Big Data is complementary and not replacing traditional systems. This is similar to the “flash” estimates of GDP, for example.

15. In summary, the conference provided an overview of promising Big Data applications for official statistics and contributed to a better understanding of the methodological, privacy and access challenges. The conference fostered sharing of experiences between national statistical offices and

other important stakeholders and partners on issues of building partnerships, access to data, communication and advocacy, and on the need for training, updating of skills and capacity building.

III. Meeting of the Global Working Group on Big Data for Official Statistics, 31 October 2014, Beijing, China

16. The Global Working Group (GWG) reviewed at its first meeting its terms of reference and discussed the work programme for 2015. The GWG consists currently of 28 members from developed and developing countries and from various international and regional organizations. The list of members is shown in Annex II.

17. In the margins of the conference, UNSD, UNECE, UNESCAP, SIAP, ITU and Eurostat met to ensure coordination of the work on Big Data among international organisations. As a result, all agencies committed to share their information and outputs in this area, and UNECE will create and maintain a calendar of events. One way to stimulate the cooperation will be to closely link the various organizational websites on Big Data.

A. Terms of reference and mandate of the GWG on Big Data for official statistics

18. The statistical community has in general the obligation of exploring the use of new data sources to meet the expectation of the society for enhanced products and for improved and more

efficient ways of working. However, the terms of reference and mandate of the GWG on Big Data for official statistics should certainly also be understood within the context of fulfilling the new data demands posed by the monitoring and reporting requirements of the Post-2015 development agenda.

19. The use of Big Data for official statistics is fully supported by the Fundamental Principles of Official Statistics. By way of example, Principles 1, 5 and 6 state (a) that official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information; (b) that data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records; (c) that statistical agencies are to choose the source with regard to quality, timeliness, costs and burden to respondents; and (d) that individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes. Based on these principles, it is not only a possibility but almost an obligation for statistical agencies to investigate and pursue using Big Data sources for statistical purposes, as long as privacy and confidentiality are strictly observed. Big Data have the potential to be relevant, timely and be more cost effective than traditional data collection methods.

20. The terms of reference of the GWG also refer to the report of the Independent Expert Advisory Group (IAEG)[2] on the Data revolution for sustainable development, which emphasized that statistical offices will need to change, and more quickly than in the past, and continue to adapt, abandoning expensive and cumbersome production processes, incorporating new data sources, including

[2] The Secretary-General appointed this Independent Export Advisory Group (IAEG) on 29 August 2014 to draft a report on the Data Revolution within the context of the post-2015 development agenda. The IAEG released the report "A World that counts: mobilising the Data Revolution for sustainable development" on 6 November 2014.

administrative data from other government departments, and focusing on providing data that is human and machine-readable, compatible with geospatial information systems and available quickly enough to ensure that the data cycle matches the decision cycle[3]. The IAEG report also explicitly refers to Big Data proposing that national capacity for data science must be developed to leverage opportunities in Big Data and complement official statistics. Increased domestic resources and international support for developing countries are needed to have the data revolution contribute to sustainable development. Through bilateral and international coordination and collaboration, common and standardized applications of Big Data for official statistics must be developed and scaled up transparently, demonstrating full compliance with applicable laws[4].

21. Among the many proposals in the IAEG report is also the proposal to create a global "Network of Data Innovation Networks", to bring together the organisations and experts in the field. This would: contribute to the adoption of best practices for improving the monitoring of the new sustainable development goals (SDGs), identify areas where common data-related infrastructures could address capacity problems and improve efficiency, encourage collaborations, identify critical research gaps and create incentives to innovate.

22. In summary, statistical agencies should choose data sources with regard to quality, timeliness, costs and response burden, and Big Data sources fall within this scope. Big Data have the potential to be relevant, timely and be more cost effective than traditional data collection methods, and could make the data cycle match the decision cycle. The work on Big Data should contribute to the adoption

[3] IAEG, "A World that counts: mobilising the Data Revolution for sustainable development", page 9
[4] IAEG, "A World that counts: mobilising the Data Revolution for sustainable development", page 23

of best practices for improving the monitoring of the new SDGs under the Post-2015 development agenda. Some of the new indicators or proxies of those indicators could be based on Big Data sources with improved timeliness and granular social and geo-spatial breakdown.

