

December 2023

منظمة
الأغذية والزراعة
للأمم المتحدة联合国
粮食及
农业组织Food and Agriculture
Organization of the
United NationsOrganisation des
Nations Unies pour
l'alimentation et
l'agricultureПродовольственная и
сельскохозяйственная
организация
Объединенных НацийOrganización de las
Naciones Unidas para la
Agricultura y la
Alimentación

AFRICAN COMMISSION ON AGRICULTURAL STATISTICS

Twenty-eighth Session

Johannesburg, South Africa: 4 – 8 December 2023

AGENDA ITEM 9

The FAO Food and Agriculture Microdata Catalogue: Unlocking the power of Microdata

NAME OF AUTHORS

Clara Aida Khalil; Oluwakayode Anidi

ORGANIZATION

FAO Office of the Chief Statistician

SUMMARY

Launched in July 2019 with the aim of becoming a one-stop-shop for agriculture-related microdata and metadata, the [Food and Agriculture Microdata \(FAM\) catalogue](#) provides an inventory of datasets collected through farm and household surveys and censuses containing information related to FAO's mandate (e.g., agriculture, food security, nutrition, emergencies, etc.). After three years from its creation, the catalogue is now populated with over 1300 studies. Majority are datasets collected by the FAO (directly or indirectly by providing technical and financial support to other institutions) or by external organizations granting FAO the license to redistribute them. After highlighting the various benefits of microdata dissemination and the substantial progresses made in terms of FAM maintenance and enrichment since the 27th session of AFCAS, the paper presents the statistical workflow and standards adopted by the FAO for microdata dissemination, and outlines ways in which AFCAS Members can benefit from this initiative.

1. Introduction

The FAO and Member countries increasingly rely on microdata for monitoring and evaluation, designing programs and interventions, and assessing important trends. Furthermore, the adoption of the 2030 Agenda for Sustainable Development has resulted in a significant increase in the demand of highly granular statistics and microdata, which are needed to better monitor phenomena such as food insecurity (Target 2.1), small-scale food producers' productivity and incomes (Target 2.3), agricultural sustainability (Target 2.4), women's rights on agricultural land (Target 5.a) and many others. In this regard, the FAO devotes significant efforts to promote the collection of household and farm surveys microdata in Member countries.

Nevertheless, implementing surveys and censuses and collecting microdata are very expensive endeavours. Hence, ensuring maximum return from these investments by promoting data use is a key responsibility of data producers and sponsor organizations.

The benefits of microdata dissemination are manifold. To mention some, the dissemination of microdata allows increasing the return of the investment in data collection activities and enhance the transparency of the entire statistical production process. Furthermore, when well documented microdata is made available to the public in line with open data values, results and research are made reproducible by a larger audience fostering the credibility of producing institutions. Finally, dissemination opens the way to new uses of information and can serve as a leverage to create statistical innovations.

Despite the importance of microdata dissemination, until 2019 the FAO lacked the policies and the IT infrastructure for promoting the corporate dissemination of microdata and metadata. To fill this gap, the Food and Agriculture Microdata (FAM) catalogue was launched in July 2019 at <https://www.fao.org/food-agriculture-microdata/en/> with the financial support of the Bill and Melinda Gates Foundation, and the United States Agency for International Development.

At launch, the FAM catalogue disseminated 331 datasets on the Food Insecurity Experience Scale, as well as few studies related to food and agriculture. Now, after three years from its development, the FAM hosts more than 1300 surveys and censuses from 189 countries around the world and is continuously updated as new datasets from FAO and member countries become available.

From July 2019 to June 2023, Google Analytics statistics indicate that about 68,000 unique users accessed the FAO microdata catalogue. Furthermore, at the time of writing, there are 2824 registered users which created an account to access licensed datasets. Cross-referencing these users with licensed requests indicates that students and researchers from the academia account for the majority (69%) of users of FAO's microdata catalogue, followed by users from public administrations (6%), UN Agencies (6%), International Organizations (5%), as well as other lesser user groups. All this is evidence of the fact that FAM is gradually becoming the reference global hub for accessing microdata and metadata related to food, agriculture and other themes relevant to the FAO, with great usage.

