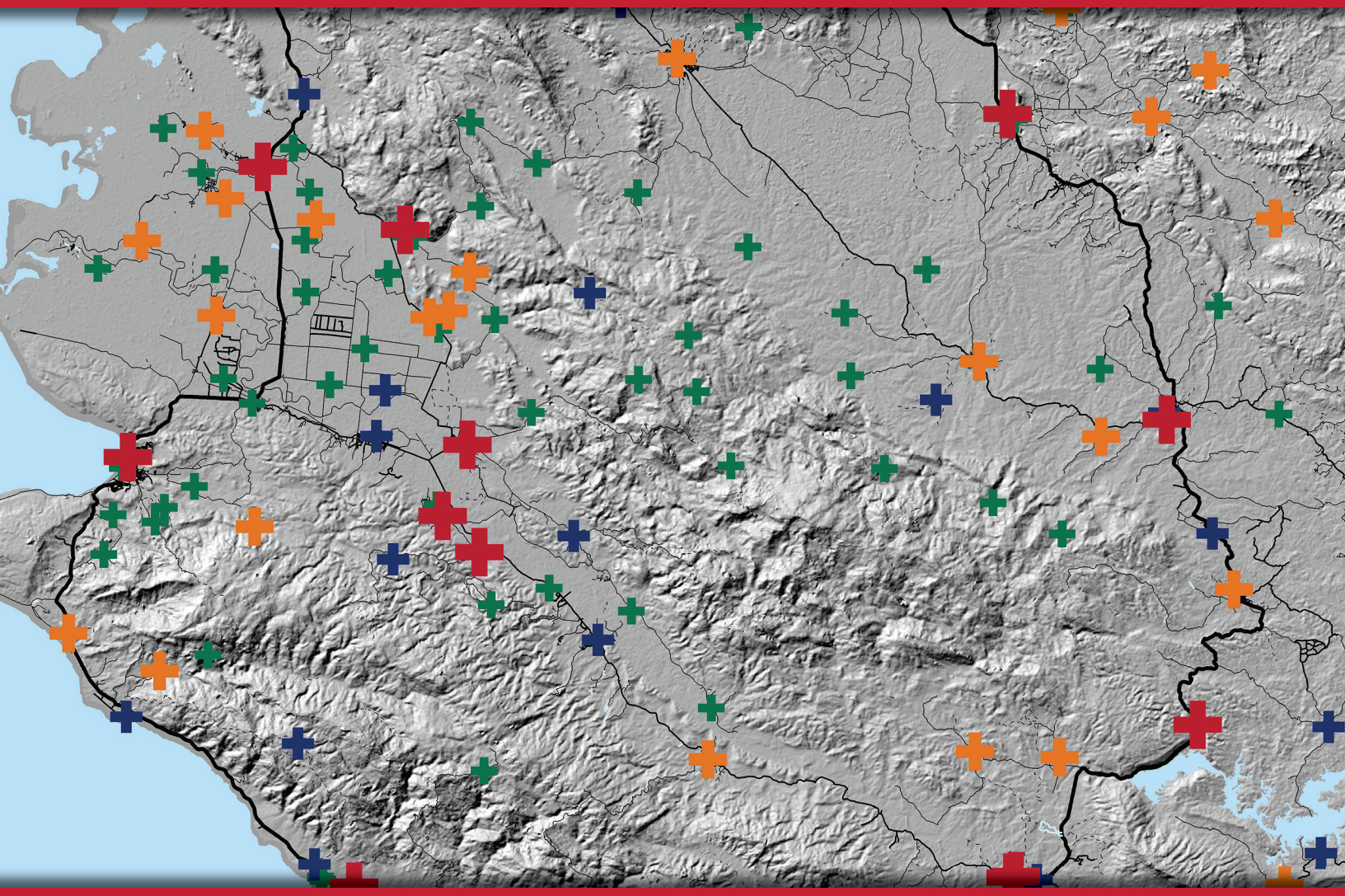
USAID
FROM THE AMERICAN PEOPLE

MASTER FACILITY LIST RESOURCE PACKAGE:

Guidance for countries wanting to strengthen their MFL

Module 5: MFL Data Content



May 2017

DRAFT

MFL DATA CONTENT

This module describes the data that should be included in an MFL. It covers both the minimum data fields to include in the MFL as well as the optional data fields that are commonly included. The module is useful both when setting up an MFL and when considering modifications to the content of an existing MFL. Additionally, the module will help guide those involved in assessment of data in an MFL.

