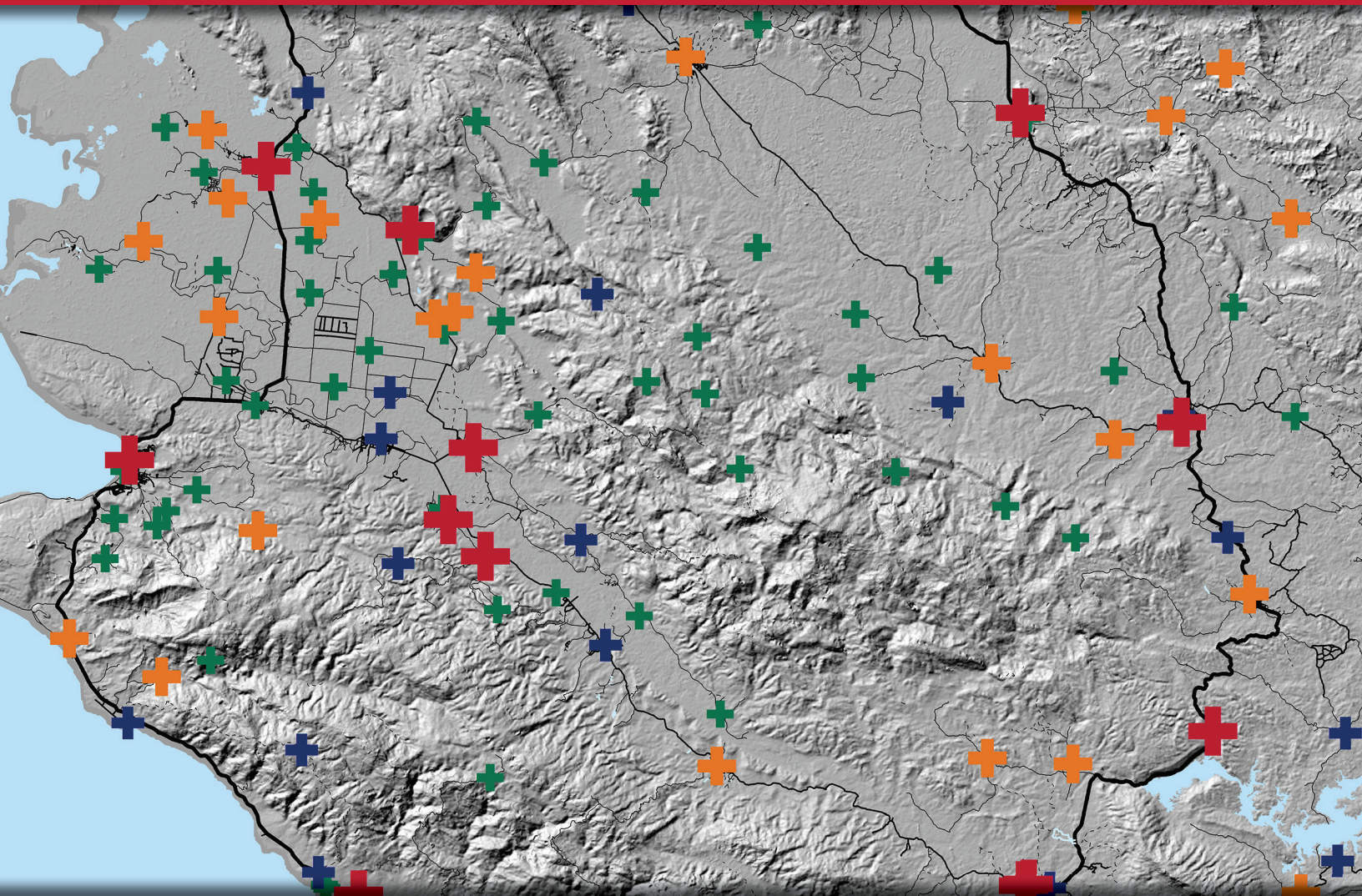# MASTER FACILITY LIST RESOURCE PACKAGE:

Guidance for countries wanting to strengthen their MFL

Module 10: Sharing the MFL

May 2017

DRAFT

SHARING THE MFL

This module describes the processes and considerations necessary for sharing the Master Facility List (MFL). It presents information on (1) the importance of a data sharing policy, (2) decisions regarding what to share and with whom, and (3) what additional documentation needs to be shared along with the data. Finally, the module describes considerations around integration of the MFL with other information systems.

