Economic and Social Council

Statistical Commission
Fifty-first session
3–6 March 2020
Item 3 (t) of the provisional agenda*
Items for discussion and decision: big data

Report of the Global Working Group on Big Data for Official Statistics

Note by the Secretary-General

In accordance with Economic and Social Council decision 2019/210 and past practices, the Secretary-General has the honour to transmit the report of the Global Working Group on Big Data for Official Statistics. Following Statistical Commission decision 49/107, the report presents the governance structure, the business model and the establishment of the regional hubs of the United Nations Global Platform. It also contains an overview of the range of projects on data solutions and training activities undertaken and the partnerships established for the platform that reflect its future sustainability. With the establishment of the regional hubs in 2020 and the extensive network of international experts, the necessary building blocks are in place to scale-up the number of data projects and to launch an extensive training programme for developing new skills and capabilities in the use of big data and modern technology. The report also includes information on the release of new handbooks, training workshops and collaborative projects on new data solutions. The Commission is invited to comment on the points for discussion.

---

* E/CN.3/2020/1.

Please recycle

I. Introduction

1.   In its decision 49/107, the Statistical Commission reaffirmed that the use of big data and other new data sources is essential for the modernization of national statistical institutions and requested the Global Working Group on Big Data for Official Statistics (a) to further develop the Global Platform as a collaborative research and development environment for trusted data, methods and learning; and (b) to make the business case for sustainability of the Platform by 2020.

2.   The Global Platform has evolved from a concept into a reality as a collaborative environment for the delivery of data, methods and learning. The Global Working Group is now working out the details for control and access to the various data sets on the Platform, the degree to which data and algorithms should be made open and how the software, services and tools available will be vendor independent. Those questions have a direct bearing on the business model for the Platform, which is outlined in the present report.

3.   Section II of the report includes information on the Global Working Group's Bureau, its Board and its task teams. The functions, organization and network of the Global Platform are provided in section III, and section IV contains information on the events relating to the Global Working Group over the past year. The next steps to be taken by the Global Working Group to advance its work programme are described in section V, and section VI contains a conclusion and points for discussion.

II. Progress of the Global Working Group on Big Data for Official Statistics

A. Bureau of the Global Working Group

4.   The Bureau of the Global Working Group consists of two co-Chairs (Denmark and the United Kingdom of Great Britain and Northern Ireland) and 10 members (Canada, Colombia, Indonesia, Mexico, Netherlands, United Arab Emirates, the African Development Bank, Eurostat, the Economic Commission for Europe (ECE) and the Statistics Division). The Bureau has held teleconferences every two weeks since March 2018, and all of the meetings have been documented on the Global Working Group's website (https://unstats.un.org/bigdata/bureau).

5.   The Bureau has been instrumental in formalizing the organization and management of the operations of the Global Platform and in obtaining commitments for the establishment of the regional hubs and the wider Global Platform network. Moreover, the Bureau has fully promoted the use of the Platform and supported the work of the task teams and the organization of the major regional conferences.

B. Board of the Global Working Group

6.   The Board of the Global Working Group consists of the full membership of the Global Working Group and meets once a year. The most recent annual meeting of the Board[1] was held in the margins of the fifth International Conference on Big Data for Official Statistics, held in Kigali (see paras. 45–47 below). The meeting was chaired by Denmark and attended by 20 countries and five international organizations. Agenda items included the progress reports of the task teams and of the Global Platform, the adoption of the Kigali Declaration (see annex II below) and an initial

_____

[1] Meeting report available at https://unstats.un.org/bigdata/bureau/, under the "Documents" tab.

discussion on the report of the Board to the Commission. In view of the progress made on the Global Platform, the Board agreed on 1 May 2019 to a set of recommendations, which are contained in the Kigali Declaration.

7. The three main recommendations contained in the Kigali Declaration have been implemented as follows:

(a) The Advisory Board for the Global Platform was created in September 2019 and will address the overarching policies and principles for the set-up and use of the platform;

(b) An independent, not-for-profit entity was created in August 2019 tasked with setting up an agreement with the United Nations for the management and operation of services for the Global Platform;

(c) Regional hubs are being created in Brazil, China, Rwanda and the United Arab Emirates to build capacity and deliver regional support for data innovation in official statistics;

(d) Memorandums of understanding between the United Nations and the respective regional hubs are scheduled to be signed at the time of the 2020 session of the Statistical Commission.

8. In the Kigali Declaration, it was reiterated in particular that the Global Platform should provide technological infrastructure as a service for all States Members of the United Nations, and emphasis was made on the need to support the least developed countries and small island developing States to facilitate their access to global data sets and state-of-the-art tools and services. The platform should also be an environment for capacity development activities in new areas like data science, artificial intelligence, machine learning and privacy-preserving techniques.

9. In the Declaration, the Government of Rwanda offered to host a hub of the Global Platform at the National Institute of Statistics of Rwanda. That regional hub would support capacity development in the area of big data for official statistics for Africa, in collaboration with international and regional agencies. The Economic Commission for Africa (ECA) has fully endorsed the regional hub for the Africa region and has become a full partner in the Rwanda regional hub.

