

Rome Data Champions meeting summary - January 2024

A meeting of Rome Data Champions, hosted by the Global Partnership for Sustainable Development Data included guest speakers: Tom Orrell, Founder and Managing Director of DataReady, along with Valerie Bizier, Senior Statistician at the Office of the Chief Statistician (OCS) at the Food and Agriculture Organization of the United Nations (FAO) and Piero Conforti, Deputy Director of FAO Statistics Division (ESS).

Data and statistics within FAO

Producing statistics is a core mandate of the FAO, embedded in the first article of its 1945 constitution. Many offices within FAO collect, process, and use data. The FAO's Office of the Chief Statistician (OCS) coordinates and sets standards for the production and dissemination of statistics while the Statistics Division (ESS) implements about 60% of FAO's statistical program subject to these data practices and manages regular updates to FAOSTAT, the agency's primary database. As custodian of statistics related to food and agriculture, FAO works with many types of data though a few data products are of primary importance, specifically data related food security¹ and linked to the <u>UN Sustainable Development Goals</u> (SDGs).

FAO collects data directly from countries through a set of thematic questionnaires usually sent on an annual basis. However, country data published by FAO does not always come directly from what country-level officials report. In some cases, timely data is not available from countries or countries do not respond to these surveys in a timely manner. In these cases, FAO statisticians seek to fulfill the organization's mandate by supplementing country-level data in other ways, including accessing publicly available data on the web or through alternative data sources. In other cases, FAO seeks to minimize the reporting burden on countries by, for example, sourcing data on European Union countries from regional officials instead of asking individual countries to report the same data twice.

FAO works traditionally with official statistics and data. These days, however, much useful information is available through innovative data sources and new, informal, or non-official channels. This data may be more timely than official statistics, which are often reported and processed by FAO annually with a 12- to 18-month delay. The FAO leverages machine learning, web scraping and artificial intelligence, for example, to collect, process and analyze data to supplement official statistics. Such techniques and innovative data sources may

¹ Access to food is a particularly important dimension of food security, as observing that enough food is present in a given country or community does not ensure that people have means to acquire food or sufficient food to eat. The Food Insecurity Experience Scale (FIES) measures access to sufficient food, and establishes the basis on which FAO reports on SDG indicator 2.1.2.

produce more detailed information on country-level food security, crop availability, etc. than what is available at country-level.

Statistics produced by FAO on the basis of non-official data sources are submitted to countries for validation. This is the case, for example, for SDG indicator 2.1.2 (the prevalence of moderate and severe food insecurity), which is collected by FAO through the Gallup World Poll and for SDG indicator 15.4.2 (green mountain cover), which is produced directly by FAO using Earth Observation (EO) data.

This validation process does not apply to indicators derived from national official data that are publicly available or collected by FAO or other international organizations. This includes indicators produced from publicly-available national microdata files resulting from agricultural censuses or surveys or from household surveys, indicators produced from official data and published by other international organizations (e.g. employment in agriculture estimates taken from the International Labour Organization (ILO) official database), and indicators derived from variables already collected by FAO through its questionnaires. Data produced by FAO in the context of early warning systems are currently not subject to country validation as they are usually time-sensitive and/or predictions that will be replaced by official data once they become available.

Geospatial data is opening up new opportunities for more timely statistics and new information related to crop estimates and mapping, etc. Yet new data sources also raise questions around how FAO establishes new, often non-statistical rules and parameters for data estimation, validation, publication, collection, sharing and more. Such questions about rule-setting and data management are typically classified as *data governance*.

FAO has responded to these needs through internal processes, including creating a role for an Executive Data Champion and proposing more integrated internal governance for data and statistics (this document will be presented to the 170th session of the Council). A Data Coordination Group will be created to ensure governance and strategic oversight over data and statistics falling into the mandate of FAO in line with the Data Protection and Intellectual Property Rights Policies. Renewed interdepartmental working groups on data and statistics will continue to handle more detailed statistics and data management issues such as the development and implementation of standardized concepts, classifications, methods, and IT tools at the technical level. In the midst of the new data paradigm and governance, FAO is also updating its Quality Assurance Framework, a set of principles to ensure the quality of data within FAO. These adjustments are in line with the recommendations of the *Data Strategy of the [UN] Secretary-General for Action by Everyone, Everywhere (2020-22)²* described in detail in the following section.

² https://www.un.org/en/content/datastrategy/images/pdf/UN_SG_Data-Strategy.pdf

FAO within the larger context of UN systems and issues of data governance

The UN Secretary-General's <u>Data Strategy</u> provides a primary incentive for the FAO to become a data-driven organization. The strategy directs agencies within the UN to improve their data skills and methodologies and to increase the use of data for evidence-based decision making.

The FAO finds itself on a similar path as other UN agencies seeking to adapt to the rapid increase in digitalization and expansion of digital technologies while also grappling with existing mandates and constitutions to govern the interactions between the UN agencies and member states. Such agreements, crafted in the 1940s and 50s, could not foresee the opportunities and challenges that would arise with rapid technological change in the digital age.