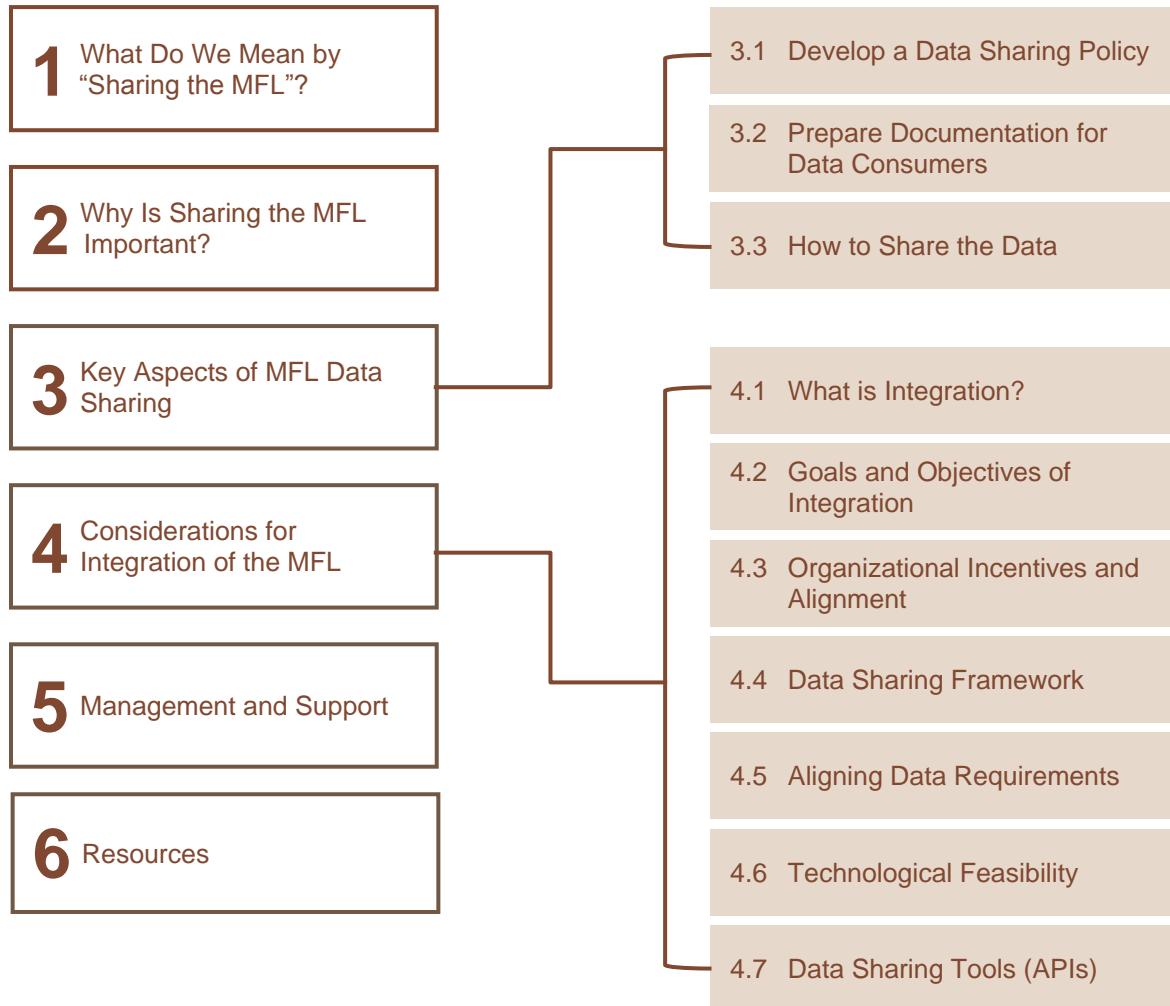| Checklist of things to do before using this module	Module where information is located
☐ Set up a steering committee to lead the process of developing and strengthening the MFL	**MFL Governance Module**
☐ Determine the requirements for the MFL	**Key Considerations Module and Establishing a Facility Registry Service Module**

