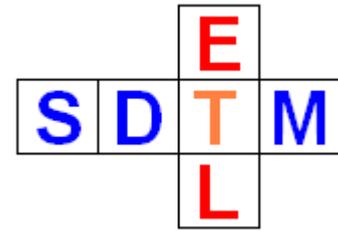


# SDTM-ETL 4.4 User Manual and Tutorial

Author: Jozef Aerts, XML4Pharma

Last update: 2023-11-12



## Tutorial: Save define.xml for batch execution (and more)

## Table of Contents

Table of Contents .....	1
Introduction.....	1
Generating the slimmed down define.xml .....	2
Conclusions.....	4

## Introduction

SDTM-ETL has already a feature "Save cleaned define.xml", which essentially is meant to generate a define.xml that is cleaned for everything that is not needed for a regulatory submission. This includes the template domain definitions, unused variables, codelists that are not used, and especially, the mapping scripts. The latter (unfortunately) are not of interest to e.g. the FDA and PMDA.

The "Save cleaned define.xml" also removes the ["GLOBAL" set of variables](#), variables that are defined once and used many times in calculations in mapping scripts. Examples are RFSTDTC (Subject Reference Start Date/Time (in DM) for the calculation of all --DY (study day) variables, and also RFXSTDTC (Date/Time of First Study Treatment - also in DM), for the calculation of baseline flags (--LOBXFL: Last Observation Before Exposure Flag), and for when timepoints relative to the first exposure are used.

This means that essentially, such a "cleaned define.xml" is not usable for [batch execution](#) of the mappings to generate SDTM/SEND datasets without the need of using the Graphical User Interface (GUI). One can of course simply use the normal saved define.xml (which includes the template, and all the codelists, including the ones that are never used), but under some circumstances, this has some disadvantages:

- as the template and the unused codelists are still in the define.xml, loading takes considerably more time
- this may also have an effect on memory usage

Therefore, some of our users have asked to provide functionality that produces a "slimmed down" define.xml that can easily be used in batch executions, among others.

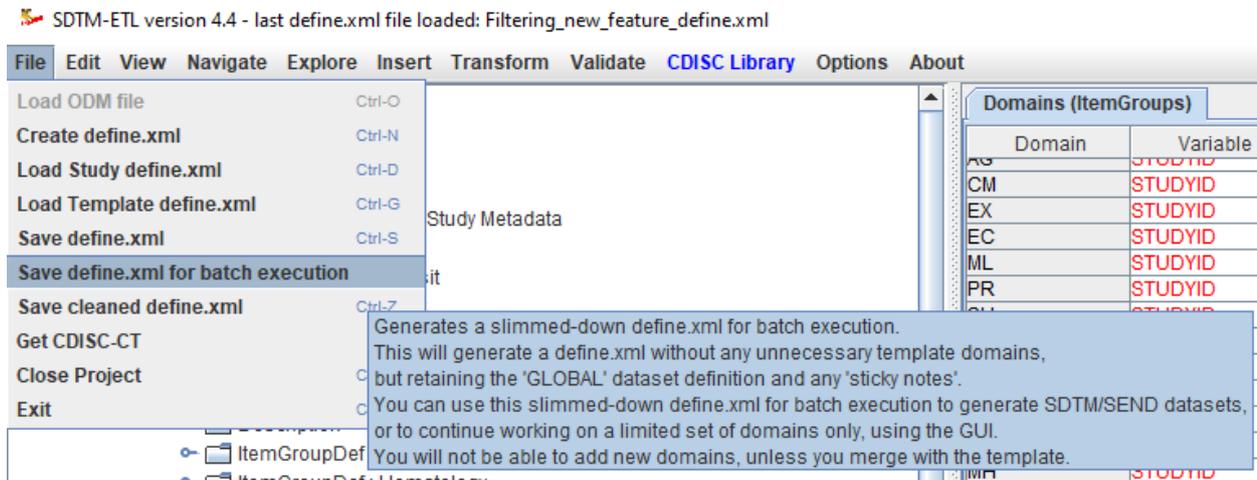
One of the reasons they asked for it, is that with each new version of SDTM and SEND Implementation Guides, the number of domains is growing<sup>1</sup>, and with each new publication of CDISC controlled terminology, the number of codelists is growing.

---

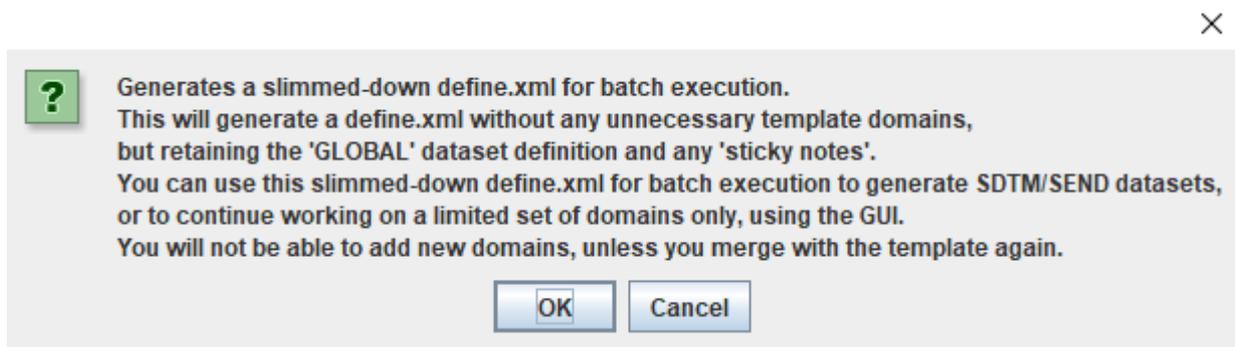
<sup>1</sup> About 15 years ago, Wayne Kubick (CDISC CTO) stated "we will never have more than 25 domains" ...

