PEPFAR

U.S. President's Emergency Plan for AIDS Relief

Data for Accountability Transparency Impact Monitoring (DATIM)

Data Import Reference Guide (DRAFT)

U.S. Department of State

U.S. Office of Global AIDS Coordinator (OGAC)

Data Import Reference Guide

1 DOCUMENT SCOPE

This DATIM Data Import reference guide serves as a resource outlining the steps involved with importing data into DATIM. Data intended for import into DATIM must satisfy strict requirements with respect to the format of the data as well the relationship of the data to the current metadata of the system. Data must be valid with respect to the destination sites, indicators, disaggregation and mechanisms. Additional checks such as valid types of data (integers, options, etc.) are also expected to be satisfied. Data intended for data import into DATIM should also be free of validation rule error violations. In short, the data imported into DATIM through an import process should satisfy all of the requirements as if it had been entered manually through the data entry screen.

This guide will provide guidance to DATIM data importers explaining these key concepts in more detail. While detailed technical reference material will be referred to, this guide will not provide detailed step-by-step instructions on how to transform data from an external system into a format suitable for import into DATIM due to the wide variability and differences in use cases. However, the basic principles of preparing data for import into DATIM will be the same regardless of the details of the source system.

Currently, data submitted by operating units will need to be validated first by the DATIM help desk prior to import. If there are errors in the data, this can lead to delays. Ideally, data submitted to the DATIM help desk have been tested thoroughly for both formatting as well as content errors, prior to submission. **Data should be submitted to the help desk AT LEAST five days in advance of any reporting deadline** to provide enough time for validation, resolution of any problems, and import testing.

2 CHARACTERISTICS OF DATIM DATA

Data values in DATIM are always associated with four dimensions, which describes different aspects of the data (reference *Figure 1*).

- **Where** – The Organization Unit. This dimension describes the location of the data.
    - Example: Clinical Facility, Community Site, or OU Level
- **What** – Data element + Disaggregation: This dimension describes the phenomena which the data value is attached to, such as the *Number of HTC tests* for *Females 1-4*. In this case the data element has a disaggregation of *Female 1-4*. Data elements in DATIM typically have many different disaggregations with respect to age and gender. In addition to that, different classes of data elements such as *Targets* and *Results* and *DSD* and *TA* exist.
- **When –** Reporting Period: This dimension describes the time period of the data being reported.
    - Example: January through March 2016 or April through June 2016
- **Who –** Funding mechanism: Also known as the "attribute option combination" this is an extra custom dimension in DATIM, which describes which implementing mechanism the data is being reported for.

All data in DATIM contain those four dimensions, also known as meta-data.

> ***NOTE:** No data entry, regardless of the method, is possible unless ALL meta-data are properly associated and valid.*****
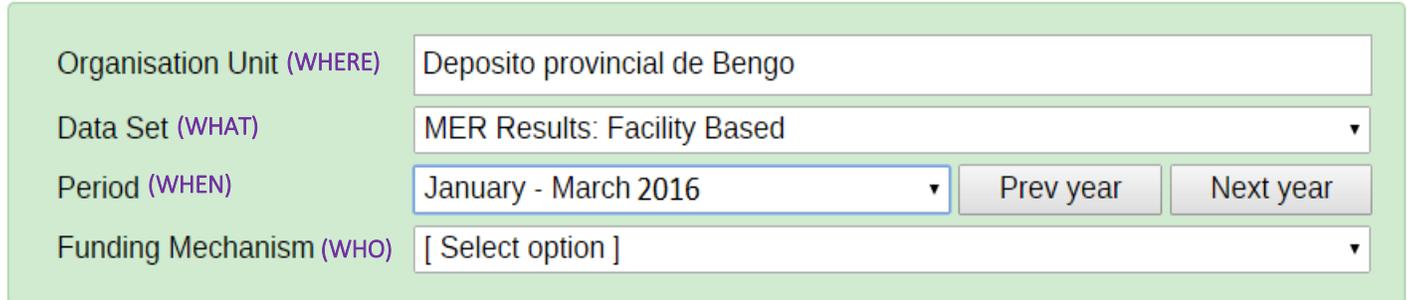
2.1. Manual Data Entry in DATIM

Implementing Partners currently enter a majority of the data in DATIM. Prior to data being entered, USG staff will use the data entry interface to enforce the selection of all meta-data to ensure that only valid organization units, data sets, periods, and funding mechanisms are selected.