In addition, by tracking back on the usage of requested licensed data files, we have been able to discover the outputs of the data use, by the different user groups. For example, a professor in Canada used the dataset in FAM to produce a [scientific report](#) on how food insecurity relates to poor mental health and reduced wellbeing of individuals (Elgar et. al., 2020). Many international organizations like the World Bank (WB) and World Food Program (WFP) have also produced various outputs, based on their use of microdata available on FAM. One of such use, resulted in a [quantitative measure report](#) for examining the relationship between food security and different dimensions of gender empowerment (WFP, 2020). A list of selected publications based on microdata available in FAM, is presented in Annex 1.

2. FAM Standards and workflow

The FAM catalogue follows well-known international standards and best practices for microdata anonymization, documentation and dissemination. In particular, the internal FAM workflow is regulated by a [Quality Standard on microdata dissemination](#) – which was endorsed by the FAO Inter-Departmental Working Group on Statistics (IDWGS) in January 2021 – including protocols for both the implementation of statistical disclosure control techniques and the curation and documentation of microdata. In addition, in order to tackle the legal aspects of data dissemination and re-distribution, the Office of the Chief Statistician (OCS) liaised with FAO's legal department to redact a template of license to redistribute datasets produced by other Organizations. In the following paragraphs, the main technical features of the FAM catalogue and its workflow are illustrated. Similar information was provided at the 27th AFCAS session.

2.1 The FAM catalogue IT platform

The FAM catalogue leverages on the [National Data Archiving \(NADA\)](#) platform, an open-source web-based cataloguing application developed and maintained by the World Bank. NADA is used by numerous international organizations and national statistical offices (NSOs) to create portals allowing users to browse, search, compare, apply for access, and download relevant data. Thanks to the adopted

technology, studies in the FAM catalogue can be filtered by multiple dimensions (year, country, collection, license), which enhances the findability of disseminated microdata and metadata.

The NADA IT infrastructure was selected because, as of now, it is the most widely used cataloguing system worldwide by many countries and international organization, which makes the data exchange between different platforms relatively easy to implement. In addition, the NADA offers several back-end features providing IT security protection and rendering the platform safe against possible crashes.

Figure 1: Overview of the NADA platform of the FAM catalogue

The screenshot displays the 'Food and Agriculture Microdata Catalogue' interface. At the top, there is a search bar with the placeholder text 'Keywords...' and a 'Search' button. Below the search bar is a 'Browse Citations' button. The main content area shows a list of survey results. The first result is 'Comprehensive Food Security and Vulnerability Analysis and Nutrition Survey 2012' for Rwanda, 2012, with details including the National Institute of Statistics of Rwanda, collection name 'Food Security', ID 'RWA_2012_CFSVANS_v01_EN_M_v01_A_OCS', last modified date 'Sep 02, 2020', and views '531461'. A button indicates 'Data available from external repository'. The second result is 'Annual Agricultural Area and Production Survey 2010' for Ecuador, 2010, with details including the National Institute of Statistics and Censuses (INEC), Agricultural Statistics Unit, collection name 'Agriculture Census and Surveys', ID 'ECU_2010_ESPAC_v01_EN_M_v01_A_OCS', last modified date 'Feb 18, 2020', and views '516885'. A button also indicates 'Data available from external repository'. The third result is 'Impact, Maintenance, and Sustainability of Irrigation Impact Evaluation Survey 2015-2018, Baseline, First, Second, Third Follow-up'. On the left side, there are filters for 'Years' (From and To dropdowns), 'Countries' (a list of countries with checkboxes and counts, and a search filter), 'Collections' (a dropdown menu), and 'License' (a dropdown menu). At the top right, there are view options for 'Study view' and 'Variable view', and a 'Popularity' dropdown menu. The page shows 'Showing 1-15 of 1,314' results.