23. Against this background the GWG formulated its terms of reference and mandate as follows (1) to provide strategic vision, direction and coordination of a global programme on Big Data for official statistics, including for indicators of the post-2015 development agenda; (2) to promote practical use of Big Data sources, including cross-border data, while building on the existing precedents and finding solutions for the many challenges; (3) to promote capacity building, training and sharing of experience; (4) to foster communication and advocacy of use of Big Data for policy applications, especially for monitoring of the post-2015 development agenda; and (5) to build public trust in the use of private sector Big Data for official statistics.

24. The full terms of reference including the mandate are given in Annex I.

B. Programme of work and deliverables

25. Given the discussions during the international conference and the GWG meeting itself, the GWG agreed that its programme of work would consist of a number of work streams carried out by specific task teams, and have one team in addition for the overall coordination. Firstly, it was agreed to continue the work on the three classes of Big Data sources of the conference, namely on Mobile

phone data, Satellite imagery and Social media data. It was understood that advancing the work in each of these three task teams will include engagement with at least one pilot project.

26. Secondly, one of the recurring themes of the conference was access to data and building partnerships with private sector and other communities. Therefore, a task team was created to explore these linked topics in more depth. This team will also investigate the possibility of establishing umbrella agreements for access to data with globally operating Big Data providers. Access to and use of Big Data are also linked to public trust and to communicating the benefits and challenges of Big Data in general. For this purpose a task team on Advocacy and Communication was created, which will also look into fund-raising strategies to allow developing countries active participation in pilot projects.

27. Thirdly, given the context of the post-2015 development agenda it was further agreed to have one team especially tasked with keeping track of the links between the indicators needed for monitoring the Sustainable Development Goals (SDGs) and the Big Data applications. Finally, two more teams were proposed and agreed, namely one on training, skills and capacity building given that different skills are necessary for processing Big Data than the skills currently available in statistical offices; and one on cross-cutting issues such as classification and frameworks. This last team will work among others on the classification of Big Data sources building on the classification so far proposed by the UNECE Big Data group. The overall coordination of the work of these 8 task teams would be done by a coordination team consisting of the team leaders.

28. It was pointed out that these GWG task teams should build on the experience of the UNECE Big Data project and the material that is being produced by the UNECE Big Data task teams. Those UNECE task teams will soon complete their work, while the GWG task teams are only starting up. This provides the possibility of continuation of some of the work flows that have already started. Each GWG task teams will need to develop its own work programme, time schedule and outputs in close consultation with the coordination team.

29. The deliverables of the GWG are directly related to the work of the 8 task teams. Therefore, it is expected that the GWG will deliver a number of proposals and reports as shown below, and will participate in some pilot projects as well. The GWG will also continue working on an inventory of Big Data projects, probably through its coordination team. Results of the first survey are shown in the next section. So, the concrete deliverables for the year 2015 are:

- Proposals and activities for training, skills and capacity building to exploit Big Data sources for official statistics;
- A report on links between Big Data and the sustainable development goals;
- A strategy on advocacy and communication of the use of Big Data for official statistics;
- A report of the use of mobile phone data for official statistics, which clarifies and provides solutions to methodological, IT and privacy challenges, and includes possible uses of these data for monitoring the Post-2015 development agenda;
- A report of the use of social media data for official statistics, which clarifies and provides solutions to methodological, IT and privacy challenges, and includes possible uses of these data for monitoring the Post-2015 development agenda;

- A report of the use of satellite imagery and remote sensing data for official statistics, which clarifies and provides solutions to methodological, IT and privacy challenges, and includes possible uses of these data for monitoring the Post-2015 development agenda.
- Proposals for enhancing access to Big Data sources and partnerships with private sector;
- A report on cross-cutting issues, classification, frameworks and taxonomy.