The World Health Organization (WHO), for example, set out to establish data governance principles in 2020, faced with highly-politicized health data amidst the COVID-19 pandemic. The <u>principles</u> outline the values that guide how data is collected, processed, shared, and used by WHO. Such principles serve as a starting point for follow-on data governance and management conversations and decisions as the work on data governance in WHO is ongoing.

Similarly, UNICEF recently transitioned the role of the Chief Statistician to the Chief Data Officer to align with both the larger UN system strategy and with the organization's <u>principles</u> on the responsible use of data for children. Moving forward, UNICEF is starting to establish networks of data stewards across its country and regional offices to uphold accountability standards to promote transparency around data use.

FAO shares data through agreements with other organizations, including other Rome-based agencies. Data from the International Fund for Agricultural Development (IFAD), for example, is currently disseminated through FAO's Food and Agriculture microdata catalog, and FAO can share operational-level data with the World Food Program, which is outside the scope of FAO data and statistics. FAO is seeking to improve the interoperability of data and data tools among UN agencies through the adoption of a one-stop-shop statistical data warehouse for all FAO data—an objective that was part of a proposal to improve FAO data and statistical governance submitted to the Council at its 168th and 170th sessions.

FAO's current data licensing policy does not allow commercial use of its data. This constitutes, on one hand, a limitation of FAO's compliance to open data principles and, on the other hand, a restriction that is difficult to implement, since FAO has no means to limit or track access to its data on its website. This policy is currently being reassessed and possibly revised in the context of the development of the new FAO policy on intellectual property rights and the review of FAO's data license.

Data governance best practices

Tackling the challenges of governing data within FAO must start by acknowledging that data are political. Data governance issues are by nature both political and contentious, not primarily technical—and that is as it should be, Mr. Orrell explained. Once this is acknowledged, tackling data governance challenges becomes easier as they are demystified and moved from the technical to the policy realm.

Building on this understanding, it's easy to see that all good data governance is built on compromise, trust, and mutual accountability. Social and diplomatic factors impact conversations around data governance much more than technical constraints or standards.

To distinguish between the political and technical aspects of data, it's helpful to think of data governance as the practice of exercising control over data while data management is the technical policies and decisions that are based on data governance agreements. Decisions of FAO's Council, Conference, and committees as well as the decisions and policies that guide the role of FAO's Chief Statistician are data governance tools distinct from the practice of implementing these rules, otherwise known as data management. Such a distinction separates the diplomatic and political elements of setting the rules from the day-to-day practice of implementation, ensuring that data is used consistently and appropriately.

In establishing practices and resolving questions related to data governance within FAO, it's important to consider—not only FAO's internal processes—but also the role that member states should play within FAO to ensure mutual accountability. Once trust is established, resolving questions around data use, validation, and more becomes a routine process based on mutually agreed upon data governance guidelines. How to strengthen the interlinkages between data governance processes for member states and those within FAO is a key question facing the organization and its governing bodies.

Mr. Orrell advised against getting "stuck in the weeds," i.e. focusing on specific issues and losing track of bigger issues on these topics. It's important to first cooperatively establish parameters of data governance before tackling technical questions around data use, privacy, and more. Questions, for example, around what data should be anonymized in geospatial imaging are related to data management. Once the mechanisms for establishing data governance rules are in place and agreed upon, these questions become relatively easy to address.

A human rights-based approach to data governance at FAO

In response to a participant question on this point, Mr. Orrell stated that applying a human rights-focused lens is integral to determining data governance practices within FAO, whose mandate includes considerable data collecting powers to benefit all people. At the same time, the organization is constrained in that it must use its powers in the best interest of member

states; this can sometimes lead to tensions in the use of 'official' versus 'non-official' data. Compromises must be found that accommodate both FAO's duties to the public and to Member States. Transparency is key in this respect. It is central to FAO's role as a custodian of data to provide transparency about what datasets are being collected and processed, even if the content of those datasets is not made publicly available. Such transparency is important to build trust among FAO and both Member States and the general public.

Two further pertinent concerns around data governance from a human rights perspective include: (1) ensuring that people whose data is being collected reap benefits from the insights that are generated by such data, and (2) ensuring that vulnerable communities are not harmed by data collection or use (e.g. as when geospatial data reveals information about remote tribes who may have chosen not to have contact with outsiders). For normative data governance frameworks at the level of the FAO and Member States, establishing standards must take into consideration the harm that may be caused to individuals and communities by data which are under the control of FAO.

It is important to note that FAO is not an exception in not having well-established and mature data governance processes for non-traditional data collection, processing, and use. The challenges that FAO faces are universal, with many UN agencies, governments, INGOs and private companies grappling with similar tensions and challenges. Moving forward, there is an opportunity for FAO to both learn and share lessons with other UN agencies and further deepen and strengthen its data governance and management practices.