The Concepts bulk loader has been designed to load new Concepts into caDSR via spreadsheet.

It takes 2 steps for a caDSR curator to upload new Concepts:

1.Create *Concepts.csv* file.

2.Make request for Bulk Load with attached *Concepts.csv* file.

1.In order to create Concepts.csv file you should fill out the template *Sample\_Upload\_Concepts.csv* following instructions in [*data\_upload\_instuctions.txt*](https://github.com/CBIIT/cadsr-db/blob/master/bulk-data-load/data-owners/CONCEPTS/data_upload_instuctions%20.txt).

Files locations are placed below.

The sample of required template “***Sample\_Upload\_Concepts.csv****”* is located at

<https://github.com/CBIIT/cadsr-db/blob/master/bulk-data-load/data-owners/CONCEPTS/Sample_Upload_Concepts.csv>

The Instruction file to how to create a spreadsheet for Concepts “[***data\_upload\_instuctions.txt***](https://github.com/CBIIT/cadsr-db/blob/master/bulk-data-load/data-owners/CONCEPTS/data_upload_instuctions%20.txt)” is located at

<https://github.com/CBIIT/cadsr-db/blob/master/bulk-data-load/data-owners/CONCEPTS/data-upload-instructions.txt>

When you create *Concepts.csv* file make sure it has the same format as *Sample\_Upload\_Concepts.csv*.

First row consists of columns names: PREFERRED\_NAME, LONG\_NAME, PREFERRED\_DEFINITION, EVS\_SOURCE, ORIGIN, SOURCE\_DEFINITION.

Names and order the columns must remain as given.

Columns PREFERRED\_NAME, LONG\_NAME, PREFERRED\_DEFINITION must be populated otherwise a record will be rejected.

The columns EVS\_SOURCE, ORIGIN, SOURCE\_DEFINITION have following rules:

a. If all three columns (EVS\_SOURCE, ORIGIN, SOURCE\_DEFINITION) are empty, default value (NCI Thesaurus, NCI\_CONCEPT\_CODE, NCI)

will be created for the concept.

b. If all three columns are not empty and evsSource is valid, concept will be loaded with the data in the file.

c. If one or two out of the three columns are empty, concept will NOT be loaded.

Authorized columns length:

PREFERRED\_NAME must not exceed 30 Characters.

LONG\_NAME must not exceed 255 Characters.

PREFERRED\_DEFINITION must not exceed 2000.

EVS\_SOURCE must not exceed 255 Characters.

ORIGIN must not exceed 240 Characters.

SOURCE\_DEFINITION must not exceed 2000 Characters.

After the file is created:

1.Name the created file CONCEPTS.

2.Convert an Excel worksheet to CONCEPTS.csv file by using the Save As command.

3.Click the File menu, and then click Save As.

4.In the Save as type box, choose CSV (Comma delimited).

2.Every time when Curators make request on bulk data loading to caDSR has to include 5 attached files:

[1.Newly created Concepts.csv](https://github.com/CBIIT/cadsr-db/blob/master/bulk-data-load/dba/CONCEPTS/DBA_Upload_Instructions.txt)

2.[DBA\_Upload\_Instructions.txt](https://github.com/CBIIT/cadsr-db/blob/master/bulk-data-load/dba/CONCEPTS/DBA_Upload_Instructions.txt)

3. [CONCEPTS\_DBA.ctl](https://github.com/CBIIT/cadsr-db/blob/master/bulk-data-load/dba/CONCEPTS/CONCEPTS_DBA.ctl)

4. DELETE\_TEMP\_CONCEPTS.sql

5. UPLOAD\_CONCEPTS.sql

Files 2,3,4,5 can be find at

<https://github.com/CBIIT/cadsr-db/tree/master/bulk-data-load/dba/CONCEPTS>

Please copy and them and make sure to attach to the request.

The Original request to design bulk load EVS concepts from all sources is located at <https://tracker.nci.nih.gov/browse/CADSRMETA-534>

DB script for bulk load EVS concepts from all sources is located at

<https://github.com/CBIIT/cadsr-db/blob/master/bulk-data-load/dba/CONCEPTS/UPLOAD_VALIDATE_CONCEPTS.sql>