2.2. Adopted metadata standards

The NADA platform supports the use of the [Data Documentation Initiative \(DDI\)](#) metadata standard, which is an XML-based international standard designed to document socio-economic surveys, censuses and other microdata collection activities. The DDI is developed and maintained by a consortium of research institutions, universities, and private foundations, and the International Household Survey Network (IHSN) is an active partner in its development.

Though the list is much longer, the FAM catalogue requires clear and accurate information on the following minimum set of DDI core elements listed below:

1. Title
2. Abstract
3. Kind of data
4. Unit of analysis
5. Description of Scope
6. Country
7. Geographic coverage
8. Sampling procedure
9. Weighting
10. Dates of data collection

11. Primary Investigator
12. Confidentiality
13. Access conditions

NADA further uses the Resource Description Framework (RDF) standard for cataloguing the physical datasets, and related materials (e.g., reports, questionnaires, methodological documents, etc.).

2.3 Statistical disclosure control for microdata dissemination

Most data collected by NSOs and International Organizations cannot be published directly due to confidentiality and privacy concerns. Statistical disclosure control (SDC) techniques are a set of procedures to treat microdata files so that these can be published without disclosing the confidential information it contains, while limiting the information loss due to the anonymization process.

The FAO is responsible for ensuring the quality and the proper anonymization of disseminated data. For this reason, OCS has refined a well structure SDC protocol which serves as a guide for the anonymization of all microdata files disseminated through FAM. When transmitting microdata to OCS, data providers are required to remove all potential direct identifiers, which are all variables that unambiguously reveal the identity of a respondent or a statistical unit. Besides direct identifiers, data owners can also decide to remove extremely sensitive variables, which are those that – if disclosed – would cause harm to the data subject. Then OCS, in collaboration with technical people from the unit/institution providing the data, identifies a list of key variables and defines potential disclosure scenarios. Based on these inputs, the disclosure risk is measured and one or multiple SDC techniques – extensively discussed in the FAO's SDC Protocol – are implemented. Subsequently, the outcome of the anonymization process is reviewed with the data provider in order to ensure that adequate protection is achieved without overly reducing the analytical value of the data. Finally, the anonymized files are submitted to the Chief Statistician for approval and consequent publication on FAM.

2.4 Type of datasets and access modalities

The type of studies disseminated by the OCS on FAM can be classified into three broad categories:

- 1) *Studies submitted by FAO internal data providers through a data deposit system:*
 - Impact evaluation surveys from the Inclusive Rural Transformation and Gender Equity Division (ESP), Food Insecurity Experience Scale surveys from the Statistics Division (ESS), Resilience Index Measurement Analysis (RIMA) data from the Agrifoods Economics Division (ESA), the Data in Emergency Monitoring (DIEM) surveys from the Office of Emergencies and Resilience (OER), and the Global Strategy to Improve Agricultural and Rural Statistics (GSARS) pilot surveys.
 - FAO/WHO Global Individual Food Consumption (GIFT) data from the Food and Nutrition Division (ESN), National Forestry Inventories from the Forestry Division (NFO), and the Pastoralist data from the Animal Production and Health Division (NSA).
- 2) *Studies submitted by external data providers via email:*
 - Data from international organizations, such as the International Fund for Agricultural Development (IFAD), and the private sector (e.g. Syngenta) through different collaborations and calls for engagement.
- 3) *Metadata harvested from external and compatible platforms:*
 - Surveys from NSOs and the World Bank microdata catalogue.

The microdata of studies in the first two categories listed above are disseminated as licensed files in most cases. This means that users will have to fill out a short application describing the intended use of dataset prior to being granted access. OCS evaluates these requests based on two main criteria:

1. the dataset is fit for the purpose stated in the application.
2. the user comes from a credible institution which is unlikely to violate any terms of use.

