PLANNING DOCUMENT 2/2

MARCH 17, 2022

The purpose of this document is to lay out the planned metadata architecture.

HARMONIZED EXPORT SPECIFICATION EM LOG RECORD

YES3 Export specification settings are stored as EM Log parameters, as opposed to EM project settings. The export specification EM log record definition is as follows.

setting	**description**
message	'yes3-export-specification'
user	REDCap username
export_uuid	unique id
export_name	name,
export_layout	'h' = horizontal 'v' = vertical 'r' = repeating form/event
export_selection	'1' = select all records '2' = select based on field+value criterion
export_criterion_field	REDCap field_name
export_criterion_event	REDCap event_id
export_criterion_value	see below
export_target	'download' or 'filesystem'
export_target_folder	REDCap application host filesystem folder or automount name
export_max_label_length	integer
export_max_text_length	integer
export_inoffensive_text	'1' or null
export_uspec_json	JSON upload specification string (see below)
export_items_json	JSON encoded items string (see below)
removed	'1' or '0'

NOTES

REMOVING SPECIFICATIONS

Export specifications are not permanently deleted. Rather, they are marked as 'removed' and can be restored from the Export Specification Editor.

THE EXPORT CRITERION VALUE

The `export_criterion_value` parameter can be either a single value, a list of values or an expression. Following are examples of possible entries:

```
'1'
'1, 2, 3'
'= 1'
'> 1'
'< 1'
'>= 1'
'<= 1'
'<> 1'
```

MAPPED REDCap VALUES

The `redcap_field_value` that is mapped to an upload specification value may be either a single value or a *list* of values ('1,2,3').

JSON SETTINGS

export_uspec_json

JSON encoding of:

```
{
    uspec_name: brief name,
    uspec_description: description,
    uspec_version: (string) use semantic versioning,
    uspec_version_date: (string),
    uspec_elements: [
        {
            uspec_element_name: uSpec element name,
            uspec_element_type: see below,
            uspec_element_label: label suitable for reports,
            uspec_element_valueset: [
                {
                    uspec_value: (string),
                    uspec_label: short label assoc w/value
                }
                ...
            ]
        }
        ...
    ]
}
```

export_items_json

JSON encoding of:

```
[
    {
        export_item_name: (uSpec element name or 'redcap_element_xx'),
        export_item_description: (deprecated),
        export_item_origin: 'redcap' or 'uspec',
        redcap_object_type: (if element_origin='redcap') 'form' or 'field',
        redcap_field_name: (if redcap_object_type='field'),
        redcap_form_name: (if redcap_object_type='form') REDcap form name
or 'all',
        redcap_event_id: REDCap event_id or 'all',
        uspec_element_name: (if export_item_origin='uspec'),
        uspec_element_type: (if export_item_origin='uspec'),
        uspec_element_label: (if export_item_origin='uspec'),
        uspec_element_value_map: (uSpec valueset with mapped REDCap values)
[
            {
                uspec_value: ('value' from uSPec valueSet),
                uspec_label: ('label' from uSpec valuset),
                redcap_field_value: REDCap value(s) mapped tp uSpec value
(see below)
            }
        ]
    },
    ...
]
```

DATA DICTIONARY

column name	description
var_name	The REDCap field name or uSpec element name.
var_label	The REDCap or uSpec label.
var_type	The data type (see Data Types, below).
valueset	A JSON-encoded array of dd valueset objects (see DD Valueset Object, below)
origin	'redcap' or 'uspec'
redcap_field_name	The REDCap field underlying this variable. Note that if the exported variable is derived from an upload specification, this will be the REDCap field associated with the upload specification element.
redcap_form_name	the REDCap form name underlying this variable
redcap_event_id	the numeric REDCap event_id underlying this variable
redcap_event_name	the unique REDCap event name underlying this variable
non_missing_count	the count of non-missing values for this export
min_length	the minimum length of the REDCap values for this export
max_length	the maximum length of the REDCap values for this export
min_value	the minimum value for this export*
max_value	the maximum value for this export*
min_value_formatted	the minimum value for this export, formatted as appropriate (ISO date etc)*
max_value_formatted	the maximum value for this export, formatted as appropriate (ISO date etc)*
sum_of_values	the sum of values for this export*
sum_of_squared_values	the sum of squared values for this export*
frequency_table	a JSON-encoded frequency table of observed values (nominal only)

* integer, float, date, time, datetime variables only.

The DD Valueset Object

A DD Valueset Object can be used to (1) associate a label with an exported value, and (2) if the variable is derived from an upload specification, the underlying REDCap value.

The DD Valueset Object Properties are:

property	description
value	one of the possible values that the exported value can assume
label	the label to associate with that label
redcap_field_value	if from an upload specification, the associated REDCap value *or list of values* (e.g., '1, 2, 3')

Data Types

Properties affected: `uspec_element_type` or `var_type`

data type	exported format
nominal	If from single-select REDCap field: raw value. If from multiselect field: comma-separated list of checked values. Associated Label(s) can be derived from the data dictionary.
integer	raw value
float	Base 10 exponential notation with up to 15 digits, e.g. '2.99792458E+8'
text	Raw value, possibly truncated and stripped of control characters and HTML tags depending on export options
date	yyyy-MM-dd
time	HH:mm:ss
datetime	yyyy-MM-dd HH:mm:ss