

The organisation

The Food and Agriculture Organization of the UN (FAO) [<http://www.fao.org/home/en/>] is an international organization dedicated to addressing the problem of hunger in the world. It is part of the "UN family" of agencies and is membership based. FAO is present in 130 countries worldwide, its HQ being in Rome, Italy. One of the activities of FAO is the collection of statistics in all areas of interest to the organization, such as agriculture, forestry, fisheries, climate change and food security. The statistics collected by FAO are typically provided by National Statistical Offices, then compiled/analyzed by FAO and disseminated through a few platforms, such as FAOSTAT [<http://www.fao.org/faostat/en/#home>]. The number and type of users of those statistics is very wide, including governments, NGOs, researchers, journalists and private citizens.

What was the challenge that you addressed?

Statistical classifications are central to all phases of statistical data collection and analysis, as they define the "meaning" of the data and, as a consequence, the possibility of reusing the data in different information systems and for different purposes. However, while statistical data are receiving great attention in finding new ways for data collection, analysis, visualization and to improve their interoperability, relatively less attention is being devoted to statistical classifications. For example, no standardized ways exist to express their structure and their content in formats that are readily machine actionable, and the common approaches to dissemination still largely ignore the potential of global unique identifiers and machine readable metadata.

Within Caliper we are applying state-of-art best practices to improve the management and use of international statistical classifications. Some of the challenges we face in our project are the move from "local" classifications based on codes to "global" classifications based on global identifiers, the definition of uniform modelling approaches to express statistical classifications in RDF, and the consistent rendering of their metadata in machine readable formats. Ultimately, such a project also touches on the need to look at statistical classifications as a common good that all actors producing statistical data should be able to smoothly use and reuse.

The work done on the project so far is available on the Caliper website : [<http://stats-class.fao.uniroma2.it/caliper/>]

Which RDA recommendation or output did you adopt & why?

We largely implemented the Recommendations from the Agrisemantics WG [<https://rd-alliance.org/group/agrisemantics-wg/outcomes/39-hints-facilitate-use-semantics-data-agriculture-and-nutrition>], especially those addressing semantic experts (those converting existing classifications into semantic resources) and IT developers (using classifications to produce and manipulate data). The Agrisemantics recommendations were important to us because they originated from a community of experts that were closely aligned with the area of interest of this project.

Caliper

Caliper - Statistical classifications
in an open linked world



Also, the fact that I was directly involved in the Agrisemantics WG as co-chair ensured a good two-ways communication between the expertise available in RDA and our needs as an adoption use case.

While doing so, we kept a vigilant eye on a number of other activities ongoing in RDA, primarily those on metadata, persistent identifiers and versioning.

The adoption process

We had the opportunity to work within a 4-year grant from the Bill and Melinda Gates Foundation, which featured a component aimed at improving the interoperability of agricultural data, especially improving the way common vocabularies (aka classifications) are used in the international context. In fact, our first area of activity was to focus on some of the classifications maintained by OECD for the reporting of development aid.

Within those 4 years, we had a chance to experiment with different (types of) classifications, to prove the scalability of the approach followed. Now in its final phase, we are working to make the project become an official service maintained by the Organization.

The project involved a small team in FAO (up to 3 people mostly working part-time), plus the technical and scientific support by the ART team at the University of Tor Vergata (Rome).

Benefits of adoption and impact

Caliper is bringing forward the idea that statistical classifications are a public good that should be made available in formats immediately consumable by machines and humans alike. The desired long lasting impact of Caliper is that organizations maintaining statistical classifications will adopt open standards for their publication and use, and that they may eventually form a network of classifications users and providers, making their products available to both individual and institutional users.

What lessons did you learn from the adoption process?

The general idea of a project like Caliper has been taken with favour by internal and external users, especially by all those working day-to-day with statistical data. However, since our effort on expanding the contents of the platform is ongoing, it is crucial that informed users keep providing their feedback on the various aspects involved - RDF modelling and conversion, publication APIs, editing platform to mention only a few. To this end, we issued a survey to gather user feedback in October 2020, the results of which are available to view on the Caliper website [https://stats-class.fao.uniroma2.it/caliper/sites/default/files/inline-files/2020_11_Caliper_Survey_summary.pdf]