Only 5 FAO studies on FAM have microdata disseminated as public use files under the collection of nutrition-related datasets. In both cases, requesting users need to agree to a disclaimer, in which they subscribe to follow general terms of use, and acknowledge that the original collector of the data, the Microdata Library and the relevant funding agency bear no responsibility for any use of the data or for

interpretations or inferences based upon such uses. The general agreement and conditions to access and use datasets in the Food and Agriculture Microdata (FAM) catalogue can be found in the [Terms of Use for Datasets](#).

3. **How countries can use the FAM**

The channels through which Member countries can benefit from the FAM catalogue are substantially two. First, the catalogue can be used to enhance the visibility of implemented surveys and censuses by disseminating or redistributing the related microdata and metadata through it. Second, Members can access the FAM to download and explore the data and documentation of data collection activities implemented by other countries and institutions.

3.1 **Publishing microdata and metadata on FAM**

Many of the data platforms and statistical databases published by the FAO are well-known to the international community and receive millions of hits per year. For example, Google analytics indicate 68,000 users and more than 383,000 page views in the first three years of the FAM. Being an aggregator of datasets all over the world, the FAM catalogue is likely to attract a much broader audience than any single national platform. Thus, countries can benefit from this by making their surveys and censuses discoverable on FAM, via two main approaches: sharing only the metadata of the study; or sharing both the microdata and the metadata.

Case 1: Sharing metadata only

It should be noted that many – if not most – AFCAS members are already publishing their microdata files, in some cases even using NADA¹. Under this circumstance, it is possible to re-disseminate the metadata from countries through FAM.

The workflow adopted to redistribute countries' metadata is the following:

- **Step 1:** A collaboration or communication is established between FAO and the data owner.
- **Step 2:** The FAM team in OCS harvests the DDI metadata from the national platform. In cases where the original metadata is not in English, the FAO translated the metadata document.
- **Step 3:** The metadata is reviewed and published on FAM, along with the data dictionary and relevant documents.
- **Step 4:** A link to the microdata on the national platform is provided, for users to access the data (**Figure 2**).

¹ Without the intention of being exhaustive, examples of AFCAS Members using a NADA catalogue to disseminate microdata and metadata are [Burkina Faso](#), [Cameroon](#), [Ghana](#), [Mali](#), [Mozambique](#), [Niger](#), [Nigeria](#), [Rwanda](#), [Senegal](#), [Tanzania](#), [Uganda](#).

Figure 2: Redirecting users to external/national data platform using a link.

The screenshot displays a web page for the 'Annual Agricultural Survey 2021-2022' in Senegal. The page includes a header with the breadcrumb 'Home / Food and Agriculture Microdata Catalogue / AGRICULTURE-CENSUS-SURVEYS / SEN_2021-2022_AAS_V01_EN_M_V01_A_OCS'. A main image shows people in a field. The title 'Annual Agricultural Survey 2021-2022' is prominent, with a 'GET MICRODATA' button. Below the title, there are fields for Reference ID (SEN_2021-2022_AAS_v01_EN_M_v01_A_OCS), Producer(s) (Directorate of Analysis, Forecasting and Agricultural Statistics), Collections (Agriculture Census and Surveys), and Metadata (Documentation in PDF, DDI/XML, JSON). On the right, statistics show 'CREATED ON Apr 13, 2023', 'LAST MODIFIED Apr 14, 2023', 'PAGE VIEWS 11456', and 'DOWNLOADS 114'. A navigation bar at the bottom contains 'STUDY DESCRIPTION', 'DATA DICTIONARY', 'DOWNLOADS', and 'GET MICRODATA'. A section titled 'Data available from external repository' states that data is available from an external website and provides the URL: <http://anads.ansd.sn/index.php/auth/login/?destination=catalog/272/get-microdata>. A 'Back to Catalog' button is located at the bottom left.

At the time of writing, the FAM catalogue redistributes almost 370 studies for which the microdata is accessible on the microdata catalogue of the producing institution or the World Bank.


