Multi-stakeholder meeting on data interoperability for the SDGs

Sunday, 5 March 2017
Room DC2-2330, 2 United Nations Plaza, New York.

Concept Note

Background

On 17 January 2017 at the first UN World Data Forum (UNWDF) held in Cape Town, South Africa, the Global Partnership for Sustainable Development Data (GPSDD) and the United Nations Statistics Division (UNSD) convened a meeting on the theme of *Advancing SDG Data Interoperability.* The meeting brought together a wide range of stakeholders to introduce their respective programmes of work and to promote the sharing of experiences and knowledge in the area of data interoperability. The meeting concluded with a proposal to establish a mechanism for multi-stakeholder collaboration on *Data Interoperability for the SDGs* with a view to forming a more organised working group and facilitating communication and consultations with the UN Statistical Commission's High-level Group for Partnership, Coordination and Capacity-Building for statistics for the 2030 Agenda for Sustainable Development (HLG-PCCB) and other relevant bodies of the official statistics and geospatial information communitites. The *multi-stakeholder meeting on data interoperability for the SDGs,* scheduled to take place on 5 March 2017 in New York, is the next stage in this process.

What is meant by data interoperability

The effective use of data for sustainable development requires the establishment of reliable and well-documented data services,[1] which are publicly available and accessible to all stakeholders (including citizens, governments, international organisations, businesses and researchers, amongst others).

Data interoperability is the ability to access and process data assets from multiple sources and in multiple formats, and to integrate them into coherent information products or services (e.g. for mapping, visualisation or other forms of analysis). Interoperability enables multiple stakeholders to make best use of existing data standards, classifications and methodologies by bringing together different sources and forms of data, thus enhancing the usability, and

[1] The concept of *data as a service* builds on the principle that data should be accessible to users on demand, in the form of self-contained "services" that bundle it with any metadata and software needed to interpret and use it.

importantly re-usability, of any given dataset by making it part of a more holistic interlinked data ecosystem; ideally in machine-readable formats.

Forms of data interoperability

The concept of 'data interoperability' is not homogenous and can refer to a number of different types, including:

- **Search interoperability**: the ability of two or more data collections to be searched by a single query;
- **Syntactic interoperability**: use of specific data formats and communication protocols to enable the exchange of data across computer systems;
- **Semantic interoperability**: the ability to automatically and unambiguously interpret the meaning of information exchanged;
- **Business process interoperability**: the ability to address the organisational and governance issues of data sharing, including data ownership, information silos, skills and usability.

In addition to these different types of interoperability, various forms also exist, including across: domains, different data providers, past and future versions of the same data service, different levels of detail or aggregation (e.g. from national to international).

Key enablers of data interoperability

- <u>Standards for data interoperability</u>

  Content and technology standards are key enablers of data interoperability. The former provide commonly agreed code lists and definitions which ensure there is no ambiguity regarding the information content of data services provided by different sources. The latter allow system components developed and maintained by multiple parties to be seamlessly integrated and to work together. Such standards enable multi-lateral exchange and use, in contrast to bi-lateral exchange where an agreement needs to be reached between each pair of parties, or the data consumer needs to adapt to the practices and standards of each data provider.

- <u>Geospatial information systems</u>

  By enabling multiple statistical data items to be linked by location and analysed using the description of geographically-related features, geospatial standards and information offer crucial opportunities for implementation and scalability of data analysis tools.

- Application Programming Interfaces (APIs)

  An API is a reusable interface to which different client applications can easily connect. APIs provide essential capabilities for extending and integrating various data and software components by allowing users to link data and processes from multiple sources. APIs also facilitate the discovery of new uses for existing data assets and make them easily available to external third-party developers; thus fuelling new innovations.

- Data governance

  To allow citizens, governments, international organisations, civil society organisations, businesses and other stakeholders to build their own applications on top of interoperable data services, it is necessary to develop and open and service-oriented data architecture that facilitates public access to a wide range of data assets from custodian organisations at different jurisdictional levels. Moreover, data management technologies and practices are rapidly evolving and require an increasingly collaborative approach. In this context, a data governance framework based on effective multi-level organisational arrangements to facilitate coordination, negotiation and collaboration across different territorial levels and stakeholder networks is urgently needed. The primary challenge when it comes to governance is reconciling collective and individual needs and interests to achieve common goals through collective decision-making. Coordination provides the critical link between the 'steering' process of governance and the 'rowing' activities of individual actors that move the community in the necessary direction.[2]

The need for interoperability to achieve and monitor the SDGs

The ambition and scope of data needs to achieve Agenda 2030 have made it more urgent than ever for stakeholders from across different communities to reflect on and evaluate existing data principles and standards, with a view to improve upon their usability and use at international, regional, national and sub-national levels.