Key audiences for this module:
• Steering committee for MFL strategic planning • Managers of the MFL • Technical Working Group responsible for establishing the facility registry service • Developers

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

Figure 1: Sharing the MFL – Module Outline

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

1 What Do We Mean by "Sharing the MFL"?	**3.1** Develop a Data Sharing Policy
	3.2 Prepare Documentation for Data Consumers
2 Why Is Sharing the MFL Important?	**3.3** How to Share the Data
3 Key Aspects of MFL Data Sharing	**4.1** What is Integration?
	4.2 Goals and Objectives of Integration
4 Considerations for Integration of the MFL	**4.3** Organizational Incentives and Alignment
	4.4 Data Sharing Framework
5 Management and Support	**4.5** Aligning Data Requirements
	4.6 Technological Feasibility
6 Resources	**4.7** Data Sharing Tools (APIs)

1. WHAT DO WE MEAN BY "SHARING THE MFL"?

Sharing the MFL entails making MFL data accessible to **data consumers**. There are a number of ways to share the MFL, some more effective than others. The following are ways that MFL data can be shared:

- The dataset is shared using a spreadsheet that can be downloaded or emailed.
- A public portal is established through which users can access and manipulate the data.
- The MFL is integrated with other information systems.

While data sharing itself can be fairly straightforward, a number of issues need to be considered before implementation. This module discusses the various aspects of data sharing that need to be taken into account.

2.   WHY IS SHARING THE MFL IMPORTANT?

The primary value of an MFL lies in people's ability to access and use the facility data it contains. Many types of stakeholders can benefit from accessing a comprehensive and up-to-date list of health facilities, including government ministries, donors, and development partners. It is advisable to share the MFL as broadly as possible. The following are additional advantages that typically come with sharing MFL data.

- The greater the number of data consumers able to access and benefit from using the MFL, the more valuable the MFL becomes. With accruing value, the MFL is more likely to become a government priority and more likely to receive the resources needed to remain current and to function optimally.

- Sharing tends to improve the quality of the MFL data because, with a wider set of "eyes," there is increased likelihood that someone catches outdated or erroneous data. Further, data consumers who value the MFL may be more likely to scrutinize the quality of the data they are using.

- A widely shared MFL ensures consistency of facility data across systems. Stakeholders who use the MFL as their primary facility list will have the benefit of standardized information as well as the same set of unique facility identifiers across organizations and information systems. They can then more readily link up the data and exchange information.

- Sharing the MFL can potentially be a source of revenue if an access fee is applied to selected private sector users (e.g., health insurance organizations).

3.   KEY ASPECTS OF MFL DATA SHARING

3.1.   Develop a Data Sharing Policy

A key aspect of the planning process for MFL data sharing is development of a data sharing policy. Having a comprehensive data sharing policy is important to assure that the following activities and requirements are adequately considered.

- Sharing procedures and decisions are transparent and known to current and potential users of the MFL

- Sharing of MFL data complies with existing policies governing national data

- Management of the MFL is more efficient because of clear processes and procedures

- Requests for MFL data are addressed and handled consistently

A data sharing policy should specify the types of MFL data that will be shared and who has access to the data. In addition, the policy should describe in detail the decision-making and approval processes associated with MFL sharing. A number of factors need to be considered in deciding these issues; these are explored detail below. (Considerations specific to integration are discussed in section 4.3.)