30. In addition to the concrete deliverables of the different task teams, the GWG will participate in a number of pilot projects on the implementation of Big Data for official statistics and will also continue working on an inventory of Big Data projects, building on the UNECE Big Data project inventory and on the results of the initial survey conducted by UNSD and UNECE, as shown in the next section.

IV. Results from the survey on Big Data strategies and project inventory by UNSD and UNECE

31. In September 2014, UNSD and UNECE teamed up to conduct a joint survey on Big Data projects in official statistics. The goal was to provide an overview of active Big Data projects in order to facilitate a more informed discussion within the community at large and to further develop the programmes of the global working group. The survey was sent to all participating offices in the UNECE Big Data Project and to the members of the GWG on Big Data.

32. The survey had two objectives: collecting information about completed, ongoing or potential Big Data projects including information about partnerships, data sources and tools, and collecting

information about the organizational setting of those projects. Two distinct questionnaires were distributed. The first aimed to collect information on overall Big Data organization, strategies and governance structures, while the second focused on actual Big Data projects. For this survey a fairly wide definition of Big Data was adopted, namely that *"Big data are data sources with a high volume, velocity and variety of data, which require new tools and methods to capture, curate, manage, and process them in an efficient way"*.

33. In total, the survey was sent to 78 national statistical offices and 28 international organizations.[5] For the questionnaire on the organizational context, 32 countries responded, while 24 countries responded on the individual project questionnaire, submitting information about a total of 54 projects. In addition, international organizations submitted information about 3 projects. It is fair to assume that these countries and international organizations probably cover a substantial number of completed and ongoing Big Data projects world-wide. However, the GWG considers conducting a global follow-up survey to collect more information about more projects for the Big Data project inventory. The full analytical report of the findings of this survey is provided as a background document. Below is a summary of some of the main findings.

A. Main findings – Organizational context

34. The questionnaire on the organizational context focused on the general experience and frameworks of handling Big Data, in terms of Big Data strategy, governance and management

[5] International organizations were emailed the format in large part to inform them that the survey was taking place.

structures, assessment of quality, privacy and confidentiality issues and skill shortage. The responses collected provide support to the conclusions at the international conference on Big Data for Official Statistics, regarding the major challenges facing NSOs.

35. The survey revealed that while only a few countries have developed a long-term vision for the use of Big Data, a number of countries are currently on the brink of formulating a Big Data strategy. Furthermore, a number of countries have established internal labs, task teams or working groups to carry out pilot projects to determine if and how Big Data could be used as a source of Official Statistics, but most countries have not yet defined business processes for integrating Big Data sources and results into their work and do not have a defined structure for managing Big Data projects.

36. In many countries steps are taken to build partnerships for the purpose of exploring Big Data opportunities as part of the formulation of a Big Data strategy. In order to minimize the risk of failure in exploring these new technologies, countries have decided to actively participate in international collaborations such as the UNECE Sandbox project, Eurostat's Big Data Task force and the Global Working Group on Big Data established under the Statistical Commission.

37. Only a few countries have a specific privacy framework for Big Data, while most survey respondents mention that the privacy framework for traditional statistics is currently applied to Big Data as well. Many organizations stress the importance of protecting privacy and confidentiality when dealing with Big Data, even beyond what is strictly required by law, given the importance of their public image. Some organizations avoid the issue of privacy and confidentiality by doing all the testing and manipulation of the data at the location of the Big Data provider, and then only transmit

aggregates to the NSO. Some organizations argue that while strict data protection regulations are needed, these often also impose a barrier to getting access to the data.