# Generating the slimmed down define.xml

As of SDTM-ETL v.4.4, a new menu item under "File" has been added:

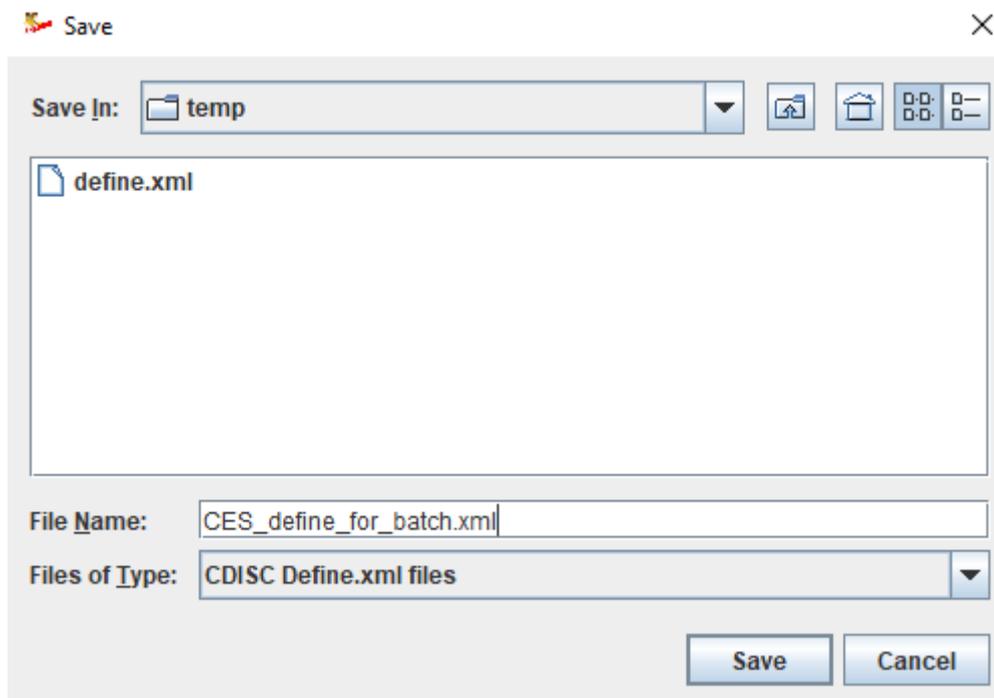


When one uses it, more information is provided in a dialog:



explaining the possible usages of the slimmed-down define.xml.

When one then clicks OK, after a few seconds (in which time the define.xml is generated) the system asks where to store the slimmed-down define.xml:



That's it! There are no options such as when using "Save cleaned define.xml", so everything is easy...

## Uses of the slimmed-down define.xml

Just as any SDTM-ETL define.xml file with mappings, the "slimmed-down" can be used for [batch execution](#) of the mappings to generate SDTM or SEND datasets. See the [tutorial](#) for more explanation and the command parameters. The advantage of using the slimmed-down define.xml is that it will load faster and use less memory.

One can however also use the "slimmed down" define.xml in the GUI. The advantage is that it will load much faster than the conventional SDTM-ETL file that still contains the template and all the codelists (including the ones never used). There are however also some severe advantages:

- No mappings for additional domains can be added, as the template is not present anymore
- One cannot start generate codelist-codelist mappings using the wizards for codelists that have not been used so far, and thus are not present.

Thus, it is recommended to only use the "slimmed-down" define.xml in the GUI, when the mappings are in an advanced stage.

Of course, one can always re-establish the template by merging with a template file, using the menu "File - load Template define.xml" and allowing to merge, or to to use "File - load Study define.xml" and allow to merge with a define.xml with mappings (e.g. for another domain), that still has the template included. In this case, also all the codelists will be reloaded.

When using "File - load Template define.xml", no additional codelists will be added, and one will need to use the menu "Insert - Codelist Definitions from File into define.xml" and then select a file containing all the codelists, e.g. from the folder "CDISC-CT" where all by CDISC published codelists are located.

## Conclusions

The new (as of SDTM-ETL v.4.4) feature to save a "trimmed-down" define.xml is especially interesting for speeding up batch execution of the mappings in order to generate SDTM or SEND datasets without needing to use the GUI.

The "slimmed-down" can also be used with the GUI when the mappings for that specific (set of) domain(s) is in advanced stage, but there are some drawbacks.