Checklist of things to do before using this module	Module where information is located
☐ Determine key requirements of the MFL	**Key Considerations Module**
☐ Identify available resources	**MFL Assessment Module**

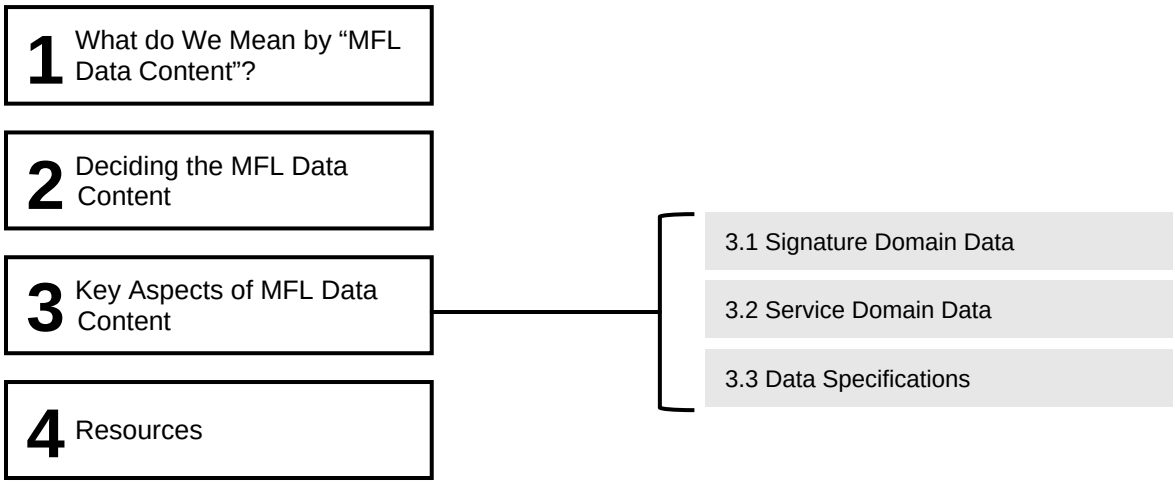
Key audiences for this module:

- MFL Steering Committee
- MFL managers
- Technical Working Group assigned to establish the MFL dataset

Note: words in **bold** are defined in the glossary.

Figure 1: MFL Data Content—Module Outline

(Press Control and click on any of the boxes to be taken directly to that section)



1. WHAT DO WE MEAN BY "MFL DATA CONTENT"?

The MFL data content refers to the information, or **data elements**, that relate to each facility included in the MFL. Typically, an MFL includes both administrative information that can be used to identify and contact the facility (**signature domain data**) and information on the service capacity of the facility (**service domain data**). Both signature domain data and service domain data are described in this module.

2. DECIDING THE MFL DATA CONTENT

It is important to carefully select and define the data elements to include in the MFL. Data elements for the signature domain (see section 3.1) are required while others, including service domain data, are desirable but optional. The inclusion in an MFL of additional or optional data may be useful to MFL **data consumers** but, ultimately, the more data included in the MFL, the greater the cost and the effort required to update and maintain the MFL. Therefore, the decision regarding what data to include in the MFL must necessarily balance the needs of data consumers with the practical consequences of collecting additional data on all facilities, and regularly updating and verifying and that data.[1]

The following are recommended best practices for deciding which data to include in an MFL:

- *Consult potential MFL data consumers prior to deciding on the facility data to include in the MFL.* It is helpful to have a formal process for identifying the types of facility data that stakeholders

[1] For additional information about the process required to keep an MFL up-to-date, see the *Maintaining the MFL Module*.

want or need from the MFL, and the ways in which they plan to use it. MFL data requirements can be captured through **user stories**.[2]

- *Review the data being collected by existing facility lists* and determine how the data are being used, how important these data are to users of the list, and what difficulties are encountered in collecting these data.

