# 1. Introduction

## 1.1 Context

There is a desire by UNAIDS to encourage countries that have implemented DHIS2 to generate indicators that use Spectrum[[1]](#footnote-1) data.

It has been successfully demonstrated that the extracts generated from Spectrum can be manually manipulated into a format that can be imported into DHIS2. Furthermore, an analysis of the Spectrum extracts revealed that certain steps in the manual manipulation process could be eliminated if minor adjustments were made to the extract format.

This document describes the requirements for implementing a new ‘export to DHIS2’ function in Spectrum should such a capability be desired.

## 1.2 Simplifying the Spectrum to DHIS2 import process

Of the several types of Spectrum extracts, the list format is the one that most closely resembles the format that DHIS2 expects. The following manipulation tasks were undertaken to prepare this extract for import to DHIS2:

* Remove ‘total’ rows for disaggregated data
* Replace the labels of data elements with their equivalent DHIS2 UIDs[[2]](#footnote-2)
* Insert new DHIS2 UIDs representing the level of disaggregation in the Spectrum data (gender/age-bracket UIDs)
* Replace the labels representing the organizational hierarchy (State/District etc.) with their equivalent DHIS2 UIDs.
* Change the order of the columns.

It should be possible to eliminate two to three of the above manipulation steps and semi-automate another through some minor changes to the export format.

## 1.3 What this document contains

This document provides the following:

* Narrative explanation of the changes required to the extract format
* A high-level list of requirements
* Appendices providing samples of the original and manipulated versions of the extract to demonstrate the requirements

It also arranges the changes in likely order of development complexity, so that Spectrum developers can rapidly prototype simpler versions of the DHIS2 compatible extract and assess the impact/overhead of developing the more complex changes.

# 2. The Narrative View

## 2.1 Spectrum-DHIS2 Background

DHIS2 has three core concepts around which it organizes its aggregate data collection: data elements, indicators and organizational hierarchies. The organizational hierarchy is typically used to breakdown a country into administrative areas, sub-areas and facilities, matched to the needs of the implementer (e.g. a Ministry of Health). Data elements are the atomic level items for holding data, while indicators are calculated items derived from data elements (e.g. prevalence for a certain disease).

Data elements and indicators are linked to organizational unit(s) and data is typically collected at the lowest level possible, and aggregated by the system to generate higher organizational level values for data elements and indicators. A data element may also be defined in such a way that it can be disaggregated into relevant sub-component parts (e.g. by gender and age-brackets), in which case, these sub-components would be loaded and the overall total for the given data element calculated automatically from these sub-components.

Each object in DHIS2 (e.g. data element, user, indicator, disaggregation category, organization unit etc.) is allocated a GUID that is unique across all instances of DHIS2. When importing data into DHIS2, GUIDs can be used to precisely identify which objects a specific data item will be linked to.

## 2.2 Desired Extract Changes

### 2.2.1 Order of Columns

The order of columns that DHIS2 expects is different than the list format extract file that is generated by DHIS2. This is likely the easiest change to implement.

In the sample extract below, the original order of the columns was:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| File name | Subnational region | Country | Indicator | Year | Estimate |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Male+Female | 2014 | 373231 |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Male+Female | 2015 | 374219 |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Male+Female | 2016 | 377019 |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Male | 2014 | 223883 |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Male | 2015 | 223923 |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Male | 2016 | 225170 |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Female | 2014 | 149348 |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Female | 2015 | 150296 |
| ..\abc.PJN | Andhra Pradesh | India | HIV population (15+) Total Region; Female | 2016 | 151850 |

The DHIS2 ‘import ready’ version of this file (in terms of column order) is as follows (several further columns are not displayed in this sample, including a **‘comment’ column that should include an ‘estimates as of *datevalue’* entry.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dataelement | period | orgunit | Categoryoption  combo | Attributeoption  combo | Value |
| HIV population (15+) Total Region; Male+Female | 2014 | India/Andhra Pradesh |  |  | 373231 |
| HIV population (15+) Total Region; Male+Female | 2015 | India/Andhra Pradesh |  |  | 374219 |
| HIV population (15+) Total Region; Male+Female | 2016 | India/Andhra Pradesh |  |  | 377019 |
| HIV population (15+) Total Region; Male | 2014 | India/Andhra Pradesh |  |  | 223883 |
| HIV population (15+) Total Region; Male | 2015 | India/Andhra Pradesh |  |  | 223923 |
| HIV population (15+) Total Region; Male | 2016 | India/Andhra Pradesh |  |  | 225170 |
| HIV population (15+) Total Region; Female | 2014 | India/Andhra Pradesh |  |  | 149348 |
| HIV population (15+) Total Region; Female | 2015 | India/Andhra Pradesh |  |  | 150296 |
| HIV population (15+) Total Region; Female | 2016 | India/Andhra Pradesh |  |  | 151850 |