C. Task teams of the Global Working Group

10. The Global Working Group delivers its work through several task teams, which develop methods, prepare handbooks, conduct capacity-building activities and collaborate actively on the Global Platform. The Global Working Group has active task teams on the use of satellite imagery data, mobile telephone data, scanner data and automatic identification system (vessel tracking) data, and a task team on training, skills and capacity-building. The progress that each of the task teams makes is presented in the present section. A more detailed overview of their work and their training workshops is posted on the Global Working Group's website.[2]

1. Task team on satellite imagery data

11. The task team on satellite imagery data is led by Canada and has delivered a handbook with information on sources of earth observation data, as well as methodologies for producing crop and other statistics. The team has also conducted workshops (in Colombia in 2017 and in Thailand in 2018) to provide a better understanding of the strengths and limitations of the use of satellite data, as well as

_____

[2] See, for example, https://unstats.un.org/bigdata/taskteams/mobilephone/.

the management and manipulation of such data and its application to estimating crop statistics.

12. New deliverables include (a) the use of analysis-ready data and production-ready data for the estimation of crop types, acreage and yields, land-use change, biodiversity and greenhouse gas emissions; (b) the development of training (such as e-learning) material on the use of algorithms and methods to estimate crop types and yields; and (c) the preparation of accompanying methodological guidance material, including a knowledge library of existing practices in the area. The new estimates support measurements for Sustainable Development Goals indicator 2.4.1.

2. Task team on mobile telephone data

13. The task team on mobile telephone data is led by the International Telecommunication Union (ITU) and has delivered a handbook on the use of mobile telephone data for official statistics, describing applications, data sources and methods. The team deals with access to mobile telephone data through partnership models with mobile network operators. The team has conducted several workshops, including in Colombia in 2017 and in Georgia, Rwanda and Indonesia in 2019.

14. New deliverables will include the delivery of a second handbook, on methodologies for using mobile telephone data in measuring statistics and Sustainable Development Goals indicators; the preparation of a synthetic training data set; participation in and documentation of data projects; and the performance of training workshops and outreach events. The task team has several members [3] from the statistics community, as well as from the private sector, civil society and academia.

3. Task team on scanner data

15. The task team on scanner data is led by Canada and has delivered open-source methods and an enduring data set, which can be shared among all partners in the statistics community. The methods take cleaned and classified scanner data and enable the estimation of price indexes. The exact method can be specified by the user. The task team has also delivered training and instructional material on the use of the application, as well as additional methodological guidance material that (a) summarizes the relevant literature on methods; (b) identifies internationally agreed recommendations on the methods that apply in different situations; and (c) catalogues existing practices across national statistical offices in this domain.

16. The lessons learned included that the Global Platform provided an easy-to-use interface and that the availability of a range of trusted index methods had expedited learning, once task team participants became familiar with how to use the Platform. In phase two of the work of the task team, scanner data will also be used to calculate consumer price index expenditure weights, and more analysis-ready data sets will be provided, plus more training materials. Furthermore, the membership of the task team will be expanded to include the United States Federal Reserve, Eurostat, the data analytics company Nielsen and the European Central Bank.

4. Task team on automatic identification system vessel tracking data

17. The automatic identification system task team was created in 2019 and is led by the Statistics Division. The team aims (a) to develop algorithms and methodologies for measuring freight transportation, traffic within harbours, economic trade indicators, carbon dioxide emissions, the amount of fishery and other experimental indicators that are fit-for-purpose; and (b) to conduct training on the use of automatic identification system data in those domains. The team will use the Global Platform

_____

[3] See https://unstats.un.org/bigdata/taskteams/mobilephone/.

for global collaboration on projects involving automatic identification system data. In that regard, the first training session on automatic identification system data was conducted in September 2019.

18. The deliverables include a handbook on the use of automatic identification system data for statistics and indicators, training materials and the delivery of training activities, and the development of links between automatic identification system data and related data sources, such as trade data, shipping registers and remote sensing data.

5. Task team on privacy-preserving techniques

19. The task team on privacy-preserving techniques is led by the United Kingdom and develops or proposes principles, policies and open standards for encryption within the Global Platform to cover the ethical use of data and the methods and procedures for the collection, processing, storage and presentation of data, taking full account of data privacy, confidentiality and security issues. Using those open standards, algorithms, policies and principles will reduce the risks associated with handling proprietary and sensitive information. Privacy-preserving computation comes at a cost: current versions of those technologies are costly, rely on specialized computer hardware or are difficult to programme and configure directly. Privacy-preserving techniques are being developed quickly, and the task team will help statistical offices and other platform users assess their utility and benefits.

20. The task team has developed a handbook[4] aimed at allowing national statistics offices to gain access to new sources of (sensitive) big data in a secure and appropriate way to enable collaboration on sensitive big data across multiple offices. The handbook provides a technical overview for each privacy-preserving technique, gives examples of applied uses and explains the modelling challenges and security arguments that typically apply. It further gives an overview of the maturity of privacy-preserving technologies, utilizing Wardley mapping, as well as the availability and costs of using the available technologies. In its next steps, the task team will work to develop practical cases and a second handbook, which will look into the legal aspects of using sensitive big data sources.