Data Entry ⓘ

Organisation Unit (WHERE)	Deposito provincial de Bengo	
Data Set (WHAT)	MER Results: Facility Based	▾
Period (WHEN)	January - March 2016 ▾	Prev year Next year
Funding Mechanism (WHO)	[Select option]	▾

Figure 1: *Example of manual data entry selection in DATIM*

2.2. Import Data into DATIM

Data import is the process of importing the data into DATIM using the *Import-Export* application, which allow users to bypass the need to enter data manually and to import the data values in bulk (which typically have been exported and transformed from an external system).

There are currently two methods for importing data into DATIM:

- **DATIM Web Application Programming Interface (API) -** Submit the data file exchange file electronically through the DATIM Web API. This implies that the target system is capable of submitting data to the DATIM Web API using a HTTP POST request.
- **Import-Export Application -** Upload appropriately formatted files through the DATIM user interface via the *Import-Export* application.

As mentioned previously, the current policy of DATIM requires all import files to be validated by the DATIM help desk prior to submission. The help desk will make a decision as to which method of import is most appropriate for a given data import file.

2.2.1.   Construct Appropriately Formatted Files for Import

During the initial setup, Country Teams and implementing partners will need to document the processes undertaken and resources used in the data import effort. This may include:

- Concurrence from the Inter-Agency, Agency, and / or Implementing Partner Lead
- Structured coordination of importing efforts
- Resources for the process (time, expertise, financial, human)
- Documentation for data mapping, instructions, or rationale for mapping
- Process for test file creation
- Data import test plan and results
- Country data exchange process flow diagram
- Long-term data import/exchange strategy for phased implementation.
- Ongoing documentation of challenges and lessons learned

Consider time for corrective action, planning, and updating these processes as they are considered iterative. The documents created and used by countries for planning and implementing data exchange will serve as great resources for other countries. Countries with well-documented and set processes should consider sharing their resources.

Note that the current policy requires each OU to submit a single data import file. This means that if multiple implementing partners of an OU plan to avail DATIM data import services, they must coordinate with their country USG POCs to create a single consolidated data import file for submission to the DATIM help desk. The country USG

Datim Support

Data Import Reference Guide

POCs should submit a help desk ticket to seek technical support from the DATIM help desk data import team. **The in-country SI advisor should submit the final consolidated data import file to the help desk.**

2.2.2.   Communicate Data Import Plans to OGAC

Implementing Partners (IP) seeking to import data must express their interest to their in-country SI contacts. When consent is given, IPs must work with their own developers, or others familiar with data import. Please note that an implementing partner can complete the testing and validation of data import files individually but after completing the validation tests, the importer should coordinate with the USG POCs to create a single consolidated data import file for the concerned OU.

Have all the appropriate and necessary documentation (as seen in the previous section), along with the necessary technical capacity to support an import. Request support from the DATIM Technical Support team by selecting the *DATIM Support* icon found under *Apps* in DATIM. From here, complete the following:

1.   Select *Submit a Request* found under the Technical Support Section. This will take the user to a new page where they can select a topic.

2.   From the dropdown menu, select *Data Import / Exchange*.

3.   After selecting *Data Import / Exchange*, additional questions will appear. Answer them accordingly, and be sure to complete the mandatory fields or the request can't be submitted. Under *Description*, indicate the plan for importing data into DATIM.

4.   Click *Submit* when ready to send your request to the helpdesk.

2.2.3.   Data Import Testing

After submitting and receiving approval the Data Import / Exchange request, data importers should thoroughly test the data import through DATIM's data exchange testing (DEV-DE) environment. This is a dedicated server which is a copy of the actual DATIM production system which can be used for testing purposes. It is simply a testing environment which can be used to validate data prior to import into the actual production system, without actually affecting the data stored in production.

Contact the DATIM Help Desk to ensure that your user has been enabled in the DEV-DE environment.   Please note the following while testing in the DEV-DE environment:

- DO NOT conduct any testing in DATIM PROD
- User credentials will be the same in DEV-DE as they are in DATIM.
- Data should only be submitted to the DATIM Help Desk after you have successfully imported the data payload into DEV-DE. Submitting data which has not been able to be imported into DEV-DE will lead to delays of your import.

Data Import Reference Guide

Figure 2: *Example of Data Import Summary*

Although the exact number of values will vary based on the actual import, when importing data with the Import/Export module of DATIM, no conflicts should be reported (similar to the picture above). If you receive any conflicts from the server, they should be resolved prior to submitting your data to the DATIM help desk for further validation.

