Summary of the Fifth UN World Data Forum: 12-15 November 2024

"Datutopia" – the vision of an innovative, inclusive, and enabled data ecosystem supporting sustainable development for all – was one of the defining ideas of the Fifth UN World Data Forum (UNWDF 2024). The vision inspired lively debates on the key pillars of such a system, addressing questions such as how to build trust and develop standards for ethical data use or how to craft lasting and effective partnerships, as well as discussions on the next big steps that could be enabled by artificial intelligence (AI) to mobilize data for communities.

Participants noted that, seven years after its inception, the World Data Forum has come of age. Once characterized by novel and sometimes confusing discussions on how National Statistical Organizations (NSOs) can open themselves up to work with diverse stakeholders and alternative data sources, the event is now teeming with activity led by firmly established partnerships and speakers from diverse backgrounds sharing success stories and new ideas on how data can support the Sustainable Development Goals (SDGs) while leaving no one behind. Panel sessions were dominated by presentations on issues that used to be on the fringe of the discussions, such as mainstreaming citizen data and integrating geospatial data and statistics to better serve marginalized communities. Plenary sessions tackled the big questions on how to build effective, trustworthy, and holistic data ecosystems that benefit from AI while safeguarding individual and collective rights to privacy and data ownership.

Meeting for the first time in a country of the Latin American and Caribbean (LAC) region, UNWDF 2024 was organized into four thematic areas:

- Innovation for better and more inclusive data;
- Maximizing the use and value of data for better decision making;
- Building trust, protection, and ethics in data; and
- Effective Partnerships for better data and a more equitable data ecosystem.

Each theme was addressed in a high-level plenary and a series of panel sessions. The event also included a large exhibition space showcasing partnerships, innovative data uses and successful projects from Medellín, Colombia and the LAC region.

The Forum also provided an opportunity to continue discussions on the Digital Compact adopted as part of the Pact for the Future during the Summit of the Future in July 2024. The Medellín Framework for Action – the UNDWF 2024 outcome document – intends to update the Cape Town Global Action Plan (CTGAP) to reflect the dynamic and evolving discussions on data for sustainable development.

UNDWF 2024 met from 12-15 November 2024, in Medellín Colombia, bringing together more than 2,000 participants from NSOs, academia, civil society and global data partnerships.

Brief History of the World Data Forum

Following a recommendation from the UN Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development, contained in the 2014 report, A World That Counts: Mobilizing the Data Revolution for Sustainable Development, the UN Statistics Division (UNSD) decided to convene a World Forum on Sustainable Development Data. It was envisaged as a platform to intensify cooperation between representatives of government, intergovernmental organizations and civil society, and various professional groups, such as national statistical offices (NSOs), information technology (IT) and geospatial information managers, and data scientists, among others. The UN World Data Forum consists of high-level plenary sessions and parallel breakout events clustered around specific themes.

First UN World Data Forum: Participants at the first Forum, which took place from 15-18 January 2017, in Cape Town, South Africa, agreed on the Cape Town Global Action Plan (CTGAP). The Action Plan sets out a framework for member countries to assess, build, and strengthen NSO capacity, and is divided into six strategic areas:

- coordination and strategic leadership on data for sustainable development;
- innovation and modernization of national statistical systems;
- strengthening of basic statistical activities and programmes, with a particular focus on addressing the monitoring needs of the 2030 Agenda for Sustainable Development;
- dissemination and use of sustainable development data;
- multi-stakeholder partnerships for sustainable development data; and
- mobilization of resources and coordination of efforts for statistical capacity building.

The UN Statistical Commission (UNSC) adopted the Plan in March 2017.

Second UN World Data Forum: This Forum, which convened from 22-24 October 2018 in Dubai, United Arab Emirates, adopted the Dubai Declaration on "Supporting the Implementation of

the Cape Town Global Action Plan for Sustainable Development Data." Through the Declaration, Forum participants resolved to ensure that quality, relevant, timely, open, and disaggregated data "by income, sex, age, race, ethnicity, migration status, disability, geographic location, and other characteristics relevant in national contexts" are available and accessible to all users. In addition, the Dubai Declaration called for establishing an innovative funding mechanism that would be responsive to the priorities of national data and statistical systems, with the goal of mobilizing funds, and activating partnerships and funding opportunities to strengthen the capacity of national data and statistical systems.

