**Data Catalog**

**Metadata Guide**

***Ministry of Health, eHealth Africa***

***April 2018***

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Author** | **Date** | **Note** |
| 0.1 | Qing Gong | 2018.01.05 | draft |
| 0.2 | Qing Gong | 2018.04.02 | Updated with inspirations from training sessions, new Bika metadata, metadata for excel tools |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Table of Content

[Metadata of Software Applications](#_blnkiu1sb9u6)

[CBIS (DHIS2)](#_eu3tqxnc7472)

[HMIS & Offline Tracker (DHIS2)](#_jm82tcr3xtuw)

[LIMS/Bika](#_ambsr91fj9ax)

[iHRIS](#_otycb6jddut6)

[Metadata of Excel Tools](#_463x8f5vozsw)

[Tools](#_z6jzk54pddnq)

[SPSS](#_pl6hmho7hhjk)

[Notepad++](#_y08lwlnbm0kb)

[Python](#_kvbo82mvg9ap)

[Simplejson](#_ua1ld07iil5h)

[Python Scripts](#_rjmeqcik1x1t)

## Metadata of Software Applications

### CBIS (DHIS2)

Since CBIS is implemented in DHIS2 (hosted by MoH), the DHIS2 metadata will be used to extract CBIS related metadata. The following steps can be followed to get the CBIS metadata.

* Export DHIS2 metadata in json format.
* Run the code CBIS\_dataElements.py, to extract the *dataElements* of CBIS related metadata. The output of this step is CBIS related metadata presented in HTML format.
  + Startup button -> enter “cmd” in the search bar -> press “enter”
  + Navigate to the directory where CBIS\_dataElements.py is located (command to use “cd ..” and “cd path\_to\_the\_directory)
  + Run the code: type CBIS\_dataElements.py -> press “enter”

Note: another way to run the CBIS\_dataElements.py code could be to double click on the file.

* Follow the same steps as described above to run the code CBIS\_indicators.py, to extract the *indicators* of CBIS related metadata.
* Add CSS style to the HTML file by adding the code in CSS-styleCode.html in front of the HTML code in CBIS\_dataElements.html. (open the files with Notepad++)
* Add CSS style to the HTML file by adding the code in CSS-styleCode.html in front of the HTML code in CBIS\_indicators.html. (open the files with Notepad++)

The pre-condition of this method is that all CBIS related metadata has the keyword “CBIS” in its name.

### HMIS & Offline Tracker (DHIS2)

Both HMIS and eIDSR Offline Tracker are implemented with DHIS2, but in two different instances. HMIS is hosted by MoH, and eIDSR Offline Tracker is hosted in eHealth Africa, as for the time when this document is written.

Since the amount of elements in the metadata is big, we will separately present each element of the metadata. To do so, the following steps can be followed:

* Export DHIS2 metadata in json format.
* View and restructure the json file in a JSON editor (e.g. <http://jsoneditoronline.org/>).
  + Copy and paste the metadata in json format to the window to the left
  + Click the arrow pointing to the right , the metadata will be displayed in structured objects in the window to the right.
  + Click the arrow point to the left , the json code will be displayed in a structured way in the window to the left.
* Copy, Paste and Save each element of the metadata in a separate file in JSON format.
  + Open a new file in Notepad ++
  + Type: {

}

* + Copy the element name part, e.g. "organisationUnitGroupSets":
  + Double click on the [
  + Copy and paste the highlighted part into the file
  + Save with the element name, e.g. organisationUnitGroupSets
  + Choose the file type to be saved as .json
* Run the code in JsonToHtml.py to convert each of the JSON files into HTML files.
  + Open JsonToHtml.py with Notepad++
  + Follow the instructions in the comments (green text) to
    - Replace to file and path to the file that is to be converted (.json)
    - Replace the location and file name of the result (.html)
    - Type
* Add CSS style to the HTML files by adding the code in CSS-styleCode.html on top of the code of each HTML file.

### LIMS/Bika

LIMS is implemented with Bika Health System and is hosted and managed by Accel. We obtained the metadata from Accel in JSON format in google documents.

To present the metadata in HTML, the following steps shall be followed:

* Save the metadata in .json documents.
* Run JsonToHtml.py to convert each of the metadata file from JSON to HTML.
  + Modify JsonToHtml.py
  + Open Windows Prompt (cmd, press Enter)
  + Navigate to the folder where JsonToHtml.py is located (use cd to navigate)
  + Run JsonToHtml.py (type JsonToHtml.py and press Enter)
* Add CSS style to the HTML files by adding the code in CSS-styleCode.html on top of the code of each HTML file.

The source code of Bika Health System metadata can be found on these two pages:

* Bika: <https://github.com/bikalims/bika.lims/tree/master/bika/lims/content>
* Bika Health: <https://github.com/bikalims/bika.health/tree/master/bika/health/content>

Since Bika Health System uses a object oriented database, we haven’t found a way to extract the metadata from its source code yet.

### iHRIS

Complete metadata of iHRIS is available on page <https://wiki.ihris.org/wiki/Liberia-Forms>

## Metadata of Excel Tools

Two documents shall be provided for Excel tools.

* A template of the excel tool, ideally with some sample data.
* Data fields or data dictionary including a list of data fields and their metadata.

To generate the Data Fields/Data Dictionary:

* Use SPSS to open the Excel template
* Copy the information on “Variable View” to the Data Fields/Data Dictionary document.
* Note: if the name of the Data Fields are not displayed proper in SPSS, the following method can be used to produce the data field names:
  + Copy the data fields name in the excel template
  + Open a new excel sheet -> right click on the new excel sheet -> select “Paste Special” -> select “Transpose”

For the SARA questionnaire and Health Workers Census Survey, in addition to the metadata, a copy of the questionnaire is also provided.

## Tools

### SPSS

SPSS is used to obtain the metadata of Excel files.

### Notepad++

Notepad++ can be used to view, edit and create codes in a number of formats, e.g. json, html, python, etc.

Json Online Editor

A JSON Editor can be used to view JSON data in an easy-to-understand structure. It can also be used to restructure the JSON code in a easy-to-read way.

One recommended Online Editor: <http://jsoneditoronline.org/>

### Python

The code mentioned in this guide are written in Python programming language. To run the code, Python needs to be installed.

To run python files in Windows:

* Download Python 2.7.14 from <https://www.python.org/downloads/>
* Install Python 2.7.14
* Put Python in your Path
* Control panel -> System -> Advanced system settings -> Advanced -> Environment Variables -> System variables -> Path
* Add “C:\Python27” to the Path
* In the Command Prompt, navigate to the location of your python file, and run file\_name.py

json2html

For the scripts that converts json to human readable html to work, the json2html Python wrapper shall be installed.

To install json2html:

* Check and make sure that pip.exe exists under C:\Python27\Scripts
* Include pip in your Path ()
* Control panel -> System -> Advanced system settings -> Advanced -> Environment Variables -> System variables -> Path
* Add “C:\Python27\Scripts” to the Path
* Run (type in Prompt): pip install json2html

### Simplejson

For the json2html.py code to work when converting Bika metadata, simplejson needs to be installed.

To install Simplejson:

* Run (type in Prompt): pip install simplejson

### Python Scripts

All python scripts can be obtained from the Git Hub page: <https://github.com/QingGongEHA/HIS-Metadata-Scripts>