2.2.3.1. Verification of Data Elements and Disaggregations

Data in DATIM must have a data element and disaggregation (also known as a category option combination). Many data elements have multiple levels of disaggregation (age break downs, gender, service type, etc). Note that the validation rules will validate the data for data elements which have an option set and if the value provided does not match a valid option, the data will be considered to be invalid. DATIM also contains a number of legacy data elements, which are present in the system, but which may not currently be active for data entry. If you are using the Web API of DATIM to prepare your file, you should take great care to ensure that the data element / disaggregation combination used for the data values present in the data import file are current. The ultimate reference for this are the data entry screens of DATIM itself. If there are any invalid data element / disaggregations present in the import file, a conflict will be reported during testing. Code lists have been made available to help you to develop mapping files for data elements and disaggregations below.

Dataset	HTML	Excel
MER Results Data Elements 2015	Code List	Code List
MER Targets Data Elements 2015	Code List	Code List

Table 1: *2015 MER Data Elements Code Lists*

If you unsure about which version of the code lists to use for your data, contact the DATIM Help Desk for further guidance.

2.2.3.2. Verify Implementing Mechanism (IM) Codes

The source for DATIM implementing mechanism is *FACTS Info IMs*. Data from FACTS is regularly synchronized with DATIM. All mechanisms which are present in a given data import file must be active for the period for which data is being imported. Once a mechanism is "promoted" or made active in FACTS, it will take up to 24 hours for the data to be updated in DATIM. Please ensure that all mechanisms for which you will be reporting data for are active in FACTS prior to attempting to test any data in DEV-DE. If any

mechanisms are not active or reported against mechanisms which the user does not have access to, conflicts will be reported by the server.

The *Agreement End Date* attribute in FACTS Info may inhibit a mechanism being visible or available in DATIM. The date must be valid for data to be entered appropriately. *NOTE: IPs DO NOT have access to this system. Funding agencies should provide the Mechanism ID to partners interested in data exchange.*

A nightly synchronization between FACTS Info and DATIM systems in which mechanism information from FACTS Info updates in to DATIM takes place. Country Teams should be diligent in maintaining FACTS Info as an updated mechanism resource for DATIM. Each code required for data exchange is obtained from a system of record seen in Table 2.

Code	System of Record
Implementing Mechanism Code	FACTS Info via DATIM
Site ID	DATIM
Indicator ID	DATIM
Dataset Code	DATIM
Reporting Period	DATIM

Table 2: Codes and their corresponding System of Records

2.2.3.3. Periods

Each data value in DATIM must have a valid period. Results are currently reported on a quarterly basis, targets once a financial year. SIMS data should be reported as monthly data. All periods refer to calendar periods. Thus, Jan-March 2015 should be referred to as "2015Q1". A complete reference of how periods should be specified is available here.

2.2.3.4. Data Validation Rules

Data validation rules define how data values of different data elements should relate to each other. As a very simple example, the number of HTC tests should not exceed the number of positive HTC tests. Validation rules are evaluated for a given organisation unit – period – attribute option combination. Thus, data is not validated for two different periods or for two different mechanisms. Data validation rules provide additional restrictions and quality checks on the data which can be imported into DATIM. Users should be exceedingly careful in choosing the correct period for validation and restrict themselves to a relatively small number of organisation units to analyze data for at the same time. If one attempts to validate too many sites over multiple time periods, it may result in a very long process.

In addition to validating the data on the server, the DATIM development team has created a set of scripts in the R programming language which can be used to validate data completely externally of the DEV-DE environment. Complete installation and usage instructions on the use of this package are available here. Note that the validation scripts will validate the data for data elements which have an option set and if the value provided does not match a valid option, the data will be considered to be invalid.

In summary, once data has been successfully imported into the DEV-DE environment, it should be subjected to validation rule violation testing. Large number of validation rule violations (when compared to the overall total amount of data) indicate there could be additional problems in the way which the data import file has been prepared. These violations should be reviewed and resolved prior to sending the file to the DATIM help desk for import into the production system.

2.2.4. Submit Final and Tested Data

Once testing is complete, the submission process begins. Follow the below steps to submit the validated file:

- Prepare import, and submit the file to the DATIM Technical Support team for review
- Include the format of the file (XML, JSON, CSV, ADX)
- Include the coding scheme which has been used (UIDs or codes)
- Include an output from the DEV-DE environment, which shows that the data was able to be imported with no conflicts reported from the server.

