

Re-run export without re-fetching data

Contents

Scenario.....	1
Approach.....	1
Pre-requisites.....	1
Re-running export step is safe	2
Modifying the export process	2

Scenario

You need to alter how data is exported from the process, without a lengthy or difficult process of re-fetching citations, and enhancing them from external sources. e.g.:

- Different columns are needed in the spreadsheet
- The rules for processing, selecting and collating data need to be changed
- A different output format or target is needed e.g. writing directly to Excel, producing HTML, or writing to a database

Generating the export is typically very quick (less than a second for 50 reading lists) but fetching the citation data and enhancements can take many hours, and so it may be desirable to run the export step in isolation.

Approach

With the previously collected enhanced citations in place (Data/MODULE.json):

1. Modify export scripts
2. Run export step in isolation:
php batch.php -m MODULE -s export
3. Check, and debug and repeat if necessary

Pre-requisites

If you have already run the previous steps (get, alma, scopus etc) you will already have the software and files in place.

If not (e.g. if taking over the export work from someone else who has done the earlier steps):

- Install and configure the software as described in **1_1_install_configure_and_run_the_software.pdf**
- *If you're running the export step in isolation, you will **not** need the API keys in the config.ini file (although you will need to create a config.ini file), and you will **not** need to install and run the software on a machine on the wired network of a Scopus-subscribing institution*

- Copy the citation JSON files (MODULE.json) previously obtained into the directory **Data/**
- Create a text file listing the module codes for processing e.g. **Config/Modules/modcodes.txt**

Re-running export step is safe

Provided you **only** run the export step (i.e. provided you call **batch.php** with **-s export**) the citation JSON files will not be modified, and so the export can be repeated as many times as needed (e.g. when developing/debugging).

*If you forget to specify **-s export**, the batch script will run the full process, from Leganto through Alma, Scopus etc to Export, which will be time-consuming and which **will** modify the citation JSON files.*

Modifying the export process

This is a question of:

1. Identifying the business-level changes needed
2. Identifying particular lists and citations that demonstrate the changed requirements, so that you can check the resulting reports
3. Making changes in the code of **simpleExport.php** and **longExport.php**
See the documents [3_3_the_short_report_export_script.pdf](#) and [3_4_the_long_report_export_script.pdf](#) for detailed descriptions of these scripts
simpleExport.php produces the **summary** and **short** reports, **longExport.php** produces the **long** reports
4. Running the export step only, for the full set of modules we have MODULE.json files for:
php batch.php -f Config/Modules/modcodes.txt -s export
Assuming you have a list of module codes in Config/Modules/modcodes.txt
5. Checking the exported reports e.g. looking in particular at the lists and citations you identified above
6. Making any further changes needed, and re-running the export step for the full set of modules