What MFL data are shared?

The data sharing policy will need to specify what MFL data can be shared. The types of MFL data shared and how broadly the data are shared depend on a number of factors including (1) national and institutional data policies, (2) the needs of data consumers, and (3) the presence of sensitive data in the MFL that may require protection.

National and institutional policies

Countries may have existing policies that govern how data are to be stored and shared. In some cases, there may be open data mandates that require all data to be accessible; in other cases, countries may be more protective of their information and have strict rules about who can access data, what types of data can be shared, and the procedures to follow to obtain permission to access the data. Having a thorough understanding of the policy context is important when making decisions about MFL sharing.

Consider not only national policies but also policies specific to the government institutions affiliated with the MFL. Further, if the MFL is pulling data from other sources (e.g., a national health facility licensing board), it is important to determine whether the data sources have restrictions on sharing or redistributing those data.

Needs of consumers

Deciding what data to share should take into account the needs of the MFL data consumers. If the data are too restricted, the MFL ceases to be useful to stakeholders. However, the MFL may have data that only a limited number of consumers require; then it may not be necessary to make these data publicly available. Data consumers can be assigned to different levels of access; the process is described below in the section on access to data.

Presence of sensitive data

Some data contained in the MFL may be highly sensitive. For example, the location data of facilities that serve highly vulnerable populations can, if misused, pose a threat to clients of these facilities. Types of information in the MFL that may be regarded as sensitive include the following:

- Certain categories of services provided at facilities (e.g., prevention services for sex workers; rehabilitation services for people who inject drugs)

- Geo-coordinates of certain types of facilities (e.g., HIV outreach centers)

- Personally identifiable information (e.g., name and contact information for facility director)

Sensitive information should be restricted to trusted users. The types of data that are accessible to different groups of users should be clearly stated in the MFL data sharing policy.

Who has access to the data?

Different types of data consumers should have differential access to MFL data. In general, data consumers can be classified into four domains of access:

- *Trusted domain users*: These are government officials who work directly with the MFL or need the MFL to populate government information systems. They have broad access to the MFL data and therefore require additional security and logins to access the data.

- *Middle domain users*: These include development partners and insurance companies, both of whom need fairly comprehensive facility data to carry out their work, but may not need all the details in the MFL. Security measures should be in place for this group too.

- *Public domain users*: Public access to non-sensitive MFL data is recommended so that as many data consumers as possible can use the data. In addition to the general public, this group typically includes researchers, survey industries, and marketing companies. Such public access need not require credentials; users can simply access the MFL data via a public portal.

- *Administrative domain users*: A small group of individuals consisting of data curators and others directly involved in the establishment and maintenance of the MFL are granted complete access to the database, including editing rights.

The MFL data sharing policy should define these user domains for the specific country context. It should also describe processes for entering into data use agreements, if these are required. Data use agreements can be established with data consumers to clarify how the MFL data can and cannot be used. The agreements can also specify any rules for attribution (e.g., acknowledging the MFL and the institution that manages it), and any restrictions, if needed, regarding the redistribution of MFL data to third parties. Usually, these agreements are not required for the public user domain.

Approval process

The MFL data sharing policy should outline the processes for evaluating and approving requests for access to the MFL data. The approval process may vary depending on the user domain and type of access requested. For example, no approval may be needed for accessing

basic data from a public portal, whereas a more careful vetting process is needed for approving a request for system integration with the MFL.

The policy should indicate the persons responsible for approving share requests, the criteria used in making these decisions, and the expected timeframe for reviewing the request.

3.2. Prepare Documentation for Data Consumers

Data consumers will need additional information about the MFL dataset in order to understand how the data are structured and how they can be used. This background information about the MFL data is often referred to as **metadata**. The following documents should be readily accessible to data consumers, preferably through online document repositories.