- (Optional) Include any reports of validation rule violations. If validation rule violations should be ignored, please include a statement of justification why.

DATIM Technical Support team will validate your data import file and data to determine next steps. If updates or changes are needed, the submitter will be responsible. When all necessary steps have been taken and verified, the Technical Support team will facilitate the import into DATIM.

2.2.5.   Move Imported Data Through DATIM Approval Workflow

Data imported into DATIM must still follow the required approval work-flow. When data is imported into the production system, it will be in the "Pending" state. Implementing partners should submit their data to the agency, which can then approve and submit the data to inter-agency space. The DATIM Help Desk can also help you to move the data which has been imported in bulk up the approval chain. If requested, this operation should be supported with approval from the PCO, authorizing the DATIM help desk to approve the data in bulk. Clear instructions should be provided to which level the data should be submitted.

2.2.6.   Data Deduplication

Once data has been imported and approved to inter-agency level, if there are any duplicates, they will also need to be resolved by the PEPFAR country team. In terms of planning, it is important to consider that the de-duplication process can be very lengthy depending on the number of duplicates.

2.2.7.   Data Import Cleaning Guidance

Please read the below guidance for each of the data import cleaning options. Note that you are required to submit separate set (s) of files for each quarter, preferably in two separate helpdesk tickets.

2.2.7.1.   Preferred Approach for Data Updates: Submission of New/Updates + Delete File
o   If you are submitting updated data and/or new data, you should submit this as a single file. Clearly indicate in the name of the file that this is intended for updates. Use the following naming convention for this file: PEPFAR_Q1_DataImport_Update_06232016.

o   If you intended to delete data, and know what data it is which you wish to delete, you should prepare another file with the data values which should be deleted, clearly indicating in the file name that this data should be deleted. Use the following naming convention for this file: PEPFAR_Q1_DataImport_DeleteData_06232016.

If you are able to submit data via this option and the data has been validated at your end, prior to submission, with the DATIM validation scripts, DATIM Engineering anticipate that this operation should be able to be completed within 48 hours of submission.

2.2.7.2.   Bulk Removal of Data
If you are unable to provide a file which contains values which should be completely deleted from the database, or if you intended to submit a complete replacement of your data, you need to request at least three weeks in advance of the submission of your completed data file, that DATIM Engineering should remove your data from the database. Use the following naming convention for this file: PEPFAR_Q1_DataImport_BulkDataDelete_06232016.

This type of operation requires a custom SQL script which will operate directly against the database, and because of this, requires both development time, a security review, as well as testing. Even if your data is not ready, you should indicate your intent to remove the data as early as possible, **but not later than three weeks prior to the re-submission of your replacement data**.

2.2.7.3.   Selected Removal of Data by DATIM Engineering
In general, this operation is not supported by DATIM Engineering. For instance, you might request the selective removal of a certain set of data elements or values for certain mechanisms or organization units. In general, this operation will not be undertaken by DATIM Engineering, except in very exceptional cases.

2.2.7.4. Data Import Cleaning Option Summary Table

#	Task Option	Description	Required Lead Time for Engineering Team
1	Submission of new/ updates + delete file	• If you are submitting updated data and/or new data, you should submit this as a single file. Clearly indicate in the name of the file that this is intended for updates. • If you intended to delete data, and know what data it is which you wish to delete, you should prepare another file with the data values which should be deleted, clearly indicating in the file name that this data should be deleted.	• If data has been validated at your end, prior to submission, with the DATIM validation scripts, DATIM Engineering team anticipate that this operation should be completed within 48 hours of submission.
2	Bulk removal of data	• If you intended to submit a complete replacement of your data, you need to request at least three weeks in advance of the submission of your completed data file, that DATIM Engineering should remove your data from the database.	• You are required to submit request at least three weeks in advance of the submission of your completed data file.
3	Selected removal of data by DATIM Engineering	• In general, this operation is not supported by DATIM Engineering team.	Not Applicable.