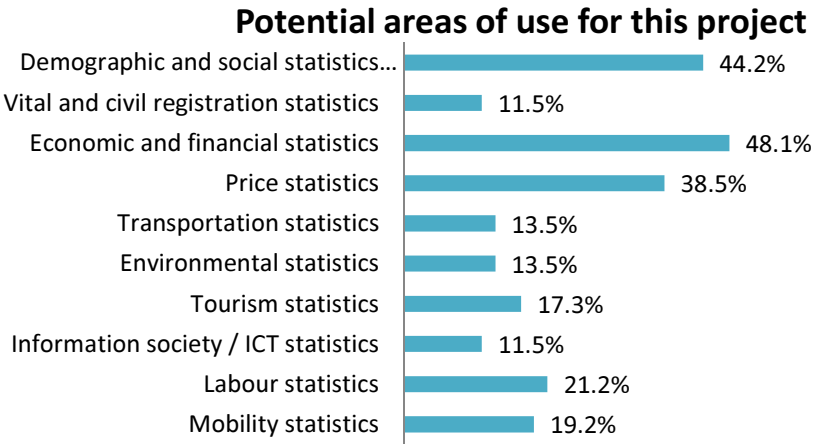
38. While most respondents recognize the challenges related to IT, skills, legislation and methodology, the biggest challenge for most Big Data projects is the limited or restricted access to potential datasets. Big data are, to a large extent, owned by the private sector (e.g. on-line companies, mobile phone operators, banks, etc.), thus, it is very important to build close collaboration with the private sector. Many of these players are global companies; hence the global statistical community could use their collective bargaining power to obtain the access to these data sets.

B. Main findings – Individual projects

39. The questionnaire on the individual projects had more detailed questions on the potential areas of use, status of project, outcomes and lessons learned, as well as detailed questions on the partnership arrangements, data sources and data analysis, tools and skills used. Many of the projects collected by the survey were ongoing projects for purposes of testing and experimentation of the use of Big Data, while some were completed. The other projects are in more or less advanced stage of exploration.

40. In terms of the potential areas of use, most of the projects are related to economic and financial statistics, demographic and social statistics, and price statistics, see graph below. It is worth noting that many projects often apply to several statistical areas. These are projects that exploit one data sources for different statistical areas. For example, the Mexican National Institute of Statistics and

Geography (INEGI) collected a sample of messages posted on Twitter, and used the data for different applications, namely for subjective well-being indicators, and tourism and border mobility statistics.



41. Just over half of the projects are based on partnerships. The most common partner is a commercial enterprise, followed by other government agencies. The most common type of partnership is with a data provider for access to data. This includes partnerships with mobile phone companies, smart meter data providers or social media data aggregators. For example, to conduct a project for the production of daily mobility patterns, the Italian National Statistical Institute (ISTAT) partnered with a mobile phone data provider. The Swiss Federal Statistical Office engaged in a partnership with the four major retail stores in Switzerland to receive scanner data twice a month to produce consumer price indices for food and near-food commodities.

42. Besides partnerships with commercial data providers, for many projects the NSOs have also established partnership with analytical partners. This could be to acquire specific know-how of the Big Data source, transform raw data into the intended data structure, or apply analytical techniques to produce certain outputs. However, this kind of partnership could also be to provide human resources

to perform labour intensive tasks. For example, INEGI partnered with a university in Mexico to get assistance from a large number of students to classify content of tweets for subjective well-being analysis.

43. Overwhelmingly survey participants chose internal hosting solutions rather than purchasing external hosting services. Privacy concerns were cited as the primary reason for internal hosting. External hosts were used more often when the data was already deemed to be in the public domain. Internal hosting could become a limiting factor in the choice of a Big Data source. Investments on internal hosting would need to cover hardware, software and human skills and could represent a significant hurdle to begin a Big Data project, especially given the fact that technology is rapidly changing.

V. Conclusions and points for discussion

44. The statistical community has started to explore in earnest the opportunities offered by Big Data sources for application in official statistics. The international conference on Big Data for Official Statistics in Beijing showcased applications of several Big Data sources for a number of different types of official statistics, and fostered experience sharing between developed and developing countries. Many of the challenges facing national statistical offices are similar, such as access to data, privacy concerns, methodological issues, expanded computing resources requirements, and lack of in-house skills for Big Data analytics, and underlined the need for global collaboration. The GWG has begun its work by formulating its operating modalities and focal areas. It has also initiated the process

of preparing an inventory of ongoing activities regarding the use of Big Data for official statistics. The survey on Big Data projects collected a number of interesting examples that may offer important lessons learned for other organizations. The GWG in collaboration with the regional working groups needs to develop further guidance and conduct a number of pilot projects with the objective to consolidate and accelerate the progress in the use of Big Data for official statistics.