Virtual UN World Data Forum: The Third World Data Forum was expected to take place in October 2020, in Bern, Switzerland. However, due to the COVID-19 pandemic, a reduced virtual event took place from 19-21 October. The virtual Forum resulted in a Declaration on the "Global Data Community's Response to COVID-19," which appeals to the whole data community to come together to support the response to COVID-19 and accelerate action on the Sustainable Development Goals (SDGs). It resolves to ensure trust in data and data privacy and increase investments in data to respond more effectively to the COVID-19 pandemic and future disasters. It also renewed the call made in the Dubai Declaration for an innovative funding mechanism to help implement the CTGAP.

Third World Data Forum: The Third World Data Forum convened in person from 3-6 October 2021, in Bern, Switzerland. The Bern Data Compact for the Decade of Action on the SDGs, launched during the closing plenary, summarizes commitments to accelerated action on implementation of the CTGAP and common actions to, among others: develop data capacity; establish data partnerships; produce data to leave no one behind; and build trust in data.

Fourth World Data Forum: The Fourth World Data Forum met from 24-27 April 2023, in Hangzhou, China, under the theme "Towards Data that Empowers our World." Delegates launched the Hangzhou Declaration, which reflects on, among other issues, high-quality, timely, open, and inclusive data required to achieve the SDGs by 2030. The Declaration commits Parties to renew commitments to working as a global community under the UN and to revitalize the CTGAP through collective action to address challenges on the path towards modernized and stronger data and statistical systems that support achievement of the SDGs.

Report of the Fifth UN World Data Forum

The Fifth UN World Data Forum (UNWDF 2024) opened on Tuesday, 12 November 2024, with an opening ceremony and introductory panel discussion. The program was organized into four themes, each consisting of a High-level Plenary and breakout sessions. This briefing note summarizes discussions from the Plenary sessions.

Opening Ceremony

UNWDF 2024 Moderator Alexandra Rentería welcomed participants on Tuesday, 12 November, and provided an overview of the Forum's agenda. Opening the Forum, Piedad Urdinola Contreras, Director of Colombia's National Administrative Department of Statistics (DANE), noted that Medellín is the first city in the Latin American and the Caribbean (LAC) Region to host the Forum. She noted the event would provide an opportunity to share the Region's experiences in using data to promote human and sustainable development and quality of life of all citizens. She said DANE is strengthening capacities to enhance data quality and interoperability, make data more inclusive, and ensure citizens have access to and understand data.

Via video, UN Deputy Secretary-General Amina Mohammed said UNWDF 2024 embodies the outcomes of the UN Summit of the Future and its Global Digital Compact, noting that the UN system needs better data to strengthen its work on social support to leave no one behind.

Stefan Schweinfest, UNSD Director, previewed the Forum's four themes on data innovation, maximizing data use, building trust in data, and effective data partnerships. Ola Awad, Co-Chair of the High-level Group for Partnership, Coordination, and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development (HLG-PCCB), said the intended review of the CTGAP is an opportunity to create a full range of tools for monitoring and reporting on the SDGs.

Philipp Schönrock, CEPEI, moderated a panel on expectations. Johannes Jütting, Partnership in Statistics for Development in the 21st Century (PARIS21), presented the "Commit to Data" initiative, which aims to generate tangible data-driven impacts for individuals. He also noted that the Mayor of Bogotá submitted a commitment to create a local statistics mechanism.

Miosotis Riva Peña, Director of the Dominican Republic's National Statistics Office, outlined the LAC Region's contributions to the global data dialogue, including the development of regional standards for gender representation. Mireia Villar Forner, UN Resident Coordinator for Colombia, called attention to Colombia's progress in collecting data on marginalized communities not represented in official statistics.