- *Identify the sources available for obtaining the data you want to include in the MFL.* Once the data requirements are identified, it is important to identify the potential data sources and the procedures that will be used to collect and validate the data.[3]

- *Start with the **minimum data content** in the MFL.* To adequately manage the data collection and maintenance process, limit the initial number of data elements to those that are absolutely necessary. Add others as additional financial and human resources become available.

- *As much as possible, include facility data that changes little over time.* Information that changes frequently, such as the name of the chief medical officer, requires the MFL data to be checked and updated more often.

- *Work through the MFL Steering Committee* to engage stakeholders in decisions about the MFL data content. The Steering Committee should revisit the data content periodically to reassess new data requirements and to review the resources available for collecting and validating facility data.[4]

[2] See *Key Considerations Module* for additional information on gathering user stories.

[3] See *Establishing the MFL Dataset Module* for additional information on identifying sources and gathering data for the MFL.

[4] For more information on stakeholder engagement and the MFL Steering Committee, see the *Governance Module*.

CASE STUDY: DEFINING MFL DATA CONTENT

Kenya: Held a stakeholder meeting to determine which data elements to include in the MFL. They came up with a minimum standard of what you need to know about the facilities and how often the data need to be updated.

Rwanda: Did a formal information gathering to determine which facility data were already available and, among these, which elements they wanted to include in the MFL. Because of funding constraints, they decided to wait on introducing new data elements, adding them to the MFL in steps. They decided on a small initial list of data elements and are working on adding more data elements to the MFL as funds become available.

Philippines: Began with a large list of data elements; however, when they tried to operationalize the large amount of data content, they encountered problems and decided to reduce the list. At this point, the stakeholders met and determined the key minimum data elements that would be included in the MFL going forward.

Tanzania: Gained consensus on data content from a broad group of stakeholders through the use of key informant interviewers and implementation of a three-day workshop.

3. KEY ASPECTS OF MFL DATA CONTENT

3.1. Signature Domain Data

The signature domain contains data elements that are used to establish a "fingerprint" for a facility. It includes all the information necessary to uniquely identify and locate a specific facility. These data elements should not change significantly over time. The data elements in the signature domain constitute the **minimum data content** for your MFL.

Signature Domain Data Field	Definition of Data Field	Description of Data Field	Example
Unique Facility Identifier	A unique code that identifies a specific facility and distinguishes it from all others.	Serial numbers are often used as unique facility codes. They are simple, compact, and can be stored in any system. Ideally, they are automatically generated by the system. (Note: Additional information about unique identifiers follows this table.)	Serial Number: 125656443

(continued…)

Signature Domain Data Field	Definition of Data Field	Description of Data Field	Example
Facility Name	The official name of the facility	The implementation team will need to agree on naming standards and use a consistent format for all facilities. The facility name should be the official name of the health facility and consist of a single text field. It is recommended that the name be free of abbreviations. Facilities may go by several names, for example if different languages are spoken. In such cases, stick to one language in the main facility name field. Other data fields can be added that include additional names the facility goes by. It is important *not* to include the administrative unit's name or level in the name of the facility, unless it is part of the official name. The location or the type of facility should be included in the facility name *only* if it is included in the official name.	Louis Pasteur Hospital Nairobi Women's Hospital Lema Dispensary
Facility Type	Describes the classification of the facility	Facility types should be determined by a central authority. The MOH may already have a list of standard facility types, with criteria defining each type.	Hospital Primary Health Care Center Dispensary Mobile Health Care Facility
Ownership or Managing Authority	Refers to the entity that owns or manages the heath facility	Ownership and managing authority should be determined by a central authority. Each facility should have just one type of ownership designation. If a facility can be classified under more than one ownership category, the *more specific* designation should be selected. For example, a "military" facility can be classified under "government" and "military," but because "military is more specific, this option should be selected.	Government Military Private Nongovernmental organization Faith-based organization
Location/Address	Refers to the physical location or address of the facility	Ideally, the following specific fields can be defined: • Street Name • Street Number • City/ Neighborhood • State/Province/Region • Postal Code However, given the variability between countries in how addresses are listed, this data element will need to be defined at the country level.	Louis Pasteur Private Hospital 380 Francis Medical Center Pretoria 0001, South Africa