45. **The Commission is invited to express its views on:**

(i) the terms of reference and mandate of the GWG on Big data for official statistics;

(ii) the proposed task teams of the GWG and corresponding deliverables on training, skills and capacity building; linking Big Data and SDGs; advocacy and communication; mobile phone data; social media data; satellite imagery; access and partnership; and cross-cutting issues, such as classifications and frameworks;

(iii) the results of the UNSD/UNECE survey and the plans to extend this to a global survey;

(iv) the pilot projects, which GWG intends to conduct in collaboration with the regional working groups to consolidate and accelerate the progress in the use of Big Data for official statistics.

Annex I: Terms of Reference and mandate of the global working group on Big Data for official statistics

The statistical community has the obligation of exploring the use of new data sources to meet the expectation of the society for enhanced products and for improved and more efficient ways of working. The terms of reference and mandate of the GWG on Big Data for official statistics are based on this principle, but should certainly also be understood within the context of fulfilling the new data demands posed by the monitoring and reporting requirements of the Post-2015 development agenda.

The main deficiency of the indicators for monitoring of the Millennium Development Goals was lack of timeliness (or availability). To improve on that situation for the goals of the Post-2015 agenda, the High-level Panel of Eminent Persons called in its report of May 2013 for a Data Revolution, which would draw on existing and new sources of data to fully integrate statistics into decision-making, promote open access to, and use of, data and ensure increased support for statistical systems.

These terms of reference also refer to the report of the Independent Expert Advisory Group (IAEG) on the Data revolution for sustainable development, which emphasized that statistical offices will need to change, and continue to adapt, abandoning expensive and cumbersome production processes, incorporating new data sources, including administrative data from other government departments, and focusing on providing data that is human and machine-readable, compatible with geospatial information systems and available quickly enough to ensure that the data cycle matches the decision cycle[6].

Within this context Big Data sources are recognized as constituting an important part of the Data Revolution to support the monitoring of the Post-2015 development goals. Big Data could contribute

[6] IAEG, “A World that counts: mobilising the Data Revolution for sustainable development“, page 9

to improving some aspects of the quality of statistics, such as timeliness and relevance, without compromising the impartiality and methodological soundness of the statistics.

Also the Fundamental Principles of Official Statistics[7] encourage the use of new data sources such as Big Data, as they state that:

- Official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information (from Principle 1)
- Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents. (Principle 5)
- Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes. (Principle 6)

With these strategic considerations of the Fundamental Principles, the Post-2015 Agenda and the Data Revolution as a basis, and with reference to Decision 45/110, the mandate of the working group is then formulated as follows:

a) Provide strategic vision, direction and coordination of a global programme on Big Data for official statistics, including for indicators of the post-2015 development agenda;

[7] See A/RES/68/261

b) Promote practical use of of Big Data sources, including cross-border data, while building on the existing precedents and finding solutions for the many challenges:

- Methodological issues, covering quality concerns and data analytics;
- Legal and other issues of access to data sources;
- Privacy issues, in particular those relevant to the use and reuse of data, data linking and re-identification;
- Security, IT issues and management of data, including advanced ways of data dissemination, assessment of cloud computing and storage, and cost-benefit analysis;

c) Promote capacity building, training and sharing of experience;

d) Foster communication and advocacy of use of Big Data for policy applications, especially for monitoring of the post-2015 development agenda;

e) Build public trust in the use of Big Data for official statistics.

Annex II: Members of the global working group on Big Data for Official Statistics

Countries
Australia
Bangladesh
Cameroon
China
Colombia
Denmark
Egypt
Indonesia
Italy
Mexico
Morocco
Netherlands
Oman
Pakistan
Philippines
Tanzania
United Arab Emirates
United States

Agencies
Eurostat
GCC-Stat
ITU
OECD
UN Global Pulse
UNECE
UNESCAP / SIAP
UNSD
UPU
World Bank