4TH INTERNATIONAL CONFERENCE ON
BIG DATA
for Official Statistics
8-10 NOV 2017 **BOGOTA, COLOMBIA**

CONFERENCE AGENDA

Please note that for specific time slots there are parallel sessions that are being held in both the Auditorium and the Main Room

WEDNESDAY, 8 NOV

08:00   **Registration – Coffee**

Opening Session 09:00
Welcome **by Mr. David Luna Sánchez,** *Minister of Information and Communication Technologies, Colombia*

Mr. Mauricio Perfetti del Corral, *Director General, DANE, Colombia*
Mr. Martín Santiago, *UN Resident Representative in Colombia (tbc)*
Mr. Rafael de la Cruz, *Representative for Colombia, Inter-American Development Bank*
Mr. Niels Ploug, *Chair of the UN Global Working Group on Big Data*

Innovation and modernization of national statistics systems through "trusted data collaboratives"
The Cape Town Global Action Plan for Sustainable Development Data emphasized among others the strengthening of innovation and modernization of national statistical systems. This innovation effort calls for a rethinking of the partnerships of the community of official statistics with private sector, academia and civil society through an interconnected ecosystem of data and technology collaborations at national, regional and global level. In this context, "Trusted Data Collaboratives" are a new way of working together with proper definition of the interests of the various stakeholders, proper definition of responsibilities and access, and appropriate protocols to safeguard confidentiality.

Keynote Address 09:50
Mr. Alexandre Barbosa, *Head of the Regional Center for Studies on the Development of the Information Society*

High-level Panel discussion 10:15
Moderator: **Mr. Niels Ploug,** *Chair of the UN Global Working Group on Big Data*
Panel: **Mr. Carlos Felipe Prada Lombo,** *Deputy Director, DANE, Colombia*
              Ms. Heather Savory, *Deputy National Statistician, UK ONS*
              Ms. Sylvie Michaud, *Deputy Chief Statistician, Statistics Canada*
              Mr. Gerardo Leyva, *National Institute of Statistic and Geography, INEGI, Mexico*
              Mr. Albrecht Wirthmann, *Director Big Data, Eurostat*

11:25   **Coffee Break**

Defining the context – the 2030 Agenda, "leaving no one behind" and official statistics
On 25 September 2015 world leaders committed to the 2030 Agenda for Sustainable Development, including many ambitious goals and targets to be achieved by 2030. The statistical community was charged with defining appropriate indicators to monitor the

progress towards these targets. It was stressed that differentiation by population groups, sub-national location and smaller time intervals ("Leaving no one behind") would make the information base more useful for policy decisions. The 2030 Agenda explicitly calls for enhancing capacity building to support national plans to implement the sustainable development goals. Along with the SDG monitoring, the underlying microdata should also become accessible through national, regional and international open data platforms and become discoverable through standard metadata documentation. These platforms should apply open data protocols for the creation and use of interoperable APIs based on ISO standards to allow for trusted data collaboratives, improved governance, citizens engagement and inclusive development and innovation (see open data charter).

High-level Panel discussion 11:45

Moderator: **Mr. Ivo Havinga,** *Assistant Director, United Nations Statistics Division, UNSD*
Panel: **Ms. Diana Nova,** *SDG Working Group, DANE*
               Ms. Alla Morrison, *Program Officer, Financial Group, World Bank*
               Mr. Fessou Lawson, *Principal Statistician, African Development Bank*
               Mr. Philipp Schönrock, *International Strategic Thinking Center, CEPEI, Colombia*
               Mr. Aditya Agrawal, *Director Of Data Ecosystems, Global Partnership for Sustainable Development Data*
               Mr. Emmanuel Letouzé, *Director, Data-Pop Alliance*

13:00   Lunch

New ways of working together (Parallel Sessions)

Session A: Local hubs and City Data - Auditorium 14:30

The World Council on City Data (WCCD) is the global leader in standardized city data - creating smart, sustainable, resilient, and prosperous cities. WCCD hosts a network of innovative cities committed to improving services and quality of life with open city data and provides a consistent and comprehensive platform for standardized urban metrics. WCCD is a global hub for creative learning partnerships across cities, international organizations, corporate partners, and academia to further innovation, envision alternative futures, and build better and more liveable cities. WCCD is implementing ISO 37120 Sustainable Development of Communities: Indicators for City Services and Quality of Life.