Case 2: Sharing both Microdata and Metadata

The approach of disseminating both the microdata and the metadata, is a more suitable option when national institutions do not have a microdata catalogue in place. The workflow adopted in these cases is summarized below:

- **Step 1:** A legal agreement is established between the FAO and the external provider/national institution, using a license to redistribute form, which defines the terms of agreement between the external data provider and the FAO.
- **Step 2:** OCS supports the country in preparing the dataset to meet well established quality standards and perform suitable SDC procedures in order to protect/anonymize the microdata.
- **Step 3:** OCS also review the relevant documentation (e.g. methodological notes, reports, questionnaires, etc.) in order to draft and develop a DDI metadata document.
- **Step 4:** The DDI metadata document and anonymized microdata files are shared with the national institution for final review and validation.
- **Step 5:** The study is published on FAM as licensed file, along with the relevant documentation. Removal of the study from the FAM catalogue can be requested by the data owner at any time after publication.
- **Step 6:** Users submit a request to access the dataset by filling an application for access to a licensed dataset (**Figure 3**), which is then reviewed by OCS.

Figure 3: Applying to licensed datasets on FAM.

Home / Food and Agriculture Microdata Catalogue / FORESTRY / UGA_2016-2019_NBS_V01_EN_M_V01_A_OCS



National Biomass Study, 2016-2019.

Uganda, 2016 - 2019 [GET MICRODATA](#)

Reference ID: UGA_2016-2019_NBS_v01_EN_M_v01_A_OCS

Producer(s): National Forest Authority, Ministry of Water and Environment, District Forest Services

Collections: [Forest Inventory Data](#)

Metadata: [Documentation in PDF](#) [DDI/XML](#) [JSON](#)

CREATED ON
Nov 16, 2021

LAST MODIFIED
Nov 16, 2021

PAGE VIEWS
3955

DOWNLOADS
345

STUDY DESCRIPTION DATA DICTIONARY DOWNLOADS **GET MICRODATA**

[View all requests](#)

Application for Access to a Licensed Dataset

Fields marked with * are mandatory.

The information provided on this page will be kept confidential and will be used for internal purposes only.

First Name	kayode.anidi
Last Name	
Organization	

3.2 Accessing microdata from other countries and institutions

Countries can also access FAM to explore microdata and metadata of data collections implemented by other countries or organizations, which may offer several benefits. First, when designing a survey and developing a questionnaire it could be useful to look into questionnaires and sampling designs adopted by other organizations to collect similar data to get insights on how it was done. In addition, accessing survey reports and microdata could help assessing how a given approach has worked in the past, and identify areas for potential improvement of the questionnaire. Second, by better understanding the data collection practices in similar or neighbouring countries, opportunities to harmonize data collection instruments and build south-south collaborations may be identified. Lastly, countries may use microdata to perform comparative analysis on similar contexts to inform the design of policy in absence of national microdata.

4. Conclusions and next steps

Since its launch in 2019, the FAM catalogue has become a reference point for the dissemination of microdata related to food and agriculture. The Office of Chief Statistician has put in place different standards and protocols to guide the curation, anonymization and dissemination of microdata on the FAM Catalogue to make sure that microdata and metadata are disseminated in line with international standards and best practices. Over the years, the FAM Catalogue has increased in content and gained more visibility. This has great potential for Member countries that can use FAM to increase the accessibility and findability of their microdata and metadata.

AFCAS Members are encouraged to take advantage and benefit from this activity, by disseminating or re-distributing their data in FAM, as well as exploring the data catalogue. Organizations that would like to contribute to FAM, can send an email to FAM-Catalogue@fao.org. OCS will also be in touch with countries that are currently disseminating microdata and in which a collaboration has not yet been established.

5. Questions and Invitations to AFCAS members

AFCAS members are requested to express their views and recommendations to FAO on the following:

- The primary legal, political, and technical (e.g., IT infrastructure, anonymization, documentation) constraints for microdata dissemination, and the role that the FAO can play in helping to address them.
- Their perceived utility and relevance of the FAM catalogue, and their willingness to disseminate or redistribute their microdata and metadata through the platform with the support of FAO.

