The main script **curation_4.m** calls five functions that represent an arm of the data design.

First arm (arm_organismv5.m) opens and creates a specimen file and sets up the groups and datasets associated with the organism;

Second arm (arm_devicev5.m) is associated with instruments used to create data;

Third arm (arm_anatomicalv5.m) is the anatomy of the organism where the cells where isolated;

Fourth arm (arm_cellv5.m) is the description of the cell that was used to make the electrical recording and the

Fifth arm (arm_assayv10.m) is the details of the assay that were conducted.

The five function calls are

arm_organismv5(dirname, filename_fits,k_adult_male,count);

arm_devicev5(dirname, filename_fits,k_adult_male,count);

arm_anatomicalv5(dirname, filename_fits,k_adult_male,count);

arm_cellv5(dirname, filename_fits,k_adult_male,count); and

arm_assayv10(dirname, filename_fits,k_adult_male,count)

where

dirname: the path where the MATLAB data to be translated is found;

filename_fits: the name of the MATLAB data file with extension .MAT to be translated it is an array of structs;

k_adult_male: number of experiments within the original MAT file that are to be translated;

count: is starting number of the specimen file that will be created.

The following additional seven (7) functions are called from the arms.

write_attribute_for_group(group_id_1a,dadefinition,ATTRIBUTE)

This writes metadata that is associated with group_id_1a; the metadata is described in dadefinition and defined in ATTRIBUTE.

DATASETID=create_and_write_string_dataset(group_id_2a,space,type,name_def,vendor);

This creates and writes dataset of type string that will be found under group_id_2a; with space and type defined, the description of the string is found in name_def (i.e., description of the dataset coined vendor) and the contents of the string are found under vendor. The reference associated with the dataset is DATASETID needed for compiling and is returned to main script.

DATASETID
=create_and_write_double_dataset(group_id_3a,space,type,name_def,weight);

This creates and writes dataset of type DOUBLE precision that will be found under group_id_3a; with space (SCALAR or NULL) and type (DOUBLE precision) defined. The description a character string is found in name_def and the double precision number is found in weight. The reference associated with the dataset is DATASETID needed for compiling and is returned to main script.

DATASETID =create_and_write_int_dataset(group_id_3a,space,type,name_def,age);

This creates and writes dataset of type INT that will be found under "group_id_3a"; with space (SCALAR or NULL) and type (INT precision) defined. The description of the dataset is found under character string name_def and the integer is found under age. The reference associated with the dataset is DATASETID needed for compiling and is returned to main script.

DATASETID
=create_write_array_of_dble_dset(group_id_4a,space_id,type,dim0,dimw,name_def,Imlf)

This creates and writes an array of type DOUBLE with dimensions dim0 and dimw that will be found under group_id_4a; with space (SCALAR) and type (DOUBLE precision) defined. The description of the dataset is found under character string name_def and the double precision data array is found in Imlf. The reference associated with the dataset is DATASETID needed for compiling and is returned to main script.

METADATA

Attributes are METADATA attached to the description of the dataset.

attribute_general(DATASETID,researcher,dofexp, cellnumber, datasteward)

To every dataset within the collection the name of the researcher (researcher) that conducted the experiment, the date of the experiment (dofexp); the original cell number (cellnumber) and the name of the person responsible for the data (datasteward) are attached as attributes to the dataset DATASETID.

specific_string_attribute(DATASETID,ATTRIBUTE,dadefinition)

This is description of the dataset where ATTRIBUTE is a string like "definition', "units' or 'note' and the "dadefinition" is the actual description of the strings. They are attached to dataset with id: DATASETID