Notes:

* The ‘File name’ column is eliminated
* The ‘Year’ and ‘Estimate’ columns become the ‘Period’ and ‘Value’ columns respectively (in a different order).
* The ‘Indicator’ column becomes the ’dataelement’ column and the concept of Indicator is represented as ‘dataelement’ and ‘Categoyoptioncombo’ in DHIS2
  + The ‘dataelement’ in the import-ready version would be replaced with the UID code that represents the object ‘HIV Population’
  + The ‘Categoryoptioncombo’ in the import-ready version would be replaced with the UID code that represents the disaggregation object ‘15+ Male’ and ‘15+ Female’ respectively.
* The ‘country’ and ‘subnational region’ columns are replaced by a single ‘orgunit’ column. The ‘orgunit’ in the import-ready version would be replaced with the UID code that represents the object ‘India/Andhra Pradesh’.

### 2.2.2 Remove Totals Rows

In DHIS2, ‘Total’ values are generated from their disaggregated components, and it is therefore not necessary to generate Totals in the extract. Using the previous sample from above, the ‘import-ready’ version of this file (in terms of column-order and removal of total options’) would look similar to:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dataelement | period | orgunit | Categoryoption  combo | Attributeoption  combo | Value |
| HIV population (15+) Total Region; Male | 2014 | India/Andhra Pradesh |  |  | 223883 |
| HIV population (15+) Total Region; Male | 2015 | India/Andhra Pradesh |  |  | 223923 |
| HIV population (15+) Total Region; Male | 2016 | India/Andhra Pradesh |  |  | 225170 |
| HIV population (15+) Total Region; Female | 2014 | India/Andhra Pradesh |  |  | 149348 |
| HIV population (15+) Total Region; Female | 2015 | India/Andhra Pradesh |  |  | 150296 |
| HIV population (15+) Total Region; Female | 2016 | India/Andhra Pradesh |  |  | 151850 |

### 2.2.3 Replace extract labels with DHIS2 UNAIDS-SPECTRUM GUIDS

To facilitate the import of Spectrum data, a set of standard Spectrum data elements (and GUIDS) have been defined that will be instantiated by a special DHIS2 UNAIDS-SPECTRUM BOOTSTRAPPING App.[[3]](#footnote-3) Given that the App will instantiate the same object GUIDs across any DHIS2 instance, these identifiers (if they form part of the Spectrum export) will eliminate one of the manual steps previously highlighted.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dataelement | period | orgunit | Categoryoption  combo | Attributeoption  combo | Value |
| G5XS6gqIrIX | 2014 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 223883 |
| G5XS6gqIrIX | 2015 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 223923 |
| G5XS6gqIrIX | 2016 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 225170 |
| G5XS6gqIrIX | 2014 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 149348 |
| G5XS6gqIrIX | 2015 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 150296 |
| G5XS6gqIrIX | 2016 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 151850 |

Notes:

* **This example Spectrum output represents an ideal version that eliminates several manual steps of preparation.**
* The ‘dataelement’ value is replaced by the DHIS2 GUID that represents the Spectrum Estimate ‘HIV Population’ in DHIS2 (which will be universal to all DHIS2 instances that implement this feature).
* Since it will still be necessary to manually set the Categoryoptioncombo GUIDs to correctly identify the ‘age/gender’ disaggregation’s for import into DHIS2, then the original ‘Indicator’ label should be entered in this column to enable users to identify the indicator and its disaggregation’s.