6. Task team on training, competencies and capacity development

21. The task team on training, competencies and capacity development is led by Poland. The general objectives of the task team are to develop a baseline needs assessment for the big data competencies required within national statistics offices and an assessment of their institutional readiness to use big data;[5] to build a competency framework for big data acquisition and processing for those offices; and to identify and analyse the existing supply of training on big data available to data scientists in those offices, academia and other centres.

22. The task team will deliver reports on the skills assessment survey and data science governance across national statistics offices, as well as on the competency framework and the training market for data scientists. It will also be involved in the development of training courses in the use of big data for official statistics.

7. Other task teams

23. The Global Working Group has additional task teams that are currently being formed and are in the process of defining their objectives and deliverables. Those task teams are on subjects that include the use of social media data for statistics and

---

[4] See https://marketplace.officialstatistics.org/privacy-preserving-techniques-handbook.
[5] A report on the global assessment of institutional readiness to use big data is provided as a background document.

indicators, integrating statistical and geospatial information, the use of administrative data and linking big data to the Sustainable Development Goals. The latter task team will systematically identify those Sustainable Development Goals indicators that are or can be computed with the support of big data and will explicitly address the question of how big data can help to "leave no one behind".

24. The task team on integrating statistical and geospatial information will build on the ideas emerging from a data integration workshop held in Belgrade in May 2019.[6] The main outcome of that workshop was a request for a "sandbox" environment where statistical and geospatial data could be shared, along with methods, tools and algorithms for integrating them. The Global Platform could provide such environment for joint projects, experimentation and testing of new ideas. The task team will create links with relevant international projects[7] and will bring together experts from the statistical and geospatial communities to share ideas and experiences in a mutually beneficial way.

III. The United Nations Global Platform

25. A marketplace capability has been developed for the Global Platform to enable international collaboration on and sharing of trusted data, trusted methods and trusted learning between trusted partners. The platform currently has over 400 users, including the experts of the Global Working Group task teams, who are working to develop new statistical methods from big data sources such as automatic identification system shipping data or remote sensing data. The platform can provide all 193 Member States of the Statistical Commission with friction-free access to those global data sources using cloud services, and enables virtual international collaboration in the use of the data to develop new methods for economic and social indicators, including many of the Sustainable Development Goals indicators. The sections below provide a more detailed explanation of the Global Platform as a digital platform, an organization and a network.

A. Global Platform: a digital platform

26. The Global Platform has developed greatly in the friction-free delivery of services. It currently contains several alpha services, such as access to cloud services including Alibaba Cloud, Amazon Web Services, Google Cloud Platform and Microsoft's Azure. It offers a range of services for code collaboration, methods publishing, Earth observation and location data analysis. Users of the platform can search, build, deploy and consume algorithms and statistical methods and can further develop methods in the main programming languages used by the statistics community, such as R, Python, Java and Scala. The platform can also host machine learning models and publish application programming interface endpoints to these.

27. The platform marketplace[8] gives users access to data sets, such as (a) global automatic identification system data feeds containing the most complete situational picture of global vessel activity by drawing data from satellite and terrestrial automatic identification system receivers; (b) global automatic dependent surveillance-broadcast data feeds with the world's largest source of unfiltered flight data; (c) Sentinel-2 satellite imagery, which provides a global coverage of the Earth's land surface every

---

[6] Organized by the Economic Commission for Europe, the United Nations Initiative on Global Geospatial Information Management (Europe), Eurostat and the European Free Trade Association.

[7] Including the proposed new Geostat 4 project in the European Union.

[8] See https://marketplace.officialstatistics.org/.

five days; and (d) OpenStreetMap, a free, editable map of the world, created and maintained by volunteers with regular data archives made available by Amazon.

28. The platform also makes statistical methods available as trusted algorithms and supports machine learning for the statistical analysis of raw research data. For example, the "XGBoost_SDG_classifier" algorithm automatically identifies Sustainable Development Goals referred to in a text, whereas the "HighwayScrapeR" algorithm samples a queried area for OpenStreetMap highways and returns points at user-defined intervals (in metres) along the matching highways. The learning area[9] of the marketplace already contains about a dozen handbooks on topics related to big data, which can be used in training settings. These include the handbook on the use of mobile telephone data for official statistics or the handbook on privacy-preserving techniques.

B. Global Platform: an organization and business model

29. The digital Global Platform needs to be operated, maintained, supported and further developed. To date, a team at the United Kingdom Office for National Statistics has done most of this work, assisted by staff from the Statistics Division,[10] Statistics Netherlands, Statistics Canada and others. As of the second quarter of 2020, those services are scheduled to be provided by a non-governmental, not-for-profit institute called the Global Platform Institute and by four government-related regional hubs. Together, those five hubs have committed to signing individual agreements with the United Nations for the purposes of operating, maintaining and further developing the Global Platform.

1. Governance structure

30. The present section contains a description of the overall governance structure for the Global Platform, as well as descriptions of the functions of the regional hubs and the Global Platform Institute.

31. The Platform has been created and is operated to support the national statistical systems of all Member States under the auspices and guidance of the Statistical Commission. The Commission will oversee the decisions taken in the operation of the Global Platform, including policies to ensure the ethical use of data, policies on access to data on the platform and policies on the scope of work conducted on the platform.