(continued…)

Signature Domain Data Field	Definition of Data Field	Description of Data Field	Example
Contact Information	Information necessary to get in contact with the facility	Separate data elements are required for each type of contact information. The most important data elements are the facility's telephone number, and email address.	+ 223 12 976 5555 xyzdispensary@gmail.com
Administrative areas	Refers to the district, province, or other administrative level in which the facility is located	There will usually be several data elements to cover the various administrative levels in a country. To assure that linkages with other data sources are possible, a standardized list of administrative units should be used. The MOH may maintain health districts or zones, i.e., administrative areas that are specific to the function of the health sector and distinct from the geographic units used in other aspects of a country's governance. In such cases, it is important to understand which administrative breakdown is used by other information systems that the MFL will interact with, and consider whether both the national and the health system administrative boundaries should be used. Each administrative unit should be assigned a numerical designation to clarify the hierarchy of levels. For example, province is level "1", district is level "2", and ward is level "3".	Southern District
Geographic coordinates[5]	Refers to the physical location of the facility, typically represented as latitude and longitude	Both latitude and longitude should be specific in decimal degrees (with positive and negative numbers). For latitude, north is considered positive and south is considered negative. For longitude, east is considered positive and west is considered negative.	The latitude and longitude (in decimal degrees) of Lusaka, Zambia are: Latitude: -15.41667 Longitude: 28.28333

(continued…)

[5] For detailed information on collecting geographic coordinates, see the *Geocoding the MFL Module*.

Signature Domain Data Field	Definition of Data Field	Description of Data Field	Example
Operational Status	Refers to the recognized legal status of a facility intended to provide health services. At any given time, a facility will have a single operational status.	The following are suggested operational categories: • Operational: Facility is open • Licensed: A facility that has been approved and licensed but is not yet operational • Registered: A facility that has been approved as an institution and has been registered • Closed: A facility that has a valid license but is permanently closed • Invalid: A facility where the defining attributes are different from those appearing on the facility license • Does not exist: A facility that has been licensed but has not been verified that it physical exists • Duplicate: The facility exists and is properly licensed but is an effective duplicate of another facility.	
Data Year	The year in which the data was collected	When possible, include the year in which the signature domain data were collected—should be specified for each facility entry. In case of duplicate entries, the latest (most recent) year is considered the valid date. If no data year is available, the field should be left blank.	2015

Further Discussion of Facility Unique Identifiers

Unique identifier codes are one of the most important components of an MFL.[6] They should consist of serial numbers, preferably randomly assigned. A unique identifier code should not include any information about the facility—for example, it should not include a part of the facility name, or reference to the administrative unit—because these characteristics can change over time. Every effort should be made to avoid having to change unique identifier codes, particularly when multiple systems rely on the codes for linkage with their data.

Serial numbers are simple, compact, and can be stored in any system. Manual generation of codes should be avoided because the process is prone to error and duplication of codes. In decentralized systems, where unique identifier codes are generated at the province level, for example, it is important to assign a range of codes to each province to avoid duplication. For instance, province A is assigned codes 0001–2000, and province B is assigned codes 2001–4000.

[6] See *Introduction to the MFL Module*

CASE STUDY: FACILITY CODES

Philippines: The MFL assigns a random unique identifier to each facility. There is no logic to the numbers; they are randomized by the system. There is no geographic association within the number. Initially, they tried to include administrative characteristics (administrative location of facility) in the unique identifier, but the administrative units changed frequently and the facility codes proved too difficult to maintain.

Tanzania: The MFL assigns a random unique identifier to each facility.

Kenya: The system assigns a random unique identifier to each facility registered in the MFL. Therefore, when the administrative divisions in Kenya were redrawn the facility codes (unique identifiers) were not affected.

3.2. Service Domain Data

The service domain contains data elements that describe the basic services available, infrastructure, and human resources at a facility. While the service domain data are important and recommended for inclusion in the MFL, they are *not considered required minimum data content*. The data elements can be included or excluded, depending on budget requirements, donor priorities, and the purpose of the MFL in the country. You will need to work with key stakeholders and the MFL Steering Committee to select which, if any, service domain data elements to include in the MFL.[7]

[7] See the *MFL Governance Module* for more Information on stakeholder engagement and the MFL Steering Committee.