- Data specifications document[1]—Describes how each data element in the MFL is defined and the parameters associated with the data.

- Summary of the process for updating the MFL and the frequency of updates—This document is important first because it informs data users how current (up-to-date) the data are; second, it alerts data users to possible changes to the dataset resulting from the update process.

- Log of recent changes to MFL data—The log includes the dates of changes to MFL data.

3.3. How to Share the Data

A number of approaches can be used to share the MFL data. The decision on which approach to use should be based on user requirements that specify in which format they need the data. It also depends on the type of facility registry service that is established to house and share the MFL.[2]

- The MFL dataset can be <u>sent electronically</u> (e.g., by email). If resources are not available to develop other means of data sharing, this sharing approach will work. However, it is not generally recommended because it is difficult to control who gets access to the data and its distribution can lead to problems of version control later if people continue circulating an outdated MFL file.

- The MFL can be made available as a <u>read-only document on a website</u> where users can view or download it as needed.

---

[1] See *MFL Data Content Module* for more information on data specifications.

[2] See the *Key Considerations for the MFL* and the *Establishing a Facility Registry Service* modules.

- The MFL can be shared <u>through an online interface</u> that allows users to query, filter, and download the data. Examples include: Kenya, the Philippines, and Tanzania.

- The MFL data can be shared <u>through integration with other information systems</u>.

Because data consumers are increasingly interested in integration of the MFL, this option is discussed in detail in Section 4.

4.   CONSIDERATIONS FOR INTEGRATION OF THE MFL

4.1.   What is Integration?

Integration is the process of physically or functionally linking multiple information systems to create a combined system or unified solution. **Data Integration** refers to the combination or exchange of data from one or multiple sources into a tool or platform that uses the acquired data for transactional or analytical purposes. Integration with the MFL is necessary to enable other systems that require the facility list to access the most recent (updated) information.

Most often, integration of the MFL aims to share the MFL facility data with other systems that need a comprehensive list of facilities. However, integration can work in the other direction, with information systems sending updates to the MFL.

Two common approaches to data integration involve the following:

Integration vs Interoperability

Integration is the process of linking multiple systems, whereas *interoperability* is an intrinsic property of the systems themselves which describes their ability to exchange and interpret data.

Interoperability is defined as the extent to which systems and devices (in this case the facility registry service) can exchange data and interpret that shared data. For two systems to be interoperable they must first be able to exchange data and second understand that data so that they can be used by data consumers without changing the data's semantics.

Note: Definition adapted from the Healthcare Information and Management Systems Society (HIMSS)

- *Data synchronization*: This is an automated process through which one system (e.g., the HMIS) updates its facility list by checking for discrepancies, and then harmonizing with the content of another system (i.e., the MFL).

- *Data warehouses*: These are repositories that store data from multiple sources. The data can then be combined for analytical purposes.

To successfully integrate a facility registry with other systems requires both technical and program management activities. Good management and governance are essential to ensuring that the interested parties are in agreement and that the technical solution for integrating the

MFL is acceptable to all those involved. The process of integration follows the five steps outlined below.

- Goals and Objectives of Integration
- Organizational Incentives and Alignment
- Data Sharing Framework for Integration
- Aligning Data Requirements
- Technical Feasibility

4.2. Goals and Objectives of Integration

After one or more systems have been identified for integration with the MFL, the first step is to define and agree with collaborators on the goals and objectives of integration. Simply put, what will be achieved through working together that cannot be achieved through the current status in which each system is independent? The reasons for integration vary among stakeholders because each group has different interests and requirements. The may include goals for integration may be dictated by implementers, programs, funders, governments, policies, caregivers, and even individuals, such as analysts who interact with each system. Creating a common vision of what is to be achieved is critical in the early stages of the process. It may be useful to organize an integration workshop to help implementers identify cross-cutting problems and goals.