32. The Global Working Group was established by the Commission to give direction to the global statistics community on the use of big data for official statistics. The actual work of the Group is delegated to the task teams and is managed by the Bureau (see section II above). The Board of the Global Working Group oversees the work of the Bureau and the task teams and reports to the Commission.

33. In September 2019, the Bureau of the Global Working Group established an Advisory Board, which provided guidance on the preparation of the memorandums of understanding between the United Nations and each of the five hubs that together manage and operate the Global Platform. The elements covered in those memorandums include a preamble, the purpose, objectives and activities of the hub, its organization and management, the support and inputs by the parties and some standard clauses. The Advisory Board is chaired by the United Kingdom Office for National Statistics and includes the senior managers of the institutes of the hubs (see para. 36 below), the co-Chairs of the Global Working Group and the Chair of the Technical Delivery

_____

[9] See https://marketplace.officialstatistics.org/learnings.

[10] As an example of the contributions made by the participating organizations, the value of the resources devoted so far by the Statistics Division to the development of the Global Platform is estimated at more than $3 million.

Board. The Statistics Division participates ex officio. The signing of the memorandums was planned for the March 2020 meeting of the Statistical Commission, to be held in New York. At that point, the Global Platform will be officially launched.

34. The Global Working Group further created a Technical Delivery Board to provide independent advice on the technology strategy for the Global Platform. The Board advises both the Bureau of the Global Working Group and the Advisory Board, and reports to the Board of the Global Working Group at its annual meeting. It has also developed a handbook[11] on information technology strategy that covers such topics as cloud computing, serverless technology, and security and data science.

35. Finally, the Global Working Group intends to create a scientific committee that will be composed of subject matter specialists and will oversee the quality assurance of all products delivered by the Global Working Group, including data, methods, algorithms, training materials and training activities. Similar to the Technical Delivery Board, the scientific committee will advise the Bureau and the Advisory Board and will report to the Board of the Global Working Group at its annual meeting.

2. Global Platform regional hubs

36. Regional hubs of the Global Platform are being created in Hangzhou, China, with the National Bureau of Statistics of China and the government of Zhejiang Province; in Kigali with the National Institute of Statistics of Rwanda and the African Centre of Statistics of ECA; in Dubai with the Federal Competitiveness and Statistics Authority of the United Arab Emirates; and in Rio de Janeiro, Brazil, with the National Statistical Office of Brazil.

37. The main objectives of the regional hubs are to facilitate projects on the use of big data and data science for official statistics and Sustainable Development Goals indicators, to share knowledge on newly developed methods, algorithms and tools, and to provide training on the use of big data and data science for the global official statistics community. Hub activities will include research projects on the use of big data and data science for the development of new statistical methods and techniques, the delivery of capacity-building and training to statisticians and data scientists from national, regional and global statistical offices, and the further development and maintenance of the Global Platform in collaboration with the other hubs. Many data sources will be openly shared across the Platform. Mechanisms are also being developed to enable selective data-sharing between the different hubs across the federated system, for example, where sensitive data is involved or where there are specific licensing restrictions. The governance of that data-sharing will be overseen by the United Nations.

38. Each hub will have a steering committee, consisting of senior staff of the United Nations and the hub, that will advise on the formulation of the annual activities in a programme of work for each hub. That information will subsequently be reported to the Global Working Group. For each hub, the steering committee will appoint a representative from the hub and a representative from the United Nations jointly to prepare a programme of activities and to oversee its implementation within the hub. Each hub will assume responsibility for making staff and other necessary resources available for the provision of agreed services and for making facilities available for the hosting of training activities and possible hosting of project staff. The United Nations and all Global Working Group members are responsible for mobilizing international support and experts to work with the hubs.

---

[11] See https://marketplace.officialstatistics.org/un-global-platform-handbook-on-information-technology-strategy.

3. Global Platform Institute

39.   The Global Platform Institute has been established as a not-for-profit organization in the United Kingdom. Like the regional hubs, the Institute's main objective is to facilitate projects on the use of big data and data science on the Platform, including the sharing of data, methods, algorithms and tools and the provision of training for the global official statistics community. In consultation with the regional hubs, the Institute will assume responsibility for the management and coordination of the services across the platform, which include facilitating access to global data sources, the onboarding of users, user authentication, data security and global billing services. The Institute will assure the interoperability of services across different cloud providers.

40.   In coordination with the regional hubs, the Statistics Division and other Global Working Group members, the Institute will set up a global support structure for the Global Platform, to ensure that collaborative projects run smoothly, with transparent accounting of the services consumed and the data used on each project. The operating costs for the provision of services will be recuperated through the global billing system, which should indicate for each project or training activity which services were consumed and at what cost.

C.   Global Platform network

41.   The five hubs of the Global Platform will operate collectively and maintain the Platform on the basis of the memorandums of understanding with the United Nations and under the overall guidance of the Statistical Commission.

