

Differential data upload

Differential upload user guide

The 1st stage differential upload is to upload a .gpg file to add, update or delete records. The existing UPDATE_MODE and STATUS fields are used to give information to do differential upload.

- [1. Organizations allowed for DIFF upload](#)
- [2. How to create the .gpg file](#)
 - [When there is more than one record](#)
 - [When there is no record](#)
- [3.Expected behavior](#)
- [4. Business validation rules](#)
- [5. Errors and warnings in the processing report](#)
 - [Report](#)
 - [New Record Errors/warnings:](#)
- [6. Organizations use DIFF upload](#)

1. Organizations allowed for DIFF upload

- The registries that upload donors **with GRID. GRID is compulsory for DIFF.**
- All registries and CBBs upload CBUs

WMDA suggests all organizations with dataset more than 100K consider to implement DIFF upload to save effort to time.

2. How to create the .gpg file

When there is more than one record

- Following the same process in the FULL upload user guide [Search & Match Service Data Submission Information](#) to prepare the differential upload .gpg file.
- The DIFF upload needs include all the fields for a DONOR/CBU exactly like the FULL upload. For no records situation for DIFF file, please check guide below.
- Changes comparing to current FULL upload (Will later be updated in the FULL upload user guide)

Field Identifier	Required	Description	Type	Length	Comment
UPDATE_MODE	Yes	Update mode of the inventory, i.e. FULL or DIFF	updateMode Type	4	"FULL" means for full upload "DIFF" means differential upload
STATUS	Yes	Status of the donor/CBU	statusType	2	statusType: "AV", "TU", "RS", "DE" * AV = Available for transplantation purposes TU = Temporarily unavailable RS = Reserved DE = Deleted, permanently unavailable* *DE is only supported in DIFF upload, will be rejected in FULL upload.

When there is no record

Accepted structure

DIFF file generation normally run automatedly, and it is possible there is no record generated. We suggest our registry do send the updated for this empty file case. But the empty record file must use the structure as below to not include the <INVENTORY> element. Otherwise the data process will generate error and block new files from your registry to be processed.

XML example as below:

XML example for file with empty record

```
<?xml version="1.0" encoding="UTF-8"?>
<INVENTORIES CREATION_TIME="2022-03-07T10:00:03Z">
</INVENTORIES>
```

Not accepted structure

Structure below will cause issue in WMDA data process, we added a ticket to add this feature to future release, but not supported it yet.

XML issue structure with empty record

```
<?xml version="1.0" encoding="UTF-8"?>
<INVENTORIES CREATION_TIME="2022-07-15T07:49:26Z">
  <INVENTORY LISTING_ORGANIZATION="0999" POOL="0999" CONTENT_TYPE="C" UPDATE_MODE="DIFF" SNAPSHOT_TIME="2022-07-15T07:49:26Z" SCHEMA_VERSION="2.3" />
</INVENTORIES>
```

3.Expected behavior

Differential Update cases	File level	DONOR Behavior (GRID is compulsory)	CBU Behavior
Upload frequency	Upload frequency must be more than 15mins for each DIFF upload. Otherwise, only one file in each 15mins will be processed.		
Add new records when STATUS is "AV" or "TU" or "RS".		<ul style="list-style-type: none">Records with duplicated ID in the upload file will be rejected.Records with duplicated GRID in the upload file will be rejected.Records without GRID will be rejected. But there is no rejection message, under implementation in next release.Invalid records will be rejected.Valid records with not existing GRID will be added	<ul style="list-style-type: none">Invalid records will be rejected.Valid records with not existing ID will be added
Update existing records when STATUS is "AV" or "TU" or "RS"		<ul style="list-style-type: none">Records with duplicated ID in the upload file will be rejected.Records with duplicated GRID in the upload file will be rejected.Invalid records will be rejected, and existing record with same ID and GRID in database will be deletedValid records with existing GRID will be updated	<ul style="list-style-type: none">Records with duplicated ID in the upload file will be rejected.Invalid records will be rejected, and existing record with same ID and GRID will be deleted in the database.Valid records with existing ID in DB will be updated
Deleted records when STATUS is "DE"		<ul style="list-style-type: none">Record exists in database will be deleted.Record does not exist in the database will be ignored. And no warning message in the report.	<ul style="list-style-type: none">Record exists in database will be deleted.Record does not exist in the database will be ignored. And no warning message in the report.
Upload records threshold limitation	Less than 200K for each DIFF upload. (Can request for more than 200K DIFF upload if needed, please contact support@wmda.info)		