Moderator: **Mr. Ivo Havinga,** *Assistant Director, United Nations Statistics Division, UNSD*
Panel: **Mr. James Patava,** *World Council of City Data*
               Mr. Andrés Ortiz, *Planning Secretary, City of Bogota*
               Ms. Sylvie Michaud, *Statistics Canada*
               City of Buenos Aires *(tbc)*

Session B: Data collaboratives on SDG monitoring, Part I - Main Room 14:30

The most promising way forward of compiling data for SDG monitoring is the integration of Big Data coming from new technologies with traditional data, in order to produce relevant high-quality information, with more detail and at higher frequencies to foster and monitor sustainable development. This implies also an increase in accessibility to data through much more openness and transparency, which should ultimately more empower people for better policies, better decisions and greater participation and accountability, leading to better outcomes for people and planet.

Moderator: **Mr. Ronald Jansen,** *United Nations Statistics Division, UNSD*
Panel: **Mr. Joaquim Barris,** *Climate Change indicators, UNFCC*
               Ms. Alla Morrison, *Collaborative Data Innovations for Sustainable Development, World Bank*
               Mr. Misha Lokshin, *Household Survey Network, World Bank*
               Mr. Manuel Francisco Lemos, *Open Data Portals for SDG monitoring, ESRI*
               Ms. Dias Rahwidiati, *UN Global Pulse*

BigData
UN Global Working Group

4TH INTERNATIONAL CONFERENCE ON
BIG DATA
for Official Statistics
8-10 NOV 2017   **BOGOTA, COLOMBIA**

16:00   Coffee Break

New ways of working together (Parallel Sessions)

Session A: National Data Centers involving official statistics - Auditorium   16:20
Digital services are becoming increasingly important in our lives. Services such as cloud, mobile apps, and other digital applications are facilitated by data centers. They have become the main enabler of the digital economy by facilitating a wide range of activities across government, business and society. Data centers therefore are an important part of the national critical infrastructure. As enablers of the digital economy, data centers play an important role when it comes to trust. Data should not only be accessible and available 24/7, secure data storage and privacy must be guaranteed. Data centers provide a platform for organizations to compute, run and store their services and data. In the Netherlands, municipalities join forces with Statistics Netherlands in urban data centers to use data more effectively in local administration.

Moderator: **Mr. Bert Kroese,** *Statistics Netherlands*
Panel: **Mr. Jorge Caldas Gallo,** *Alianza Caoba, Colombia*
                  Ms. Heather Savory, ONS, *United Kingdom*
                  Ms. Sylvie Michaud, *Statistics Canada*
                  Mr. Setia Pramana, *BPS Indonesia*

Session B: Data collaboratives on SDG monitoring, Part II - Main Room   16:20
The most promising way forward of compiling data for SDG monitoring is the integration of Big Data coming from new technologies with traditional data, in order to produce relevant high-quality information, with more detail and at higher frequencies to foster and monitor sustainable development. This implies also an increase in accessibility to data through much more openness and transparency, which should ultimately more empower people for better policies, better decisions and greater participation and accountability, leading to better outcomes for people and planet.

Moderator: **Mr. Misha Lokshin,** *World Bank*
Panel: **Mr. Alexandre Barbosa,** *CETIC/NIC, Brazil*
                  Mr. Joao Azcevedo, *World Bank*
                  Mr. Esteban Peláez Gómez, *Fundación Corona*
                  Ms. Ana Lucía Martínez & Mr. Carlos Mazariegos, *OPAL Project*

17:30   Closing Day 1

THURSDAY, 9 NOV

Opening   09:00
Mr. Niels Ploug, *Chair of the UN GWG on Big data*

Standards for Trusted Data Collaboratives
Within the community of official statistics "trusted data" is defined in terms of 'compliance with quality standards'. There are national quality assurance frameworks, statistical codes of practice and compliance with international standards, such as the System of National Accounts. IMF developed a Special Data Dissemination Standard and a General Data Dissemination Standard, which can be generally used as a measure for "trusted data". At a technical level, the statistical community has defined protocols for data exchange and

interoperability of data systems, like SDMX and DDI. More broadly, the private sector has defined security standards and associated protocols for data transmission, data storage and the like. These are taken up as ISO standards. In the business world, assurance is given through the certification that something is ISO compliant.
What kind of certification do we need for a "trusted data collaborative"?