42.   Furthermore, the Global Platform has obtained support from a wide network of organizations and is therefore referred to as the United Nations Global Platform network. First, the member organizations of the Global Working Group actively contribute to the activities and operations of the Global Platform network. That includes support from the United Kingdom Office for National Statistics, the Big Data Centre of Statistics Netherlands, Badan Pusat Statistik of Indonesia and its data scientists, the research centres of Statistics Canada and Statistics Poland, the Policy Innovation Lab of the African Development Bank, the Global Pulse laboratory in Jakarta and the algorithm bank of the Organization for Economic Cooperation and Development. Other contributors to the work of the network include the National Department of Statistics (DANE) of Colombia, Statistics Denmark, the Australian Bureau of Statistics and the Federal Statistics Office of Switzerland. Several United Nations agencies, such as the Economic and Social Commission for Asia and the Pacific (ESCAP), ECE, the United Nations Environment Programme (UNEP), the United Nations Human Settlements Programme (UN-Habitat), the United Nations Children's Fund and the Food and Agriculture Organization of the United Nations (FAO), also fall within the inner circle of the Global Platform network. In fact, all members of the statistics community – as related to the Statistical Commission – are regarded as core partners of the Global Platform network and can therefore gain immediate access to the Platform.

43.   Second, there is a growing outer circle of the Global Platform network users and contributors, consisting of organizations from academia, research, civil society and the private sector. Those organizations have become trusted partners in the Global Platform network through sponsorship by core partners. Through the work of the various Global Working Group task teams, those trusted partners now actively collaborate in the Network. They include Positium, Flowminder, the Telenor Group, the Global System for Mobile Communications Association (GSMA), the University of Tokyo, Queensland University of Technology in Australia, Nielsen, Maritime Research, Galois, Cybernetica, the University of Oxford in the United Kingdom and

Microsoft Research. The Global Platform network is in principle open to collaboration between organizations from all stakeholder communities, if they are recognized as a trusted partner by one of the core platform partners. All of the activities that take place on the Platform must have the involvement of at least one of the core partners, which means that at least one member of the global statistical community will participate in any activity taking place on the Global Platform.

44.   Third, with the approval of the Advisory Board and the Technical Delivery Board, the Department of Economic and Social Affairs is able to sign strategic partnership agreements with certain organizations on behalf of the Global Working Group and the Global Platform network. Such agreements would benefit the network in the sense that the partnership would provide access to large data sources held by specific data owners, to additional platforms or to additional networks of experts in the public or private sector. A few such strategic partnerships are currently under consideration.

IV. Events organized by the Global Working Group on Big Data for Official Statistics

A.   Fifth International Conference on Big Data for Official Statistics, held in Kigali on 2 and 3 May 2019

45.   The Fifth International Conference on Big Data for Official Statistics Conference[12] was held in the main auditorium of the new training centre and data science campus of the National Institute of Statistics of Rwanda. The Conference and the new facility were officially opened by the Prime Minister of Rwanda. Many senior managers of statistical offices – mostly from the African continent – participated, but the conference also attracted many participants from other stakeholder groups, including government agencies, the private sector, academia and civil society.

46.   The Conference showcased successful projects on the use of satellite imagery data for agriculture and environment statistics, automatic identification system vessel tracking, road traffic data and value added tax data for faster economic indicators, as well as mobile telephone data for tourism, population and migration statistics. Moreover, it conveyed strategic messages on how big data affects the way business can or should be conducted in statistical institutes in the short and medium term, including the need to develop skills across offices and create partnerships with other stakeholder communities.

47.   The Conference concluded with a show of support for the Kigali Declaration, especially from the chief statisticians of the African countries, led by the Director General of the Kenya National Bureau of Statistics, who at that time was also the Chair of the Statistical Commission.

B.   Seminar on data science campus activities, held in Kigali from 29 April to 1 May 2019

48.   The seminar on data science campus activities set out training and skills-building opportunities. It included lectures, project presentations and hands-on demonstrations, as well as discussions on the development of new methodologies alongside the upgrading of quality assurance frameworks, the deployment of new technology and considerations related to security, privacy and legal matters. Also at

---

[12] See https://unstats.un.org/unsd/bigdata/conferences/2019/default.asp.

the seminar, it was demonstrated how a national statistical office can try to attract data scientists and data engineers as part of its work force.

C. Regional workshop on the use of mobile telephone data for official statistics, held in Jakarta from 11 to 14 June 2019

49.   The workshop[13] included discussions on the use of mobile telephone data in the statistical production process. About 25 statisticians from 10 Asian countries participated. The workshop included explanations about what mobile telephone data are, why and how they need to be pre-processed and how we can derive statistics from them. Badan Pusat Statistik Indonesia demonstrated how signalling data from mobile telephones is used in the production of its cross-border inward tourism statistics as well as for domestic tourism and commuting statistics. The key benefits of using mobile telephone data were examined, including for producing subnational estimates, validating sampling frames and traditional survey statistics and improving accuracy and timeliness while reducing costs and respondent burden.

D. Symposium on data science and official statistics, held in Kuala Lumpur from 15 to 17 August 2019

50.   The Symposium[14] was organized by the Global Working Group, supported by the Statistics Division, ESCAP and the Department of Statistics of Malaysia. It was attended by more than 150 statisticians and researchers from government and academia, including from about 20 developing countries. The topics discussed included the Global Platform, innovation in official statistics in Asia, particularly in China, Malaysia and the Republic of Korea, new paradigms and new skills required in official statistics, the role of a data science campus in connection with a national statistical office and the use of mobile telephone data for official statistics. As a result of the symposium, the Statistics Division and ESCAP plan to develop data solution projects using big data with selected statistical offices in Asia and the Pacific, using the Global Platform.