Manuel Francisco Lemos, ESRI Colombia, highlighted Colombia's journey from a country where data was scarce 25 years ago to a regional leader, noting the importance of high quality and reliable data for different decision-making scenarios. Maria Fernanda Galeano Rojo, Medellín Mayor's Office, outlined data-driven projects that aim to transform Medellín, including an innovation district, a smart city strategy, a weather and air quality hazard early warning system, and a dynamic traffic management system.

Introduction

On Wednesday morning, 13 November, Piedad Urdinola Contreras, DANE, outlined the Forum's agenda and themes and invited participants to use the event to discuss ways to improve coordination and collaboration among NSOs, increase funding for data systems, and tackle challenges for inclusive data in the face of crises like the COVID-19 pandemic.

Federico Gutiérrez, Mayor of Medellín, welcomed participants and told the story of how data helped Medellín to evolve from conflict to prosperity, including increased quality of life, reduced homicide rates, and growth in areas where data guided decision making. He explained that the city's social data map is more important than political maps to allocate investments and provided examples of targeted projects such as schools and parks in disadvantaged areas.

Innovation for Better and More Inclusive Data

On Wednesday morning, 13 November, Stefan Schweinfest, UNSD, invited a panel of experts to comment on the vision of Datutopia – an innovative, inclusive, and enabled data ecosystem supporting sustainable development for all, noting it should leverage open data, AI, and other new technology to address humanitarian challenges and ensure no one is left behind.

On envisioning and inclusive data ecosystem, Ola Awad, Palestinian Central Bureau of Statistics and Co-Chair, HLG-PCCB, said such an ecosystem must ensure all individuals and communities are well recognized and empowered through data. She noted that enablers like NSOs must work hand in hand with contributors in civil society, the private sector, academia, and others. On working in Palestine, she highlighted the need for innovation and partnership in conflict situations as data cannot be collected in the field and existing surveys lose relevance.

Prem Ramaswami, Google, stressed the need to understand how to rebuild the data commons with data from different sources using a common language to enable natural language searches. He provided examples including the use of Google's Data Commons platform, which aims to combine data from different sources to develop a better understanding of how people are affected by rising temperatures due to climate change.

Luisa Franco Machado, UN Youth Office, said data colonialism commodifies human lives by using data as a tool for control, suggesting that decolonizing data requires placing people in control of their narratives. She advocated for community-led government that views community members as actors, rather than data points, and embeds bias checks and fairness matrices to capture the complexities of people's identities. Explaining that today's youth are the first generation that is subject to constant data harvesting, she called for a transformation of data governance. On inclusion, she stated that the so-called gender gap is not an absence of data, but a consequence of deep-rooted biases on whose experiences count.

Steve MacFeely, Organisation for Economic Co-operation and Development (OECD), stated that public data must be at the heart of the system, which he said raises a fundamental question about who should decide which data are public goods. He also stressed that everyone must be able to understand how the data system is governed. He added that this can lead to unexpected challenges, stating that a re-design of OECD's data explorer to facilitate access and remove data fragmentation confused users.

On innovations for more effective data use, panelists suggested: harnessing the intelligence of people through citizen science; collaboration; breaking down the wall between data scientists and data users; embedding intersectional perspectives in data; and foresting community data governance.

Maximizing the Use and Value of Data for Better Decision Making

On Wednesday afternoon, 13 November, moderator Christina Veléz, EAFIT University, Medellín, asked panel members to reflect on their experiences in using national statistics and other data for better decisions and provide examples for how data use can be improved.

Leota Aliielua Salani, Samoa Bureau of Statistics, described his country's journey from making decisions without data 25 years ago to using statistics to push policy makers out of their comfort zone for decisions that improve lives. Noting the limited capacity of NSOs in small island developing States (SIDS), he stressed that regional collaboration is key, pointing to a new regional data hub as an example. He added that using media, especially social media, is important to disseminate data to all population groups.

Emphasizing the value of diversity to societies, Piedad Urdi-nola, DANE, said data must reflect diversity to understand what is happening on the ground. She noted positive experiences in compiling data and other information in partnership with communities and outlined Colombia's pioneering economic census that uses an ethnic regional focus to capture the needs of marginalized territories.