Panel discussion 09:10
Moderator: **Mr. Brant Zwiefel,** *Principal Architect, Microsoft*
Panel: **Mr. Paulo Cunha,** *Amazon Web Services*
Mr. Dan Bogdanov, *Cybertecnica, Brazil*
Mr. Mark Craddock, *UK ONS*
Mr. Misha Lokshin, *World Bank*
Mr. Carlos Julio León, *MINTIC, Colombia*
Mr. Albrecht Wirthmann, *Eurostat*
Mr. Momar Kouta, *African Development Bank*

11:00   Coffee Break

Use cases for Data, Services and Applications

Earth observation data and official statistics 11:15
The GWG task team on satellite imagery, geospatial data and remote sensing developed a handbook that contains information on sources of Earth observation data, methodologies for producing crop statistics and other statistics through the use of satellite imagery data, outlines of pilot projects and guidance for national statistical offices in exploring the use of Earth observation data for the first time. The pilot projects include an application of satellite imagery data in the production of agricultural statistics. This session also discusses a hands-on course to teach methods for using Earth observation data in generating agricultural crop statistics and in monitoring the SDGs.

Moderator: **Ms. Sylvie Michaud,** *Statistics Canada*
Panel: **Mr. Brant Zwiefel,** *Microsoft*
Mr. Paulo Cunha, *Amazon Web Services (AWS)*
Mr. Manuel Francisco Lemos, *ESRI*
Ms. Kerrie Mengersen, *Queensland University of Technology, Australia*
Mr. Zhou Wei, *NBS, China*
Ms. Yineth Acosta & Ms. Sandra Liliana Moreno, *DANE, Colombia*

13:00   Lunch

Scanner data and on-line data and official statistics 14:30
This session will provide information on how to access online data, selecting data sources, preparing raw data for use, classifying data and processing data for use in the CPI. It also discusses different methodologies, describe the status of scanner data implementation in different countries. Further, it gives a status update on scanner and online data integration projects and provides guidance for NSOs considering using this data source for the first time.

Moderator: **Mr. Ivo Havinga,** *United Nations Statistics Division, UNSD*
Panel: **Mr. Michael Holt,** *Statistics Australia*
Mr. Niels Ploug, *Statistics Denmark*
Mr. Jonathan Wylie, *Statistics Canada*
Mr. Manuel Bertolotto, *Pricestat.com*
Ms. Dana Childerhose, *Nielsen*

BigData
UN Global Working Group

16:00 **Coffee Break**

Mobile phone data and official statistics 16:15
This session gives an overview of data generated by mobile communication technologies and choices, which clarifies the trade-offs between size, complexity and usefulness. The session will discuss the importance to understand stakeholders and partnership models for mobile data projects and the logical order of steps in the process of data extraction. It will also discuss how to calculate tourism statistics and how to identify tourism indicators, calibration and inference.

Moderator: **Mr. Margus Tiru,** *Positium, Estonia*
Panel: **Ms. Julieth Solano,** *DANE, Colombia*
Mr. Juan David Olarte Torres, *MINTIC, Colombia*
Mr. Jose Luis Fajardo, *CLARO, Colombia*
Mr. Fernando Reis, *Eurostat*
Mr. Dan Bogdanov, *Cybertecnica, Brazil*

17:45 **Closing Day 2**

FRIDAY, 10 NOV

Opening 09:00
Mr. Niels Ploug, *Chair of the UN GWG on Big data*

Use cases for Data, Services and Applications (Parallel Sessions)

Session A: Multi-source data for official statistics - Main Room 09:10
Trusted Data Collaboratives is about the use of big data and its integration with administrative sources, geospatial information and traditional survey and census data. The use of multi-source data requires collaboration by a number of stakeholders, such the statistical office, government agencies, research institutes, civil society and private sector. Assuring quality of the outcome therefore requires quality assessment at all levels of the collaborative.