E. International symposium on the use of big data for official statistics, held in Hangzhou from 16 to 18 October 2019

51.   The Symposium,[15] organized by the Statistics Division together with the National Bureau of Statistics of China, was attended by about 100 statisticians from national and provincial statistical offices in China, as well as from statistical agencies of 11 countries in the Asia-Pacific Region. The symposium included presentations on the use of satellite and remote sensing data for agriculture statistics and fisheries, with presentations by the National Bureau of Statistics, Statistics Poland and FAO; the Global Platform architecture, data and services; the use of big data for Sustainable Development Goals indicators, including a presentation by the Philippine Statistics Authority on measuring the rural access index (indicator 9.1.1); and the use of mobile telephone data for population statistics, migration and urbanization, presented by the National Bureau of Statistics, as well as for tourism and commuter statistics, presented by Statistics Indonesia. A representative of the company Alibaba gave a

---

[13] See https://unstats.un.org/bigdata/events/2019/jakarta/default.asp.
[14] See https://unstats.un.org/bigdata/events/2019/isi-wsc/default.asp.
[15] See https://unstats.un.org/bigdata/events/2019/hangzhou/default.asp.

presentation on the new automated data processing architecture within the Alibaba data ecosystem through a unified data platform.

V. Next steps

A. Bureau and Advisory Board of the Global Working Group

52.   Since signing the memorandums of understanding, the role of the Advisory Board has changed from setting up the Global Platform to managing it. Because of that change in function, the decision was taken also to adjust the composition of both the Bureau and the Advisory Board of the Global Working Group. It was agreed that, given the frequency of its meetings, the Bureau benefits from remaining a relatively small but very active group, whereas the Advisory Board could include some additional members to gain a broader view of all the relevant issues.

53.   In November 2019, the Bureau and the Advisory Board of the Global Working Group agreed to adjust their memberships and methods of work as follows:

(a)   The following decisions were made with respect to the Bureau:

(i)   Statistics South Africa took over as Chair of the Bureau from the United Kingdom Office for National Statistics; Statistics Denmark remains co-Chair;

(ii)   The number of members of the Bureau was reduced from 12 to 8;

(iii)   The Bureau would continue to meet every two weeks;

(b)   The following changes were made with respect to the Advisory Board:

(i)   The United Kingdom Office for National Statistics/data science campus would remain as Chair of the Advisory Board;

(ii)   The members of the Advisory Board would comprise the United Kingdom (Chair), Brazil, China, Denmark (on behalf of task teams), Rwanda, the United Arab Emirates, ECA, the Chief Executive Officer of the Global Platform Institute, the Chair of Technical Delivery Board and several new members;[16]

(iii)   The Advisory Board will meet three or four times per year.

B. Global Platform

54.   Since the signing of the memorandums of understanding between the United Nations and each of the hubs, the Global Platform has moved to the production phase. All of the hubs have started to formulate their programmes of work for 2020, including project, training and promotional activities. In addition, the Global Working Group task teams have been instructed to reflect their use of the platform in their workplans. More generally, the global statistical community and its related stakeholder communities are encouraged to become users of, raise awareness about and support the Global Platform, which has the potential to become the driver for innovation in official statistics.

1. Innovative data solutions

55.   Projects on innovative data solutions executed by the regional hubs of the Global Platform will include: (a) pilot work on e-commerce data for price statistics and rapid economic indicators (in China); (b) satellite and other remote sensing data (using

---

[16] See https://unstats.un.org/bigdata/.

drones) for estimating crop acreage and crop yield (in China and Rwanda); (c) mobile telephone and other global positioning system sensor data for measuring human mobility and transport (in Rwanda); and (d) the use of artificial intelligence and machine learning to improve the production of official statistics (in Brazil and the United Arab Emirates).

56.   The ability to gain access to global (big) data sources in an analysis-ready state is one of key benefits of the Global Platform. Task teams can play a key role in the identification of new (global) data sources. For example, the automatic identification system team identified the global ships register, the UN Comtrade data hub and possibly a transport data hub as data sources for the global statistical community on the Global Platform.

57.   FAO is rolling out its "50 by 2030" programme[17] for the development and promotion of new cost-effective, high-quality tools for national agricultural and rural surveys, which include research projects using satellite imagery to be executed on the Global Platform in collaboration with the Global Working Group task team on Earth observation data.

58.   A collaborative project on mapping slums and informal dwellings (to meet Sustainable Development Goals indicator 11.1.1) is under consideration for execution on the Global Platform, involving UN-Habitat, the University of Warwick in the United Kingdom, the University of Tokyo and others. The project would use a combination of data solutions, including satellite, mobile telephone and citizen science data.

2.   Methodological guidance

59.   The task team on use of mobile telephone data is formulating projects around the following five statistical domains: (a) tourism and migration statistics; (b) census and dynamic population data; (c) transport and commuting statistics; (d) Sustainable Development Goals indicators on the information society; and (e) data regarding displacement in disaster contexts. The methodological developments and the lessons learned from the projects will be published in the second handbook of the task team.