Service Domain Data Field	Definition of Data Field	Description of Data Field	Example
Services offered	Information on the types of services offered by facilities.	A series of data elements list key health services are included in the MFL and facilities are categorized as 'Yes' providing or 'No' not providing that particular service. Information should be adapted at a country level to include the package of services offered through the country's health system, and that are of interest to data consumers.	Family planning Antiretroviral therapy (ART) Labor and delivery
Human Resources	Information on the number of medical personnel by type	The categorization of health personnel is specific to the country. Possible types include, but are not limited to: physicians, non-physician clinicians, registered nurses, and registered midwives. For each type the facility reports the number available. The data should be limited to positive numbers.	Number of midwives: 4
Infrastructure	Information on the number of inpatient and maternity beds and cots present in the facility	For the MFL, it is suggested that only information on impatient beds/cots (including maternity beds) be collected. Other equipment and infrastructure details should be collected through a separate health facility assessment (SAM, SARA, SPA, HFA, etc.). However, additional equipment and infrastructure data may be added to the MFL, if you chose. Responses should be limited to a positive numbers.	Number of inpatient beds: 15

3.3. Data Specifications

Data specifications are guidelines describing how each data element should be defined and formatted for data entry. Data specifications are important for ensuring that information about facilities is collected in a standardized and consistent manner. Each facility record will comprise a series of data elements that describe the details about each facility. For each data element, it is important to clearly define the following attributes:

- Definition: A simple description of the data element
- Data Rules: A description of the format for the data element along with a list of constraints or conditions that should be applied to a data element. For example:
    - Number of characters
    - Use of letters, numbers and symbols (including accents)
    - Capitalization rules
    - Use of abbreviations if allowed, and if so which ones are permitted (e.g., use only Ave. to abbreviate Avenue)
    - Language (including when to use symbols and accents)
- Data Source: Where the information comes from (an individual, survey, organization, or other information system).
- Required, Important or Optional: Some data elements are absolutely required to create a new facility record (*required*); some are fundamental to stakeholder needs but may be difficult to acquire (*important*); and some are simply nice to have (*optional*).
- Missing Values: In all kind of data collection there will be missing values; information may be hard to get, or the respondent does not have the information. It is, however, important to *distinguish missing information from the value zero*. If the respondent does not know the number of beds in a facility, that information ("Don't know") is substantially different from there being no beds in the facility. It is necessary then to assign a code for the missing data. It should always be possible to distinguish the missing data codes from the codes for valid answers. Depending on the valid range of answers, the codes 9, 99, 999…, are recommended to use. Be careful that missing data are not confused with real data (for example, if data are missing for the number of beds in a facility, using the numbers 9 or 99 for missing data may be confusing. It is best to use a number such as 99,999, which is unlikely to reflect the actual number of beds.

These data specifications should be detailed in a data specification document easily accessible to anyone who needs to submit or use MFL data. When adding a new facility to the MFL, or

including new data to a facility record, it is important to make sure all data conform to these specifications.

Examples of different date formats

- June 2, 2002
- 2 June, 2002
- 06-02-2002
- 02/06/2002
- 2/6/02

International data standards can be used to define data specifications. Data standards are agreed upon rules for how data should formatted, defined, structured, managed and used. The use of international data standards is important for sharing data, especially for integrating the MFL with other information systems; it allows both systems to share a common language and understand what the data mean.[8] For example, dates can be entered in various ways as illustrated in the box to the right. Data standards will dictate which format to use consistently to avoid confusion and complication when exchanging data. For example, data standards may require that all dates be formatted as DD/MM/YYYY.

When selecting data standards, it is important to consult with the managers of other information systems with which the MFL will share data to determine which are already in use and which make the most sense in that country context.

4. RESOURCES

- Haiti MFL Codebook
- Tanzania Health Facility List Data Specification
- Rwanda Registry Specifications (See Page 17)

[8] See *Sharing the MFL Module* for more information on integration.