Rolando Ocampo, UN Economic Commission for Latin America and the Caribbean (ECLAC), outlined how combining national, sub-national and municipal statistics with geospatial data can generate better knowledge for decision making as it provides a more holistic understanding of relationships between data points and conditions on the ground. He described ECLAC's efforts to support the use of georeferenced data sets, including a regional platform to disseminate geospatial data and a regional Gender Observatory.

Esteban Gallego Restrepo, Comfenalco, explained how his organization uses disaggregated data to manage a private family welfare fund, stressing the importance of local socioeconomic data and data on migration movements to enable appropriate community services. He described an example of combining data from multiple sources that helped to understand that access to housing is often restricted through high individual debt, rather than housing supply.

Christopher Maloney, Hewlett Foundation, mentioned the African Data Sphere Initiative, which works with multiple partners in countries to make data more widely accessible and improve the safety, transparency, and trust in data for decision making. Noting Hewlett's support to the UN Collaborative on Citizen Data, he also described pilot projects mobilizing alternative data sources for SDG reporting in Ghana and Kenya.

Focusing on gender, Caren Grown, Brookings Institution, said how we define gender determines how our mental models and preconceptions are embedded in the institutions and data we use to make decisions. Noting challenges in collecting data that capture the nuances of how people live, she outlined how gender data has improved impacts. For example, she said using time-use data to include women's unpaid care work in national income and productivity accounts has led to more equitable targeting of support measures. She suggested using disaggregated data on violence against women to create maps displaying where women are at risk or whether they can get help in case of emergency.

On the potential role of artificial intelligence (AI) in maximizing the use of data, panelists: cautioned that AI still needs to evolve to address data safety concerns and ensure people remain in control of data use. They also stressed the need for capacity and infrastructure development to empower countries to use AI rather than react to it and underscored that AI is biased because most of the training data is in English and from rich Western countries.

Building Trust, Protection and Ethics in Data

This session took place on Thursday, 14 November. Moderator Christina Veléz, EAFIT University, Colombia, divided panelists into two teams and asked one team to speak in favor of and the other team to speak against a series of propositions regarding trust, protection, and ethics in data. Team A, speaking in favor, included the following speakers:

- Roeland Beerten, National Bank of Belgium;
- Dominik Rozkrut, Statistics Poland; and
- Priscilla Idele, United Nations Population Fund (UNFPA).
  Team B, speaking against the propositions, included:
- Natalia Carfi, Open Data Charter (ODC);
- Stefaan Verhulst, The Governance Lab (GovLab); and
- Angela Yesenia Olaya Requene, Ministry of Science, Technology and Innovation of Colombia.

To set the scene, Shaida Badiee, Open Data Watch, discussed commitments on data ethics from several organizations, including the UN's Fundamental Principles for Statistics, UNEP's commitment for better governance on data describing children, and the exchange of best practices among businesses.

On a proposition stating that the responsibility to determine ethical standards for data use lies with NSOs, Team A said NSOs are deeply rooted in an academic and ethical tradition that views data as a public good that should serve all of society. Speakers also suggested that the increased use of citizen data and new uses of official data require broadening the scope of responsible actors.

Team B responded that while NSOs have a responsibility to support ethical standards, this responsibility should be expanded to other stakeholders to build a shared national data ecosystem and enable social ownership of data. They also argued that NSOs must engage with citizens and stakeholders to acquire an additional social license to innovate and serve a rapidly changing world.

Veléz then asked panelists to debate whether data should be shared based on trust, or whether trust is misplaced as data users will prioritize political objectives over data ethics. Noting NSOs cannot control trust, Team A suggested focusing instead on responsible use of data as a principle for NSOs to become trustworthy, bearing in mind that trust is difficult to rebuild once it is lost.

Arguing that trust is hard to define and measure, Team B proposed focusing on validity in data sharing, resulting from clearly defined rationale for data sharing (legitimacy), reputation for data quality (credibility), and timeliness (effectiveness) of data sharing.