For reference, the final ‘import-ready’ version of this sample is provided below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dataelement | period | orgunit | Categoryoptioncombo | Attributeoptioncombo | Value |
| G5XS6gqIrIX | 2014 | N2bI9Abo0LM | NWzbX92mJ6L |  | 223883 |
| G5XS6gqIrIX | 2015 | N2bI9Abo0LM | NWzbX92mJ6L |  | 223923 |
| G5XS6gqIrIX | 2016 | N2bI9Abo0LM | NWzbX92mJ6L |  | 225170 |
| G5XS6gqIrIX | 2014 | N2bI9Abo0LM | ZNZZWddI4QL |  | 149348 |
| G5XS6gqIrIX | 2015 | N2bI9Abo0LM | ZNZZWddI4QL |  | 150296 |
| G5XS6gqIrIX | 2016 | N2bI9Abo0LM | ZNZZWddI4QL |  | 151850 |

Notes:

* The final manual steps entail identifying the **locally unique** DHIS2 GUIDS that represent the subnational area and age/gender disaggregations and inserting them into the appropriate cells in the extract file.

## 2.3 The Spectrum DHIS2 Ready Export Capability

The following narrative has been split between ‘should haves’ and the ‘could haves’. The ‘should haves’ define a preferred minimum requirement if such a capability were to be implemented, while the more complex requirements have been defined as ‘could haves’.

### 2.3.1 The Should Haves

The generation of a DHIS2 ready extract should at a minimum generate an output similar to:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dataelement | period | orgunit | Categoryoption  combo | Attributeoption  combo | Value |
| HIV population (15+) Total Region; Male | 2014 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 223883 |
| HIV population (15+) Total Region; Male | 2015 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 223923 |
| HIV population (15+) Total Region; Male | 2016 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 225170 |
| HIV population (15+) Total Region; Female | 2014 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 149348 |
| HIV population (15+) Total Region; Female | 2015 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 150296 |
| HIV population (15+) Total Region; Female | 2016 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 151850 |

.

This is the application of the changes described in 2.2.1 (Order of Columns) and 2.2.2 (Removal of Totals). It also ‘repeats’ the ‘indicator label’ in one of the new labels to facilitate manual manipulation in preparation of the extract for import into DHIS2. Not displayed is a **comment column that should include an ‘estimates as of *datevalue’* entry** (see Appendix 1 for the column number).

### 2.3.2 The Could Haves

The addition of a universal set of UIDs to represent the indicator in the ‘dataelement’ column (as described in 2.2.3) would eliminate a significant amount of manual manipulation. This would require the storage of these UIDs somewhere in the Spectrum system. If this ‘could have’ were implemented, the resulting output would look similar to:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dataelement | period | orgunit | Categoryoption  combo | Attributeoption  combo | Value |
| G5XS6gqIrIX | 2014 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 223883 |
| G5XS6gqIrIX | 2015 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 223923 |
| G5XS6gqIrIX | 2016 | India/Andhra Pradesh | HIV population (15+) Total Region; Male |  | 225170 |
| G5XS6gqIrIX | 2014 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 149348 |
| G5XS6gqIrIX | 2015 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 150296 |
| G5XS6gqIrIX | 2016 | India/Andhra Pradesh | HIV population (15+) Total Region; Female |  | 151850 |

The full set of columns and the indicator UIDs are provided in the Appendices.

# 3. High-level Requirements

The following provides a more formal representation of the requirements of the DHIS2 compatible Spectrum Export capability.

### REQ-1: The system should be able to extract a DHIS2 compatible import file of UNAIDS data elements

This is THE key high-level requirement. The level of compatibility is dependent on the number of ‘features’ from the Change items described in 2.2.1, 2.2.2 and 2.2.3 are implemented.

|  |  |  |
| --- | --- | --- |
| **Details** | |  |
| **Parent**: Spectrum System | | |
| **Type:** Functional | | **Status:** |
| **Assigned To:** Spectrum Developers | | **Release:** |
| **Notes** | | **Date Added** |
| The ‘list’ extract format is the one that most closely resembles that DHIS2 import format. | | 01/06/2015 |
| **Use cases that reference this requirement** | | |
|  |  | |

#### REQ-1.1: Change the order of the columns

The DHIS2 column order is listed in Appendix 1. The changes are described in 2.2.1, and a sample layout is provided.

|  |  |  |
| --- | --- | --- |
| **Details** | |  |
| **Parent**: REQ-1: The system must be able to extract a DHIS2 compatible import file of UNAIDS data elements | | |
| **Type:** Functional | | **Status:** |
| **Assigned To:** Spectrum Developers | | **Release:** |
| **Notes** | | **Date Added** |
| Changing the order of the columns, merging a couple of them, and inserting some new blank columns is the minimum level of compatibility that would eliminate several manual steps. | | 01/06/2015 |
| **Use cases that reference this requirement** | | |
|  |  | |