ACKNOWLEDGEMENTS

The MFL Resource Package was developed with extensive input from a team of persons who have been involved in various capacities in the development or management of MFLs in different countries. The content builds off of previous MFL guidance developed by the World Health Organization, MEASURE Evaluation and Open HIE. This MFL Resource Package seeks to expand and update the guidance and make it accessible to a wide audience. Development of this Resource Package included a literature review, a series of in-depth interviews with key informants, a three-day meeting attended by various experts in this area to discuss in detail the content and structure of the guidance document, and a thorough review process.

Cristina de la Torre and Clara Burgert from ICF led the development and drafting of this guidance document. Lwendo Moonzwe, and Kirsten Zalisk (from ICF) and Aubrey Casey (formerly from ICF) helped to draft the MFL Resource Package, organize resources, and document discussions during the three-day meeting. Andrew Inglis (formerly from MEASURE Evaluation/JSI) and Scott Teesdale (from InSTEDD) helped draft sections of the MFL Resource Package.

Lynne Franco led a team at EnCompass to conduct a series of in-depth interviews to inform the content of the Resource Package, and subsequently helped facilitate the three-day meeting to review the guidance proposed for the MFL Resource Package.

The following tables list persons who contributed to the MFL Resource Package by attending a three-day meeting, participating in in-depth interviews, contributing resources, reviewing drafts or providing information for the case studies.

Table 1: Persons who participated in the three-day meeting to review the content and structure of the Resource Package.

Meeting Participants	Affiliation
Tariq Azim	MEASURE Evaluation/JSI
Noah Bartlett	USAID, Bureau for Global Health
Clara Burgert	The DHS Program/ICF
Aubrey Casey	The DHS Program/ICF
Niamh Darcy	RTI
Anita Datar	Health Policy Project/Futures Group
Cristina de la Torre	The DHS Program/ICF
Mark DeZalia	PEPFAR/CDC
Lynne Franco	The DHS Program/EnCompass
Erick Gaju	MOH Rwanda
Nate Heard	US Department of State

Meeting Participants	Affiliation
Andrew Inglis	Deliver Project/JSI
Denise Johnson	MEASURE Evaluation/ICF
James Kariuki	PEPFAR/CDC
Esther Kathini	MOH Kenya
Carl Leitner	iHRIS/Capacity Plus/IntraHealth
Lwendo Moonzwe	The DHS Program/ICF
Annah Ngaruro	MEASURE Evaluation/ICF
Kola Oyediran	MEASURE Evaluation/JSI
Jason Pickering	Consultant/DHIS2
John Spencer	MEASURE Evaluation/UNC
Charity Tan	MOH Philippines
Scott Teesdale	Open HIE/InSTEDD
Kavitha Viswanathan	WHO
Sam Wambugu	MEASURE Evaluation/ICF
Kirsten Zalisk	The DHS Program/ICF

Table 2: Persons who contributed through interviews or review of the MFL Resource Package Modules.

Name	Affiliation at time of participation
Ian Wanyeki	Health Policy Project/Futures Group
Elaine Baker	Health Policy Project/Futures Group
Bernard Mitto	Health Policy Project/Futures Group
Vanessa Brown	PEPFAR/Department of State
Robert Colombo	WHO
Steeve Ebener	Gaia Geo Systems
Mike Gehron	PEPFAR/Department of State
Karin Gichuhi	Office of HIV/AIDS/USAID
Marty Gross	Bill & Melinda Gates Foundation
Jason Knueppel	BAO Systems
Rachel Lucas	USAID
Andrew Muhire	Rwanda MOH
Martin Osumba	AFYAinfo, Kenya
Alyson Rose-Wood	Office of Global Affairs/HHS
Dykki Settle	iHRIS/IntraHealth
Jim Setzer	Abt Associates, Inc
Ashely Sheffel	Consultant/WHO
Brian Taliesin	Digital Health Solutions/PATH
Ola Titlestad	DHIS2/University of Oslo

The MFL Resource Package was undertaken with support from the United States Agency for International Development (USAID) and the President's Emergency Plan for AIDS Relief (PEPFAR) through The DHS Program.