



## RAM GUIDELINES

# Using Data Library for Household Survey Data



## WHY USE DATA LIBRARY?

RAM's mission is to provide credible, relevant and timely evidence for strategic and operational decision-making, resource mobilization and accountability to beneficiaries.

[Data Library](#) is WFP's secure space for indexing, versioning and storing data. This service is easy to use and is backed by WFP enterprise security, management tools, policies and procedures.

Using WFP's Data Library serves three purposes:

1. it helps **standardize the formatting of Country Office (CO) household survey datasets** and metadata;
2. it functions as a **secure, persistent storage** for all CO household survey data and as a repository of survey reports;
3. while the **data ownership stays with the country office**, Data Library enables **sharing of specific and relevant datasets and reports with wider audiences** such as Regional Bureaus and Headquarters for global analyses and even with partners and donors for knowledge sharing and accountability.

Using Data Library and [DataViz](#) to store and share household survey data is an important steppingstone to strengthen WFP's evidence base on food security, economics and markets, climate and earth observation as well as on the effectiveness and results of our programmes.

**The ambition of RAM is to make household food security and essential needs assessment and monitoring survey data available as a public good.<sup>1</sup>** This is part of RAM's open data framework and it is in line with global practice in the

development community, such as the [World Bank Open Data](#), the [Demographic and Health Survey \(DHS\) Program](#) and the [UNHCR Microdata Library](#).

Storing and sharing data centrally will improve **RAM's effectiveness in providing evidence on food security and its root causes to WFP and its partners, enable global analyses, advance information management practices and increase transparency and accountability** to stakeholders, donors and beneficiaries. As a UN organization primarily funded through public investment, it is in fact within our responsibility to uphold the highest data management and data governance standards and minimize operational and financial risks.<sup>2</sup>

**The following standard operating procedures (SOPs) provide step-by-step instructions for RAM (VAM/M&E) Officers in the field on what datasets should be stored and shared in Data Library and when and how to do so.**

## WHAT DATASETS SHOULD BE STORED IN DATA LIBRARY?

Country Offices must store in Data Library **all datasets resulting from assessment and monitoring surveys**, including:

- Surveys led by WFP only;
- Joint surveys with partners, including food security monitoring systems, depending on context;
- Surveys conducted on behalf of WFP through third parties and outsourced to external companies.

**Data must be packaged with associated information**, such as the syntax (i.e. SPSS syntax, STATA do-files or R scripts) used for the analysis, the sampling framework, the questionnaires, and the final report.

## WHEN TO SHARE DATASETS IN DATA LIBRARY?

**Metadata should be created as soon as the assessment or monitoring survey is planned.** This will strengthen internal information management practices and provide visibility at regional and global level on the planned and ongoing assessment and monitoring surveys.

**The work should be flagged as *In Progress***, adding the Phase of the assessment as *Planned*. Any document uploaded before the analysis is completed should be set as *Private*, therefore limiting to the analyst working on the survey project. The Phase field can be updated at anytime during the assessment lifecycle.

**As soon as the analysis is completed and the full final report is cleared,<sup>3</sup>** the Country Office must upload the full survey data package in Data Library.

**The Country Office retain full ownership on their survey data** and they can decide, for each of the resources uploaded, who can access it:

- **Private:** only the staff who uploaded the resource
- **My CO:** all staff from the Country Office
- **WFP:** all WFP staff
- **Public:** partners and researchers with a Data Library account

The CO Head of RAM<sup>4</sup> is responsible ensuring that the data packages are correctly stored in Data Library with the appropriate permissions, in a timely fashion, and always accessible also in case of staff rotation and turnover. If the responsibility of the data storing is delegated, the accountability remains with the CO Head of RAM.



To publish and launch a report, the Country Office should upload the final report to GoDocs and send the link to HQ ([wfp.vaminfo@wfp.org](mailto:wfp.vaminfo@wfp.org)) for posting on VAM DataViz Report Explorer. *The integration DataViz and Data Library for reports will be developed in the course of 2023.*

## HOW TO PACKAGE THE DATASET FOR STORING?

For each completed assessment and monitoring survey, COs should upload the following information in Data Library, each in a separate folder:

- Raw Data
- Syntax
- Internal Processed Data
- External Processed Data
- Output Tables
- Reports
- TOR and Technical Notes
- Questionnaire
- Secondary Info
- Miscellaneous

See also the [Data Library Help pages](#)

### Raw Data

This folder contains the **raw data extracted from the data collection tools** (see *video tutorials on downloading data from MODA or from Kobo*). Ensure that all values and variables in the original dataset are labelled before saving it in raw data folder. Raw data is set to *Private* by default and it is **only accessible to the staff who uploaded the data**.

#### IMPORTANT

Raw and Internally processed Data contain **Personally Identifiable Information (PII) and can only be accessed by the staff who work on the survey.**

[Raw data must not be shared](#) unless there is a specified reason connected to the original purpose of the assessment or survey (e.g. panel analysis).

**The CO should ensure to delete the PII when it is no longer needed** (see [WFP Guide to Personal Data Protection and Privacy](#))

### Processed Data

This folder contains two subfolders:

#### Internally Processed Data

This folder contains the finalized dataset with results from all cleaning, transformation and analysis of data; Personally Identifiable Information (including GPS coordinates) should be retained. The suggested access permission is *My CO*.

#### Externally Processed Data

This folder contains the finalized dataset without Personally Identifiable Information, which have previously been sanitized as explained in the box below. The suggested access permission is *My CO*.