60.   The task team on privacy-preserving techniques will report on practical use cases, including the cross-border sharing of detailed trade and business statistics and the use of multi-party computation in the secure processing of data from several mobile network operators. The lessons learned plus the legal aspects of using sensitive big data sources will be published in a second handbook.

3.   Training and capacity-building

61.   In 2020, the Global Working Group will continue its training and awareness-raising activities for the Global Platform, with a seminar on big data to be held in April or May at the regional hub in Hangzhou; the Sixth International Conference on Big Data for Official Statistics, to be held in Seoul from 12 to 14 May 2020; a seminar in the margins of the Conference of the International Association for Official Statistics,[18] to be held from 19 to 21 May in Livingstone, Zambia; and a workshop on the use of mobile telephone data to be held in June in Jakarta. The Global Working Group is expected to participate in sessions at the Artificial Intelligence for Good Summit,[19] to be held from 4 to 8 May in Geneva and at the third United Nations World Data Forum,[20] to be held from 18 to 21 October in Bern.

---

[17] See http://www.data4sdgs.org/resources/50-x-2030-data-smart-agriculture.

[18] See https://www.iaos-isi.org/index.php/conferences.

[19] See https://aiforgood.itu.int/.

[20] See https://www.bern.com/en/news-events/detail/uno-weltdatenforum-2020-in-bern.

62. Furthermore, the Global Working Group is considering engaging on an ambitious training programme in close collaboration with the Global Platform Institute, the regional hubs and the private sector. The programme would aim to support capability development within the institutes of the national statistics systems to equip them with the skills necessary to use big data in an ethical way to produce official statistics. Existing materials from the task teams would be used to develop the required training manuals and examinations, and to ensure the academic rigor of the training courses and their professional delivery internationally. Training courses would be delivered both online and in the classroom, including within the regional hubs, with support for the examinations that would be taken both in person (in a classroom or an examination centre) or online (proctored).

63. The training curriculum could be prepared in a modular format using two tracks – one for technical staff and one for statisticians – and could be defined for various skills levels, starting at the foundation level, followed by the practitioner and master levels. The revenue generated through the delivery of the training courses, their certification and examinations would be reinvested into the Global Platform.

4. Outreach and fundraising

64. The operational and maintenance costs of the Global Platform are expected to be recovered from projects and training activities taking place digitally on the platform or physically at the facilities of the regional hubs. If the statistics community wants the Global Platform to succeed, the use of the platform needs to be promoted and activities need to be well funded.

65. The Global Platform is still not very well known in the wider data community, and even within the global statistics community awareness-raising is needed. Outreach and promotional activities are therefore necessary. All partners of the Global Platform network can function as ambassadors and raise awareness about the Global Platform at all of the events in which they are involved. These events also include the formal meetings of the broad network of expert groups[21] under the Statistical Commission.

66. The Global Platform will need the active support of civil society, academia and the private sector. For example, collaboration on the Data-for-Now[22] initiative by the Global Partnership for Sustainable Development Data [23] and the Sustainable Development Solutions Network[24] with the Global Platform regional hub in Rwanda is very important in order to sustain the platform. In that respect, it is very encouraging that, in the recent report entitled "Counting on the World to Act",[25] the Thematic Research Network on Data and Statistics and the Sustainable Development Solutions Network recommend that initiatives in this data space "build upon and complement the Global Platform".

67. Finally, formal agreements on strategic partnerships with certain organizations and institutes could bolster the use of the Global Platform. Such agreements are being considered with FAO, UNEP and the Centre for Data Science [26] at Queensland University of Technology.

---

[21] See https://unstats.un.org/unsd/statcom/groups/.

[22] See http://www.data4sdgs.org/data4now.

[23] See http://www.data4sdgs.org/.

[24] See http://unsdsn.org/ and https://www.sdsntrends.org/.

[25] See https://countingontheworld.sdsntrends.org/2019/time-to-act-a-roadmap/.

[26] See https://research.qut.edu.au/qutcds/.

VI. Action to be taken by the Statistical Commission

68. **The Statistical Commission is invited:**

(a) **To endorse the Global Platform as a collaborative environment for data innovation by the global statistics community and all of its partners, including United Nations agencies, government agencies, civil society organizations, research institutes, academia and private sector companies;**

(b) **To endorse the governance structure and business model of the Global Platform presented in the present report, under the overall auspices of the Commission;**

(c) **To support the planned data solutions projects and methodological guidance on the Global Platform;**

(d) **To support the envisioned training programme delivered at scale by the Global Working Group and the Global Platform;**

(e) **To urge the Global Platform network, including the global statistics community, to promote and mobilize resources for and support the Global Platform.**