Moderator: **Mr. Niels Ploug,** *Statistics Denmark*
Panel: **Ms. Mara Brigitte Bravo & Mr. Javier Mauricio Jacome,** *DANE, Colombia*
Ms. Margarita Ramirez, *DANE, Colombia*
Ms. Cornelia Hammer, *IMF*
Mr. Marcelo Pitta, *NIC, Brazil*

Session B: Enterprise and Trade Data Lake - Auditorium 09:10
A data lake is a storage repository that holds a vast amount of raw data in its native format until it is needed. While a hierarchical data warehouse stores data in files or folders, a data lake uses a flat architecture to store data. Each data element in a lake is assigned a unique identifier and tagged with a set of extended metadata tags. When a business question arises, the data lake can be queried for relevant data, and that smaller set of data can then be analyzed to help answer the question. The term data lake is often associated with Hadoop-oriented object storage. In such a scenario, an organization's data is first loaded

BIG DATA

for Official Statistics

8-10 NOV 2017   **BOGOTA, COLOMBIA**

into the Hadoop platform, and then business analytics and data mining tools are applied to the data where it resides on Hadoop's cluster nodes of commodity computers. Like big data, the term data lake is sometimes disparaged as being simply a marketing label for a product that supports Hadoop. Increasingly, however, the term is being accepted as a way to describe any large data pool in which the schema and data requirements are not defined until the data is queried.

Moderator: **Mr. Setia Pramana,** *BPS Indonesia*
Panel: **Enterprise Data Lake**
  Ms. Irene Salemink, *Statistics Netherlands*
  Trade and Transport Data Lake
  Mr. Ronald Jansen, *UNSD*
  Mr. José Anson, *Universal Postal Union, UPU (via Skype)*
  Mr. Sai Ananthanarayan, *International Civil Aviation Organization, ICAO*
  Mr. Tim de Jong, *Statistics Netherlands*

10:40   Coffee Break

Modernizing official statistics in Latin America and the Caribbean – the way forward 11:00
Moderator: **Mr. Gerardo Leyva,** *INEGI, Mexico*
Panel: **Mr. Carlos Felipe Prada Lombo,** *DANE, Colombia*
  Mr. José Antonio Mejia, *IADB*
  Mr. Statchel Edwards, *Antigua and Barbuda*
  Mr. César Vicuña, *INEC, Ecuador*
  Ms. Philomen Harrison, *Caricom*

13:00   Lunch

Proof of Concepts for Collaboratives for Data, Services and Applications 14:30
Moderator: **Ms. Heather Savory,** *ONS, United Kingdom*
  Mr. Mark Craddock, *UK ONS - Global Platform on Data, Services and Applications*
  Mr. Louis Kouakou, *African Development Bank - Regional Platform: African Information Highway*

16:00   Coffee Break

Bogota Statement on Trusted Data Collaboratives 16:20
Moderator: **Mr. Niels Ploug,** *Chair of the UN Global Working Group on Big Data*
  Ms. Heather Savory, *Deputy National Statistician, ONS, United Kingdom*
  Ms. Sylvie Michaud, *Deputy Chief Statistician, Statistics Canada*
  Ms. Philomen Harrison, *Director, CARICOM*
  Mr. Bert Kroese, *Deputy Director-General, Statistics Netherlands*
  Mr. Carlos Felipe Prada Lombo, *Deputy Director, DANE, Colombia*
  Mr. Ivo Havinga, *Assistant Director, United Nations Statistics Division, UNSD*
  Mr. Georgy Oksenoyt, *Deputy Head, Rosstat*

17:30   Closing of the Conference by the Ministry Mr. David Luna Sánchez
  Minister of Information and Communication Technologies, Colombia