There is a need to significantly strengthen the capacity of official statistical institutions to produce and make available the basic data and information that our societies rely on. At the

[2] Rajabifard, A. (2010). *Critical issues in global geographic information management with a detailed focus on Data Integration and Interoperability of Systems and Data*. Scoping Paper for the 2nd Preparatory Meeting of the Proposed UN Committee on Global Geographic Information Management, New York, 10-11 May 2010.

same time, new partnerships between all stakeholder groups need to be fostered to ensure that the digital and data revolutions are taken full advantage of. When it comes to SDG data and indicators, these needs are two sides of the same coin.

The value that interoperability can bring to enhancing usability and use is widely acknowledged and progress on joining-up SDG data is already underway. Ensuring that data is interoperable is the responsibility of all stakeholders that produce data and those who are responsible for setting the standards. Moreover, technical solutions on a global scale that make best use of all available data will only be found if the political will to do so exists. Collaborative leadership is therefore crucial to ensuring that a progressive, common agenda incorporating all stakeholders moves forward at a speed commensurate with both the urgency and aspiration of global ambitions.

Meeting Objectives

Ultimately, the aim of this joint UNSD/GPSDD initiative is to improve the availability, accessibility and use of data for SDG implementation and monitoring through enhanced data interoperability.

To this end, this meeting will pick up where the discussion at the UNWDF in Cape Town left off. The meeting will function as the first of several consultations that will help to identify what a workable mechanism for the coordination and facilitation of various work streams and initiatives to join up data across official and other processes could look like. It will also explore what system(s) need to be put in place to validate data and define appropriate standards.

In order to achieve this, a series of face-to-face meetings and web-based discussions is envisaged.

Areas of work for consideration identified at the UNWDF

Specific initial tasks include:

a. Agreement on the concept note defining the scope and types of data interoperability needs to be addressed by the consultation;
b. Agreement on initial work streams based on the priority areas identified at the meeting of 17 January 2017 at the first UNWDF; that may include, but is not necessarily limited to, exploration of:

i. Integration of geospatial and statistical information;
ii. Adoption of SDMX and other open standards through accessible and user-friendly templates, adaptors and interfaces;
iii. Development and promotion of open APIs for accessing, integrating and analysing data for sustainable development from multiple sources;
iv. The coordination of classification systems to enhance the integration of data from different sources, especially between financial measurement and programmatic data;
v. The improvement of interoperability of microdata from surveys and administrative records; and
vi. SDG data interoperability in the context of the Cape Town Global Action Plan for Sustainable Development Data.

The results of these work streams will be regularly communicated to the HLG-PCCB and Expanded Bureau of the UN Committee of Experts on Global Geospatial Information Management, as appropriate, in the form of specific proposals for their consideration.

Agenda

Time	Activity	Presenter / Facilitator
15:00	Introduction to the event: background, objectives and format	Ms. Claire Melamed Ms. Francesca Perucci
15:10	Brief introduction to the different types of data interoperability and key enablers	Mr. Luis González Morales
15:20	Participant introductions and brief statement from each on what they consider interoperability to mean and how it emerges in their work	Ms. Claire Melamed Ms. Francesca Perucci
15:40	Priority areas identified at the meeting of 17 January 2017 at the first UNWDF	Mr. Tom Orrell
15:50	SDG data interoperability in the context of the Cape Town Global Action Plan for Sustainable Development Data	Ms. Linda Hooper
16:00	Group tasks (3 groups, same questions to all): 1. **Identifying priorities:** Are the priorities identified in Cape Town the right ones to focus on? What specific datasets, classifications and standards most urgently need making interoperable and where are the quick wins? 2. **Longer term objectives:** What are the longer term goals and what (potential) obstacles (technical and political) exist? 3. **Moving forward:** Based on the discussion up to now, what are the next steps for this initiative? a. What priority work streams need to be agreed? b. Who should be involved (or invited to participate)? c. How, and by when, should work be carried out?	Facilitators: Luis González Morales, UNSD Abdulla Gozalov, UNSD Linda Hooper, UNSD Tom Orrell, Publish What You Fund Aditya Agrawal, GPSDD Shaida Badee, Open Data Watch
17:00	Round-up of discussions from break-out sessions (10 mins per group then 15 mins plenary discussion)	Ms. Claire Melamed Ms. Francesca Perucci
17:45	Next steps	Ms. Claire Melamed Ms. Francesca Perucci
18:00	**End**	