#### REQ-1.2: Remove Total Rows

It is unnecessary to include indicator totals in the extract, as the DHIS2 import will focus on the disaggregate values (and total them internally). This change in the extract is described in 2.2.2, and a sample layout is provided.

|  |  |  |
| --- | --- | --- |
| **Details** | |  |
| **Parent**: REQ-1: The system must be able to extract a DHIS2 compatible import file of UNAIDS data elements | | |
| **Type:** Functional | | **Status:** |
| **Assigned To:** Spectrum Developers | | **Release:** |
| **Notes** | | **Date Added** |
| This applies to any totals that are a result of summing more detailed disaggregated values. Only those values that are at the lowest level of disaggregation for the ‘sub-national’ level of interest should be generated (e.g. sub-national, gender and age-bracket combinations as appropriate. | | 01/06/2015 |
| **Use cases that reference this requirement** | | |
|  |  | |

#### REQ-1.3: Replace Indicator label with DHIS2 UID

It is expected that this will be the most complex compatibility change to implement. How the ‘indicators or interest’ are labeled in Spectrum are not necessarily consistent across instances, yet the target indicators for this DHIS2 compatible extract function will be consistently identified with UIDs that apply to all instances of DHIS2 that run the UNAIDS-Spectrum bootstrapping App. The label of the indicator could be changed to match the DHIS2 UID for that indicator. An attribute-value list is provided in Appendix 2.

|  |  |
| --- | --- |
| **Details** |  |
| **Parent**: REQ-1: The system must be able to extract a DHIS2 compatible import file of UNAIDS data elements | |
| **Type:** Functional | **Status:** |
| **Assigned To:** Spectrum Developers | **Release:** |
| **Notes** | **Date Added** |
| How this function is implemented should take into account the need for ‘new additions’ to the initial list of UIDs | 01/06/2015 |

### REQ-2: Specify the parameters of the DHIS2 compatible extract

It will be necessary for the user to match/select the target indicators for extraction with those for which UIDs exist. Appendix 2 provides an initial list of the indicators that are targeted for import into DHIS2. Given that these may be labeled differently in each instance of Spectrum then some sort of ‘matching’ step may need to be part of this extract function.

|  |  |  |
| --- | --- | --- |
| **Details** | |  |
| **Parent**: Spectrum System | | |
| **Type:** Functional | | **Status:** |
| **Assigned To:** Spectrum Developers | | **Release:** |
| **Notes** | | **Date Added** |
| The Spectrum developer will need to consider whether to retain the selections of the user for the next time the function is run. | | 01/06/2015 |
| **Use cases that reference this requirement** | | |
|  |  | |

# Appendix 1 – DHIS2 Import File Column Order

The following table indicates the column format that DHIS2 expects and how the original Spectrum list extract format maps to these values.

|  |  |  |
| --- | --- | --- |
| Column Number | Column Name | Original List Format |
| 1 | dataelement | Indicator |
| 2 | period | Year |
| 3 | orgunit | Country / Sub National |
| 4 | categoryoptioncombo | Indicator |
| 5 | attributeoptioncombo |  |
| 6 | Value | Estimate |
| 7 | storedby |  |
| 8 | timestamp |  |
| 9 | comment | “Estimate as on” *datevalue* |
| 10 | followup |  |

# Appendix 2 – Target Indicator / UID Mappings

The following table provides a list of the target indicators for extract and the UID mappings if the indicators are to be replaced in the output extract. This informs both REQ 1.3 and REQ 2 (indicators to match).

|  |  |  |
| --- | --- | --- |
| Target Spectrum Indicator | Extract Parameters | DHIS2 Data Element UID |
| Estimated number of people living with HIV | Sub-national  <15 M/F  15+ M/F  15-49 M/F | To come |
| Estimated number of people eligible for ART according to national guidelines | Sub-national  <15 M/F  15+ M/F  15-49 M/F | To come |
| Estimated number of HIV+ pregnant women | Sub-national | Fttd76PfAsf |

Estimated number of HIV+ pregnant women

Estimated number of people living with HIV

1. Spectrum – the software package used to model HIV/AIDS numbers [↑](#footnote-ref-1)
2. Universally unique ID’s that DHIS2 internally generates for all objects. [↑](#footnote-ref-2)
3. DHIS2 has the equivalent of an App Store that allow the creation of Apps. [↑](#footnote-ref-3)