3. FREQUENTLY ASKED QUESTIONS

Unsure of something? Reference this section for answers

API FAQs
When is it acceptable to use the API to exchange data?
The API capability is always available; however, OGAC requests that, for now, API ONLY be used in the DATIM TEST environment because meta-data can't be synced through this method, and it limits the ability to accurately and completely exchange data. It is an update that is being fully developed. Changes to this policy will be communicated.
The API source file is too large to attach. How else can I submit the file to DATIM Helpdesk?
Request access to the DATIM secure FTP site by submitting a helpdesk ticket. A member of the Technical Support team will assist in obtaining valid credentials for access to the site. Additionally, be sure to compress the file (ZIP) before you send it.
Is there any way to validate my data externally of the DEV-DE environment?
Yes, the script is written in the open source statistical programming language, R. Install the R programming environment and run the validation script there. Details of how to use the scripts are available here . All validation rules are available through the meta-data API of DATIM as well. If interested in developing unique data validation procedures based on the data validation rules of DATIM, contact the data exchange team for technical guidance.
How do we update data previously added by the API?
OGAC requests that updates to previously submitted data not use API. For instance, COP targets submitted before the COP review may now require changes (after the COP review). Data import or manual data entry should be considered for these changes.

Data Import Reference Guide

Reporting FAQs
Are all Monitoring, Evaluation, and Reporting (MER) indicators available in DATIM TEST?
Yes. Please refer to the *PEPFAR Monitoring, Evaluation, and Reporting (MER) Guidance* released in March 2015 for all indicators required for reporting. MER Indicators are stored as Data Elements in DATIM while the indicator disaggregates are stored as *CategoryOptionCombo*. Submit BOTH as a pair in file uploads. You should also double-check the data entry screens of DATIM for the absolutely latest updates to any of the indicators or disaggregations.
There are sites without IDs in our country system, how can we update and use them in DATIM?
Each country contains Site Administrators who can add and edit sites. Refer to Section 3.3 on Site Administration from the *DATIM Compendium for PEPFAR Country / Regional Operational Plan (COP/ROP) 2015* for more details.
How should I refer to periods in my data import file?
DATIM uses the DHIS2 internal date format for designating reporting periods. All periods in DATIM data exchange files must provide a valid DHIS2 period identifier. In particular, DATIM uses Financial Year October for COP targets, Quarters for MER Results, and Months for SIMS reporting (see the table below). A full discussion of DHIS2 data formats can be found here.

Interval	Description	Format	Example
Day	March 15, 2016	*yyyyMMdd*	20160315
Week	Week 10 2016	*yyyyWn*	2016W10
Month	March 2016	*yyyyMM*	201603
Quarter	January-March 2016	*yyyyQn*	2016Q1
Six-month	January-June 2016	*yyyySn*	2016S1
Six-month April	April-September 2016	*yyyyAprilSn*	2016AprilS1
Year	2016	yyyy	2016
Financial Year April	Apr 2016-Mar 2016	yyyyApril	2016April
Financial Year July	July 2016-June 2016	yyyyJuly	2015July
Financial Year Oct	Oct 2016-Sep 2016	yyyyOct	2016Oct

What happens if we are unable to import data in DATIM during this (current) reporting cycle?
Each country team that is planning for data import still needs to have a contingency plan to ensure data is available in DATIM within the set deadlines. Manual data entry by IPs and / or USG staff is the only contingency until a data exchange approach is in place. Furthermore, plans for data import should be based on available resources, not on planning or reporting deadlines. Plan realistically.

Validation FAQs
What is the process for external validation checks?
The DATIM development team has developed a script that will report findings from the following:

- Valid data set assignments for organization units, e.g. so that MER facility data is not imported into community level
- Checks for active mechanisms
- Invalid data element / disaggregate combinations based on the current data entry form by dataset
- Duplicated data (OU / DE / DISAGG / PERIOD)
- Checks against DATIM validation rules and provides a report of violations

Do logic checks run across periods or within a single period?
Logic checks only validate data within a single period. Users must still ensure that all reported data from previous periods are consistent with the current reporting period, and the overall program plans or results.

Miscellaneous FAQs
Can I or other members in my OU import data into DATIM?
No, users can only import in DATIM DEV-DE. Only members DATIM Technical Support have the ability to import data in to DATIM.
How can we get additional documentation for data exchange processes?
Join the OHIE HMIS community / DATIM sub-community to fully participate in the B2B meta-data sync process.

Data Import Reference Guide

Miscellaneous FAQs
Can we use Excel templates developed by USG for data collection to upload data into DATIM?
An Excel template can be used; however, the template needs to be in a DHIS2 CSV compliant format. A sample format from the DATIM Support page is available for download found [here](#).