Annex I

Membership of the Global Working Group on Big Data for Official Statistics

States	Organizations
Australia	African Development Bank
Bangladesh	Caribbean Community
Brazil	Economic Commission for Africa
Cameroon	Economic Commission for Europe
Canada	Economic and Social Commission for Asia and the Pacific
China	Eurostat
Colombia	Food and Agriculture Organization of the United Nations
Denmark	Global Pulse
Egypt	International Telecommunication Union
Georgia	International Monetary Fund
Germany	Organization for Economic Cooperation and Development
Indonesia	Statistical Centre for the Cooperation Council for the
Ireland	Arab Countries of the Gulf
Italy	Statistical Institute for Asia and the Pacific
Mexico	United Nations Statistics Division
Morocco	Universal Postal Union
Netherlands	World Bank
Oman	
Pakistan	
Philippines	
Poland	
Republic of Korea	
Rwanda	
Saudi Arabia	
South Africa	
Switzerland	
United Arab Emirates	
United Kingdom of Great Britain and Northern Ireland	
United Republic of Tanzania	
United States of America	

Annex II*

Kigali Declaration

Recalling the mandate[1] of the United Nations Global Working Group (GWG) on Big Data for Official Statistics to provide strategic vision, direction and coordination for a global programme on Big Data for official statistics, including for the compilation of the Sustainable Development Goal indicators, and to promote practical use of Big Data sources, while building on existing precedents and finding solutions for the many existing challenges;

Recalling the adoption of the Cape Town Global Action Plan for Sustainable Development Data[2] (CTGAP) by the Statistical Commission at its 48th session in March 2017 to support the implementation of the 2030 Agenda[3] for Sustainable Development, which requires the collection, processing, analysis and dissemination of an unprecedented amount of data and statistics at local, national, regional and global levels and by multiple stakeholders.

Highlighting that the CTGAP calls upon the global statistical community to take action on the strategic area of modernizing and strengthening the national statistical systems with a focus on modernizing the governance and institutional framework; on applying statistical standards and new data architecture for data sharing, exchange and integration; and on facilitating the use of new technology and new data sources in statistical production processes;

Reiterating the Bogota Declaration[4] of the GWG that the implementation of the global data collaborative as a federated system will place the community of official statistics at the heart of modern trusted data usage and information technology, which will offer both developed and developing countries opportunities to realize the benefits of multisource data, including Big Data, administrative data, census data and survey data, to better understand economic, environmental and societal changes without investment in expensive technologies;

Reemphasizing the recommendation of the Bogota Declaration that global collaboration, facilitated by the global platform with the potential to accommodate many different types of trusted data, trusted services and trusted applications, should (a) make it easy for all nations to gain value by participating in the global network; (b) deliver a marketplace and a flexible cloud-based technology infrastructure to allow trusted data, methods, services and applications to be shared as a public good where useful and legally possible; and (c) develop transparent partnership agreements with private- and public-sector organizations so that network partners contribute and derive value through a business model which is individually sustainable for all stakeholders and ensures access to trusted data;

The GWG at its 5th Global Conference on Big Data for Official Statistics in Kigali

Recommends:

• the establishment of an independent entity under the UN umbrella, which plans and delivers the UN Global Platform as it evolves, which is able to generate and

_____

* The present annex is being issued without formal editing.

[1] See Annex I of the Report of the GWG to the 46th session of the UN Statistical Commission in 2015 (E/CN.3/2015/4).

[2] See https://unstats.un.org/sdgs/hlg/cape-town-global-action-plan/.

[3] See https://sustainabledevelopment.un.org/post2015/transformingourworld.

[4] See Annex II of the Report of the GWG to the 49th session of the UN Statistical Commission in 2018 (E/CN.3/2018/8).

accept funding, and undertakes operational and commercial activities such as hiring staff, running projects, contract negotiation and platform evolution;

• the establishment of a Governance Board for the UN Global Platform, which will decide on the overarching policies and principles, and will report to the UN Statistical Commission;

• that the UN Global Platform should have multiple hubs in various regions of the world to build capacity and deliver regional support including for official statistics, working under a well-defined governance framework;

Underlines in this regard that the UN Global Platform:

• should support and connect the global statistical community and all its partners, including the UN agencies as well as government agencies, private sector companies, research institutes, academia and civil society organizations;

• should be based on networking and marketplace principles, to facilitate the exchange, development and sharing of data (especially global data sources), methods, tools and expertise, and accelerate data innovation for official statistics and for the compilation of Sustainable Development Goals indicators;

• should provide technological infrastructure as a service for all UN Member States, especially Least Developed Countries and Small Island Development States, to facilitate their access to global data sets and state of the art tools and services;

• should engage with the global statistical community and all its partners to provide an environment for capacity development activities to teach and develop new competencies in the areas of Artificial Intelligence, data science, machine learning and privacy-preserving techniques;

Recommends specifically for the African context:

• that one or more hubs of the UN Global Platform is established in Africa, which:

  ◦ will support the work of the international and regional agencies in Africa, especially the pan-African institutes, to innovate and build capacity in national statistical systems;

  ◦ will support and enable international collaboration on existing and new data innovation projects for official statistics and the Sustainable Development Goals indicators in Africa; and

  ◦ will facilitate capacity development activities in Africa for data innovation using global data sets, state-of-the-art methods and the latest technologies.

The Government of Rwanda has offered to host a hub of the UN Global Platform in Rwanda at the National Institute of Statistics of Rwanda Training Centre and Data Science Campus. The hub will support capability development work of international and regional agencies in Africa in Big Data for official statistics.

United Nations Global Working Group assembled for the 5th International Conference on Big Data for Official Statistics, Kigali, Rwanda, May 2019.

———————