4. Business validation rules

Reference number	Validation level	Validation Source	Date rule is valid	Effective since XSD version number	Field Name	Error in field	Reported Validation message	Action
9	File			2.1	UPDATE_MODE	Invalid update mode	Your file has been rejected as the UPDATE_MODE must be equal to "FULL" or "DIFF".	Reject file
222	File	WMDA		2.1	UPDATE_MODE	Invalid update mode for multiple inventories	Your file has been rejected as we have identified mixed update modes in your XML inventories.	Reject file
	File	WMDA		2.1	N/A	Too many record provided for DIFF upload	File could not be processed due to DIFF upload exceeded record count threshold: 600K	Reject file

5. Errors and warnings in the processing report

Report

Example report of Differential upload
<p>File: ION-0999-D.gpg 2019-09-09 11:39:22 Pool(s): 0999 Content Type: D Update Mode: DIFF Start processing: 2019-09-09 11:50:00 Schema version: 2.1 Total records processed: 14 Total records with warnings: 6 Total records rejected: 4 Total valid records: 7 Total updated records: 4 Total new records: 4 Total deleted records: 2</p> <p>List of Records with duplicated ID or GRID:</p> <p>POOL: 0999 IDs: TD-000004 GRIDs: 1234000000000203420</p> <p>W 0999 TD-000002 N/A (Warning) GRID 774800006001853603 must be 19 characters. R 0999 TD-000002 N/A GRID (Record Rejected) GRID is a mandatory field. W 0999 TD-000003 N/A (Warning) GRID 7748000060018E53612 checksum is not correct. R 0999 TD-000003 N/A GRID (Record Rejected) GRID is a mandatory field. W 0999 TD-000010 1234000000000001031 STAT_END_DATE (Warning) Status end date cannot be > 5yrs in the future. W 0999 TD-000010 1234000000000001031 STAT_REASON (Warning) Status reason cannot be provided with status AV or RS. R 0999 TD-000013 999900000TD00001129 BIRTH_DATE (Record Rejected) BIRTH_DATE is a mandatory field. W 0999 TD-000014 N/A (Warning) GRID must be 19 characters. R 0999 TD-000014 N/A GRID (Record Rejected) GRID is a mandatory field.</p> <p>Processing finished at: 2019-09-09 11:53:55 Total processing time: 0 minutes.</p>

The processing report for differential upload has 3 parts.

- 1st part is the summary of the file upload, with more details of updated, new, and deleted records.
- 2nd part is the duplication in the uploaded file, data uploader should clean up the duplication.
- 3rd part is the rejections and warnings. Data uploader should **clean up** all of them.

Summary numbers:

We define some cases:

Cases	
DE non exist, ignored	This is the records with "DE" status, but they do not exist yet in WMDA database. This will be ignored. As it is resource consuming to calculate the number based on the test implementation, WMDA decide to ignore to calculate it.
GRID Missing, rejected	GRID is compulsory for DIFF and FULL from Dec, 17, 2019. And the messages include Warning (W) type, and also Rejection (R) type. Example: W 0999 TD-000014 N/A (Warning) GRID must be 19 characters. R 0999 TD-000014 N/A GRID (Record Rejected) GRID is a mandatory field.
Calculation	Total records = Total records rejected + Total valid records + Total deleted records + DE non exist Valid records = updated records + new records + no change exist records Rejected records = Rejected records + DE records missing required fields (not listed in the R/W details)

New Record Errors/warnings:

We do not have new in DIFF upload yet.

Type	Message	Description

6. Organizations use DIFF upload

ION	Date	
ION-3553	2020-10-30 CBU, 2021-02-17 Donor	
ION-4596 includes following IONs: ION-1574, 5525,6738,7414, 9935, 9968	2020-10-08 and later for rest of the IONs that 4596 is in charge of	
ION-5103	2022-08-08	
ION-8139	2019-11-01	
ION-8486	2023-10-16	
ION-8766	2023-02-15	