On a proposition stating that ownership of data should be governed by existing laws and regulations, Team A stressed that existing laws like the European Union's General Data Protection Regulation recognize data sovereignty, which means individual ownership, noting that ownership does not preclude access. Speakers added that existing laws and regulations are a robust framework that can serve as a starting point to improve accountability and transparency and promote international collaboration.

Team B proposed regulations, norms and standards must evolve to adapt to new practices and challenges, including intersectional approaches such as citizen data, new data uses enabled by AI, and changing values on who should own and have access to data.

In closing, panelists suggested recognizing the 200-year history of NSO values and rules for statistics, emphasizing that the purpose of NSOs is not data production but information generation, building bridges between NSOs and citizens, recognizing collective rights to data sovereignty and safety in addition to individual rights, and investing in the development of an international consensus on data governance and ethics.

Effective Partnerships for Better Data and a More Equitable Data Ecosystem

This session took place on Friday, 15 November. Setting the scene, Johannes Jütting, PARIS21, reminded participants that creating a partnership is not an end, but a means to achieve SDG-related goals. Acknowledging that maintaining partnerships requires time and resources, he said they should be driven by the ability to create higher impact together.

Moderator Camillo Perez, EAFIT University, Medellín, asked panelists how to harness partnerships for better and more equitable data. Samuel Kobina Annim, Ghana Statistical Service, said the type of partnership should be driven by its purpose, cautioning against creating duplications. Eleonora Betancur González, Presidential Agency for International Cooperation of Colombia (APC), said partnerships should bridge all actors involved, noting that partnerships have been valuable to improve accountability and data access for citizens. Claire Melamed, Global Partnership

for Sustainable Development Data (GPSDD), said the power of partnerships comes from people willing to travel to jointly solve problems.

Adrienne Pizatella, Bloomberg Philanthropies, stated her foundation supports government-led partnerships that engage other actors to be inclusive. Haishan Fu, World Bank, highlighted the importance of creating incentives for partners to invest private resources for a public good so it complements other partners. Tim Hemmings, British Embassy in Colombia, said partnerships are important to build trust in actors as well as new tech. Carlos Escapa, Amazon Webservices, said partnerships can build relations between data providers to improve data quality and governance.

On strengthening partnerships, Annim stressed developing partnership skills across silos. Betancour and Melamed noted the power of South-South cooperation, with Melamed highlighting the UN's power to convene governmental and non-governmental actors including donors. Pizatella noted Bloomberg's Data for Health initiative as an example of a partnership driven by a clear purpose, with Fu adding that partnerships should be built on synergies and provide complementary values to be sustainable. Noting the role of Paris 21 and GPSDD in facilitating partnerships, Hemmings recommended starting partnerships in the national context followed by international networking. Escapa recommended tackling AI-related challenges, cautioning that current models are heavily biased towards English-language data.

Asked about new ideas for partnerships, panelists noted, among other suggestions, focusing on bridging silos, finding new allies, partnering for geospatial data, consolidating strategies, strengthening citizen data, and empowering youth.

Closing Ceremony

On Friday afternoon, 15 November, Laura Luis Saza, Diorama, presented an analysis of the UNWDF's impact on the city of Medellin, including economic benefits, carbon emissions, and waste production. Samuel Annim, NSO Ghana and HLG-PCCB Co-Chair, said the 2024 UNWDF's outcome document – the Medellín Framework for Action – will include sections on the Forum's themes. Rubina Adicary, GPSDD, reminded participants of the dire consequences of climate impacts in marginalized communities that are unmapped, calling for data systems that adequately reflect these communities' needs.

Stefan Schweinfest and Peidad Urinola then passed the Forum's Flag to Fahad Al-Dossari, President of Saudi Arabia's General Authority of Statistics and host of the next UNWDF, which is expected to convene in April 2026. Al-Dossari stated that UNWDF 2026 will focus on enhancing the quality and accessibility of statistical data and include training sessions and a contest for small and medium-sized enterprises developing data solutions.

Maria Fernanda Galeano, city of Medellín, closed the 2024 UNWDF at 5:30 pm.