6. **Annex: List of selected publications based on microdata available through the FAM Catalogue**

- Barlow, P., Loopstra, R., Tarasuk, V., & Reeves, A. (2020). Liberal trade policy and food insecurity across the income distribution: An observational analysis in 132 countries, 2014–17. *The Lancet Global Health*, 8(8). doi:10.1016/s2214-109x(20)30263-1 [https://www.thelancet.com/journals/langlo/article/PIIS2214-109X\(20\)30263-1/fulltext](https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(20)30263-1/fulltext)
- Broussard, N. H. (2019). What explains gender differences in food insecurity? *Food Policy*, 83, 180-194. doi:10.1016/j.foodpol.2019.01.003 <https://www.sciencedirect.com/science/article/abs/pii/S0306919218300824>
- Dudek, H., & Myszkowska-Ryciak, J. (2020). The prevalence and socio-demographic correlates of food insecurity in Poland. *International Journal of Environmental Research and Public Health*, 17(17), 6221. doi:10.3390/ijerph17176221 <https://pubmed.ncbi.nlm.nih.gov/32867149/>
- Elgar et. al. (2020). Relative food insecurity, mental health, and wellbeing in 160 countries. *Social Science & Medicine* 268 (2021) 113556. <https://pubmed.ncbi.nlm.nih.gov/33293171/>
- Romaello et. al. (2022). The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels. *Lancet*. 2022 Nov 5;400(10363):1619-1654. doi: 10.1016/S0140-6736(22)01540-9. <https://pubmed.ncbi.nlm.nih.gov/36306815/>
- FAO (2021). Guidelines on data disaggregation for SDG Indicators using survey data. <http://www.fao.org/3/cb3253en/CB3253EN.pdf>
- FAO (2022a). An indirect estimation approach for disaggregating SDG indicators using survey data. Case study based on SDG Indicator 2.1.2. Rome. FAO. <https://www.fao.org/documents/card/en/c/cb8670en>
- FAO (2022b). Using small area estimation for data disaggregation of SDG Indicators. Case study based on SDG Indicator 5.a.1. Rome. FAO. <https://www.fao.org/documents/card/en/c/cb8998en>
- FAO (2023). Integrating surveys with geospatial data through small area estimation to disaggregate SDG Indicators at subnational level. Case study on SDG Indicators 2.3.1 and 2.3.2. Rome. FAO. <https://www.fao.org/documents/card/en/c/cc3944en>
- FAO, IFAD, UNICEF, WFP and WHO. 2019. The State of Food Security and Nutrition in the World 2019. Safeguarding against economic slowdowns and downturns. Rome, FAO <http://www.fao.org/3/ca5162en/ca5162en.pdf>
- FAO, IFAD, UNICEF, WFP and WHO. 2020. The State of Food Security and Nutrition in the World 2020. Transforming food systems for affordable healthy diets. Rome, FAO <http://www.fao.org/documents/card/en/c/ca9692en>
- Sadiddin, A., Cattaneo, A., Cirillo, M., & Miller, M. (2019). Food insecurity as a determinant of international migration: Evidence from Sub-Saharan Africa. *Food Security*. doi:10.1007/s12571-019-00927-w <https://link.springer.com/article/10.1007/s12571-019-00927-w>
- Sinclair, K., Ahmadigheidari, D., Dallmann, D., Miller, M., & Melgar-Quiñonez, H. (2019). Rural women: Most likely to experience food insecurity and poor health in low- and middle-income countries. *Global Food Security*, 23, 104-115. doi:10.1016/j.gfs.2019.04.006 <https://www.sciencedirect.com/science/article/abs/pii/S2211912418301093>
- WFP (2020). The power of gender equality for food security: Closing another gender data gap with a new quantitative measure. <https://www.wfp.org/publications/power-gender-equality-food-security>