#### Externally processed datasets - sanitization checklist

- **Remove all personally identifiable information (PII)** in the data sets, such as first and last names of respondents, household members, enumerators, supervisors, GPS coordinates, phone numbers, and answers to open-ended questions.
- **Only keep columns referring to administrative unit up to Admin 2** (state level) or at most Admin 3 (district level). Remove lower Admin level columns. Columns referring to sampling design must be retained.
- Include in the datasets **sampling weights** and the following suggested **core indicators**, when available: demographic characteristics, food security-related indicators, food security classification (CARL), coping strategies, expenditures, shocks, income sources, water, assets, shelter, health indicators (WASH section), and child anthropometry. For monitoring, the CO should ensure the relevant mandatory indicators for the programme area in question are captured as per [WFP corporate results framework \(2022-2025\)](#).
- In the case of data from joint assessments and surveys, only share WFP-specific data, unless all member agencies agree to share the whole data set.

### Syntax

This folder contains the syntax/code files used for data cleaning and analysis to document transformations from the original raw data. Syntax should be well commented so that new users can easily orient themselves with the dataset and replicate the same results. The suggested access permission is *My CO*.

### Output Tables

This folder includes the aggregate outputs (e.g., tables, charts, etc.) generated during the analysis. The suggested access permission is *WFP*.



## Reports

Final reports, including all products and publications related to the assessment or survey, are saved in this folder. This includes executive briefs, preliminary result reports, Tableau workbooks and presentations in their final versions and in all available formats. The suggested access permission is WFP.

## TOR and Technical Notes

This folder includes all relevant documents, which explain how the survey was designed and conducted (e.g., the survey Terms of Reference (TOR) and technical notes). From the beginning of the project to the end, record and save methodological and technical notes including survey terms of reference and field notes of how the study was performed. This information is crucial to interpreting the results. The suggested access permission is My CO.

Moreover, this folder should also include documents that refer to:

- Sampling procedures, response rates, and deviations from the sample design;
- Data collection tools – including documents referring to the assessment or survey calendar, modes of collection and a description of data collection activities;
- Analysis plan and timeline;
- Budget plan and resource allocation.

See [Sample information to include in the data package for more examples](#)

## Questionnaire

This folder contains the full and final questionnaire in word/pdf format. If ODK/ONA is used, the XLS form should be saved here. Also, the final electronic version of the questionnaire should also be saved here (e.g. .xml file). The suggested access permission is My CO.

## Secondary Info

This folder includes any secondary data sources which are relevant to the report (e.g., population figures used to calculate the sampling weights). The suggested access permission is My CO.

## Miscellaneous

This folder includes any other files which may be of use but are not related to other folders. The suggested access permission is My CO.

### Unsure how to structure your files?

See a sample folder structure with files

## UPLOADING DATA AND ASSIGNING PERMISSIONS

**Users are in full control of protection levels and accessibility to the data and files.** CO users can set each of the resources to be visible to everyone in WFP (WFP), only to its own Country Office (My CO), or keep it visible only to the analyst who is working on the resource (Private).

As a rule of thumb, **privacy level should be set according to sensitivity and completeness of the resource shared**: for example, output tables, and final reports from completed report should be visible across WFP. By default, some sensitive files (Raw data) can only be set as *Private*. See more about user permissions and accessibility [here](#).

## Upload the data package on Data Library (see video tutorial)

1. Log in to the VAM Data Library with your WFP account
2. Create a new Survey in the country-dedicated container. Fill out the form with all the information requested.
3. Choose the type of resource you want to upload from the list. Choose your file(s) to upload.
4. Name your file(s) according to the guidance provided below
5. Provide a description to your file(s). For example: the name of the survey, date of data collected, data collection method etc.
6. Give appropriate permissions. You can choose between the following accesses:
  - *Private* – access only to specified staff
  - *My CO* – access to the staff in the CO
  - *WFP* – access to all WFP staff
  - *Public* - partners and researchers with a Data Library account

## File Naming System:

Each survey will be automatically named according to date of survey, country code and assessment type. Refer to the following sample name: *YYMM\_ISOCountryCODE3\_SurveyType\_SurveyModality*

- *YYYYMM*: year and month of data collection (e.g., 201805)
- *ISOCountryCODE*: 3-digit ISO name (e.g. ITA for Italy)
- *SurveyTYPE*: type of the assessment or survey, such as FSA, FSNA, FSMS, ENA, JAM, mVAM, PDM, AIM, BM
- *SurveyModality*: modality for the collection of data (F2F, CATI, web)

**Example:** *201805\_ITA\_ENA\_F2F* will identify a face-to-face Essential Needs survey completed in May 2018 in Italy.

## QUALITY CONTROL AND EXTERNAL SHARING BY HQ RAM

If the final external processed dataset is to be shared externally outside of WFP, HQ RAM will review the data packages and ensure they are in the correct format before making them available externally to approved third parties or public.

Data can be shared externally through [the Data Library](#) or other external platforms such as the [Humanitarian Data Exchange \(HDX\)](#). COs should no longer have to share household survey data externally in any other way and should refer requests to the VAM Data Library and/or HDX.

## Endnotes

1 *If released for publication by the data owner*

2 *Examples of these risks include: Poor decision-making processes/quality; Inadequate monitoring, reporting and escalation; Lack of accountability; Investment loss*

3 *Endorsed for publishing by CO management, national government authorities and/or partner agencies, depending on the context.*

4 *Alternatively, this may also be the CO Head of VAM or M&E depending on the CO organizational structure.*

## RESEARCH, ASSESSMENTS AND MONITORING DIVISION

### World Food Programme

Via Cesare Giulio Viola 68/70,  
00148 Rome, Italy - T +39 06 65131

[Photo page 1: WFP/Rein Skulderud](#)

[Photo page 2: WFP/Jonny Hogg](#)