4.3. Organizational Incentives and Alignment

With the goals and objectives of integration defined, the next step is to ensure that organizational incentives are present and that all parties are aligned regarding achieving those goals. Affected organizations need to consider the following factors:

- Whether the integration aligns with institutional mandates and policies
- The benefits to be gained individually and collectively from integration
- The real and perceived risks of integration
- Whether staff have the capacity to manage an integrated system
- How the upfront (capital) costs and ongoing (operational) costs will be funded

Maintaining integrated systems requires ongoing effort. Sometimes, the ongoing cost of the integration falls on the organization that has the technical capacity to keep things running, not the organization that has the mandate and resources, or derives the most value from the integration. This scenario highlights the potential divergence in incentives and alignment in successfully achieving and maintaining integration.

Trust plays an important role when the various teams and organizations are in the process of integrating systems. At the outset, transparency in sharing incentives and disincentives to integration is important to help groups resolve concerns related to data.

4.4.   Data Sharing Framework for Integration

When the collaborating organizations have (each) determined that there are sufficient incentives and alignment to pursue integration, the next step is to make sure there is a data sharing framework in place that adequately details the governance and contractual requirements for integration. Ambiguity in the data sharing framework is a major deterrent to integration.

In addition to the considerations discussed in the section on developing a MFL data sharing policy, a data sharing framework for integration should clearly define the following issues.

- What data are shared through integration? In which direction? How often?

- What are the assumptions about read/write/administrative access?

- How will data be curated and kept up to date?[3]

- Are there data sharing policies that affect each organization? Which ones are they? Who imposes them? Are the policies compatible?

- What are the constraints with respect to access rights and physical placement of data? Does it matter when and where data are stored, who has access to them and how, and for how long?

- Are there different guidelines for different types of data?

4.5.   Aligning Data Requirements

A major step in the integration process is determining whether integration is possible in terms of data compatibility. There are two aspects to this issue that need to be considered.

- How are the data formatted?—This is referred to as syntactic interoperability. Given the current widely used syntactic markup languages and schema standards such as XML, RDF, and JSON, differences in formatting are surmountable and syntactic interoperability is relatively easy to achieve.

- How are the data defined?—This is referred to as semantic interoperability. In a scenario in which two systems are to be integrated, it is vital to determine any differences in how data elements are defined and managed, and understand the resulting limitations. The goal is to align the definitions and constraints that are inherent with the data elements that are defined in the MFL and in other systems. Having access to proper documentation about the

---

[3] The *Maintaining the MFL Module* provides a detailed discussion on curating and maintaining the MFL.

data is necessary for completion of this step (see section 3.2 *Preparing Documentation for Data Consumers*).

There are a number of important checks that need to be made before integration of the MFL data can be implemented.

- Check that facility identifiers match and that legacy identifiers are preserved.

- Check that geographic and/or administrative hierarchies match. For example, are the same administrative boundaries, names, organizations, and levels of specificity being used?

- Check that the facility types and categories of services are defined in the same way.

- Check whether some information (e.g., facility ownership) is included in a single data element with multiple response categories, or kept in several data elements with yes/no responses for each.

- Check that there is agreement on what *empty, NA,* and *nulls* mean

- Check that there are records of when data were last updated and by whom (for quality control purposes).

- Check that there is maintenance of ontology mappings between terminology standards and project datasets.

4.6. Technological Feasibility

The last step in the integration process is to consider whether integration is feasible at the existing technological level.

For almost every scenario imaginable there are numerous proven solutions that resolve technical interoperability needs within even the most exceptional constraints. Common challenges that must be overcome with regards to integration are the following:

- System deployment and connectivity—Which systems are installed where, with access to what, and under whose control? For example, a computer running in an office of a non-governmental organization may access the Internet, but itself may not be addressable as a web service. This creates practical constraints such as "who calls whom" and "push vs pull" notifications in a given integration scenario. Also, inconsistent connectivity from mobile devices, facilities, and general Internet availability in low and middle income countries make it necessary to cache data and queue messages for reliability. It may also be required to put processes in place to resolve conflicting updates or lost messages.

