Multi-stakeholder meeting on data interoperability for the SDGs
Sunday, 5 March 2017 (afternoon side event)
Organized by UNSD and Global Partnership for Sustainable Development Data (GPSDD)

Introduction

Following an initial meeting convened at the margins of the first UN World Data Forum (UNWDF) in Cape Town on 17 January 2017, this event brought together participants from National Statistical Offices (NSOs), international organizations, private sector, and civil society, representing diverse perspectives, to discuss how to promote the use (and re-use) of available SDG-related data sets and how to make them more widely available and accessible across data ecosystems.

The group explored how technology and standards, data governance mechanisms, geospatial information systems and APIs can enable users to access, process and integrate multiple data assets into coherent information products or services across sources, domains, formats, units of analysis and time periods, through and technology standards.

After an introduction that set out the vision and expected outcomes of the meeting (see accompanying concept note), participants split into three break-out groups to identify priorities, long-term objectives, and next steps to enhance SDG data inter-operability from three different perspectives: syntactic inter-operability, semantic inter-operability, and from a governance perspective with a view to setting out a roadmap for the establishment of a multi-stakeholder collaborative, jointly led by UNSD and GPSSD, to focus on a number of inter-operability challenges at a global level.

Below is a summary of the outcomes from each break-out group:

Group 1: Governance

Finding solutions to inter-operability challenges requires policy coordination, a collaborative approach, and structured management to ensure an enabling environment for technical partnership.

Discussions in this break-out group centred on identifying the pre-requisites for such an enabling environment and the possible elements of a governance structure for a multi-stakeholder collaborative.

The discussion identified the following eight priorities to enable the effective governance of a multi-stakeholder collaborative on data interoperability:

1. **Bridging technical and policy spaces:** Inter-operability challenges are not just technical in nature, but closely related to institutional policy issues (e.g., on data ownership, access and control). Therefore, a collaborative effort for data inter-operability must include and engage both policy-level and technical experts;
2. **Setting out common working principles:** Establishing and agreeing a set of common principles will be crucial to guide the work of organizations and individuals involved to approach interoperability challenges through a common lens;

3. **Mapping available data assets to policy needs of the 2030 Agenda:** To glean a better understanding of policy-level challenges and opportunities within the data inter-operability space, it is necessary to map existing data assets, with a view to identify data gaps and interoperability obstacles across data producers and users;
4. **Linking work-streams to practical needs:** There was broad consensus that work-streams should focus on addressing specific needs for syntactic and semantic data inter-operability, as set out in the accompanying concept note. Crucially, these efforts should be complementary to, not duplicative of, work already being undertaken by the UN-GGIM, IAEG-SDGs and other official processes;
5. **Launching pilot projects to test possible solutions:** There was support for the idea of choosing and undertaking a number of pilot projects to quickly test proposed ideas/solutions to specific data interoperability challenges. Pilot projects could be thematic in nature and/or national in scope;
6. **Developing and agreeing on common vocabulary:** The group noted that terminology used by different stakeholders could sometimes be unclear or ambiguous. Therefore, the production of a glossary of terms and their meaning, building on the background text within the concept note, should be considered;
7. **Using online platforms:** The Virtual UN World Data Forum and the GPSDD website were identified as possible platforms to host information on the proposed multi-stakeholder collaborative and its work; and
8. **Maintaining free and open approach:** All outputs of the various work streams of the group should be openly licensed and free to use.

Group 2: Semantic interoperability

Discussions in this break-out group focused on the integration of geospatial and statistical domains, as well as on the relationship between the DDI (Data Documentation Initiative) standard and the Standard Data and Metadata eXchange (SDMX). Specific points included:

1. **Integrating geospatial and statistical domains:** The work of the UN Initiative on Global Geospatial Information Management (GGIM) and the UN Statistical Commission on the integration of geo-spatial and statistical information should provide the overarching framework for the activities of the group.
2. **Joining of microdata and macrodata documentation standards:** Providing a documented pathway between key microdata standards (such as DDI) and macro-data standards (such as SDMX) would be highly beneficial, particularly identifying disaggregated data sources for SDG target monitoring at national level;
3. **Focusing on the service layer:** The group noted that there will never be a single standard for describing all types of data, nor will there be a universal model for Application Programming Interfaces (APIs). The emergence of 'brokers,' offering translation and transcription services, should therefore be encouraged.
4. **Prioritizing enablers for the above:**
   - An inventory (country-by-country) of relevant datasets – including administrative, statistical, and other types of data – could create an opportunity for mapping the potential links between datasets;

- A mapping of SDG indicators to their current and potential sources at the national and sub-national levels, including their microdata sources.
- Identification and documentation of currently available APIs for dissemination of sustainable development data;
- Adherence to commonly used and established international standards;
- Identification of champions to encourage and support the introduction of open data and metadata standards as appropriate; and
- Development of a strategic plan with financing. Building links between domestic datasets and with international datasets will require an investment in human skills and tools too.

Group 3: Syntactic interoperability

Improving interoperability at the syntactic level requires an inclusive approach to capacity building, tailored to the different skill levels and technical capabilities of very diverse groups of data producers and users.

Participants in this group stressed that new tools should make it easy for people to find, explore and understand the structure and contents of available datasets, so as to promote innovation in their widespread use. In particular, more efforts are needed to facilitate the adoption of open standards at the country level that translate complex proprietary data structures into easily accessible and well-documented APIs that allow the transformation of SDG data into effective applications.

As an example, the group explored the Data4SDGs API Highways project (http://apihighways.data4sdgs.org/), which focuses on promoting the use of already existing datasets by empowering the developer community. Also, the group discussed ongoing efforts to build an open API for the global SDG indicators, aimed to facilitate their integration with additional sources of information, including geo-referenced datasets that are already available at the regional, national and subnational levels.

The group identified five key principles for APIs to be effective enablers of data inter-operability:

- **User focus** – APIs should be developed with the needs of users in mind, especially at the national level;
- **Findability/searchability** – APIs should facilitate the discovery of existing data assets and services;
- **Reliability/longevity** – Once released, APIs should be reliably maintained over time, in order to guarantee long-term support to client applications developed by their users;
- **Adaptability** – APIs should be designed in a way that allows them to evolve over time, incorporating emerging user needs and technologies, while maintaining backward compatibility for client applications; and
- **Low barriers to entry** – APIs should be well documented and allow users without overly sophisticated skills to get started quickly. Further, they should.

Conclusions

During the concluding plenary discussion, participants stressed the need to avoid duplication of efforts, as there are already ongoing initiatives in different domains on the issue of data interoperability (for instance, among the geo-spatial information community), and to focus on already existing standards, rather than trying to build standards from scratch. The plenary discussion also highlighted the crucial role of governance mechanisms and policy-level enragement in inter-operability challenges, as technology and the existence of standards alone do not guarantee the availability inter-operable data. It was further noted that capacity-building efforts are crucial for existing standards and tools for data interoperability to be implemented at the country level.

It was agreed that a multi-stakeholder collaborative should be established, to provide policy-level guidance to a number of technical work-streams tasked with working towards finding solutions to some of the challenges identified.

Next Steps

The following follow-up actions were agreed:
- Reach out to other communities of data users and data producers;
- Prepare a proposal for the multi-stakeholder collaborative, including a proposed governance structure and set of work streams; and
- Establish an online presence to allow interested parties to access information about this initiative and engage with it.

If you are interested in finding out more about this joint collaboration between UNSD and the GPSDD, please contact Luis Gonzalez Morales (gonzalezmorales@un.org) or Tom Orrell (tom.orrell@publishwhatyoufund.org).