- Accessing data securely—Typically, secure data access implies authentication (securely identifying users and systems), authorization (limiting who can see certain data or perform a particular function), and auditing (tracing what was done). Sometimes an organizational obstacle to integration emerges due to the lack of consensus about who manages these overall permissions and how, often resulting in access that has to be maintained point-to-point. While the technological considerations all need to be specifically addressed, they are generally the easiest challenges to resolve in the process of integration. For health care and other fields, technological obstacles and resulting solutions have been well documented, and can be applied to future integrations.

Programmatically, core details for successfully carrying out the technical aspects of an integration include the following:

- Reach agreement on the direction of data integration; that is, who is the source of data and who is the consumer?

- Agree on push or pull, who triggers these events, and when. This might be a manual, scheduled, or triggered process.

- Agree on where the trigger and script will be hosted or run. For example, is it in the facility registry server, a HMIS server, or a third party service?

- If the integration is done via a bridge script or point to point, make sure to have ad-hoc service credentials or use single-sign-on OpenID tokens to avoid having credentials lying around in other servers.

- Try to run the scripts against staging or replica datasets before production.

- Use interoperability profiles of standards and interoperability specifications to reduce the surface area, cost, and complexity of implementing standards.

- Integrate first and standardize later; this will increase efficiency and ensure that the standardization process is well informed.

4.7. Data Sharing Tools

Application programming interfaces (APIs) are tools that enable integration and the exchange of data. Multiple APIs are available to facilitate the automated transmission of data across systems. These are technical tools used by developers, but it is good to be aware of them as reusable options to connect systems and address interoperability with standards based tools. The following are examples of APIs that have been used for MFLs.

- Facility Registry API—Is a RESTful style API that was developed within the OpenHIE community to support integrations for facility data (see Section 6: Resources).

- Care Services Discovery API—Is a method used to share facility data along with health worker data, through the use of an interlinked registry (see Section 6: Resources).

- Other APIsand data exchange formats—It may be desirable to use or implement other API standards, depending on the use case, technical staff experience, or limitations of the methods above. These include native APIs for specific facility registry service solutions, such as the DHIS2 API or Resource Map API. Additionally, facility registries may support data exchange via file transfers in formats such as GeoJSON, RSS, and CSV. These are different data formats that particular users may want to take advantage of, particularly where there are existing data sets available but a lack of resources to develop an API for an automated process.

5. MANAGEMENT AND SUPPORT

Ongoing management and support of a facility registry's integration and data sharing needs is important for its long term success. Together, the lead implementing group, supported by the steering committee and technical partners, constitute a proven combination for success, while also supporting local ownership and sustainability. Ongoing support for MFL sharing should include the following:

- *Designated person to respond to support requests as they come in from users*. The response at times may be that there is no way to fix an issue at this time, but having a point of contact to work with users and seek work arounds in these instances is extremely valuable.

- *Technical staff that can handle technical issues and fix bugs when they come up*. A time and materials agreement here can be a cost-effective approach, so staff are engaged only as needed.

- *Routine meetings of the steering committee to plan and maintain a strategic vision for the MFL* and the associated integrations. This may include additional fundraising or petitioning for resources if substantial enhancements are required.

Funding for technical support activities can be combined with routine support for the home institution's systems. While most facility registry service efforts and related integrations have, to date, been driven by MOH and funder-related grants, other cost-sharing options may be possible, including the following:

Tiered approach to cost sharing—Access to some data is free while access to other data requires payment of fees.

Cost recovery model—Some partners or data users pay fees depending on how much they use the data. Similar models also depend on use or access rights to cover costs.

No fee for use of MFL data—Access to MFL data and services is provided by the government or owner of the registry, and no fees are charged.

6. RESOURCES

- Facility Registry API
- CSD: IHE Documentation
- OpenHIE Workflow: Query health worker and/or facility records

ACKNOWLEDGEMENTS

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