Processing AMED records pre- and post- import to Library Master

### Overview

The executable file amed\_tools.exe is used to process ETOC records before these records are edited within Library Master, and any AMED records exported from Library Master.

### Processing ETOC records before import to Library Master

#### Set up executable files

Copy the following files from [\\ad.bl.uk\bssdata\Utilities\VM tools\amed](file:///\\ad.bl.uk\bssdata\Utilities\VM%20tools\amed) to the folder in which you are working:

* amed\_tools.exe

Ensure that the following files are present in the folder [\\ad.bl.uk\bssdata\AMED\Exports\Lookup](file:///\\ad.bl.uk\bssdata\AMED\Exports\Lookup):

* AMED journal title lookup table.txt
* amed\_citations.db

#### Processing

From the command line, run:

amed\_tools.exe –i <input\_file> --LM1

where <input\_file> is the name of the input file. The input file should be a text (.txt) file with one ETOC record per line, as in the example shown in the image below:

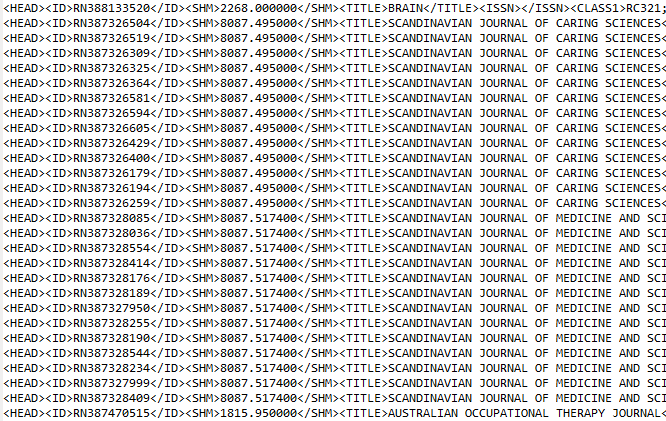
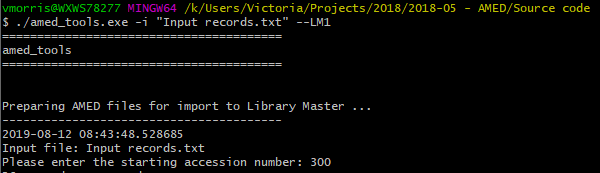
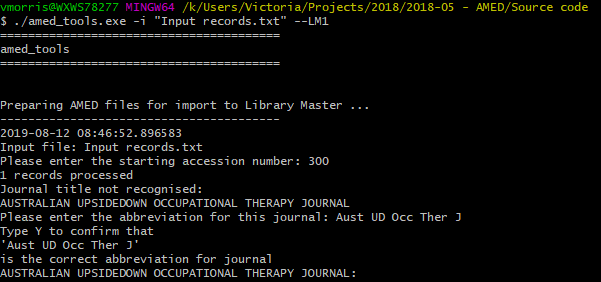


Figure 1: Extract from example file of ETOC records

You will be prompted to enter the accession number for the first record to be processed. Type an integer, then press Enter. The program will automatically pad this number with leading zeros to ensure that the length is correct.



If any journal titles (within ETOC field <TITLE>) are not recognised, you will be prompted to enter abbreviated title(s). These new abbreviated titles will be added to the list of journal titles in the file AMED journal title lookup table.txt once processing is complete.



Citations will be compared against the list stored within the database amed\_citations.db. For records where the citation does not include pagination, 20 from the title will be appended to the citation in order to check for duplication.

Records found to be duplicates will be written to the output file YYYYMMDDAMED\_duplicates.csv where YYYY MM DD are components of today’s date.

None-duplicates will be split into batches containing at most 50 records each. These batches will be written to one or more output file(s), in the same folder as the input file, with filenames of the form

YYYYMMDDAMEDN.csv

where YYYY MM DD are components of today’s date, and N is a running number (with the first file produced being numbered 1, and so on). These output files should be ready to load to Library Master.

The columns in each output file are:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Accession number** | **ISSN** | **[Blank column]** | **Title** | **Citation** | **Authors** | **Abstract** | **Abstract indicator** |
| A 7-digit integer, starting with the accession number supplied to the program, and incrementing by 1 for each record. | The contents of ETOC field <ISSN>, if present. |  | The title of the article, taken from the ETOC field <TEXT> | A citation, created by concatenating the abbreviated title for the journal with the ETOC fields <VOL>, <NUMBER>, and <PAGE> | Names of all authors, taken from the ETOC field <AUTH>, and formatted ***Last name, Initials*** | The contents of ETOC field <ABS>, if present. | Contains ‘AB’ if an abstract is present in the previous column; otherwise blank. |

The program also performs the following text-cleaning functions:

* Replaces characters from outside the basic ASCII character set with appropriate HTML entities. For example, é will be replaced by &#233. (This is necessary because Library Master cannot easily cope with non-ASCII characters);
* Replaces smart quote characters (e.g. “) with straight quotes (").
* Removes HTML tags (<p>, <br/> etc.) from the Abstract field.
* Normalises whitespace.

Once the program has run, you should check the file AMED journal title lookup table.txt to make sure that new journal titles have been added correctly. It is important that the file ends with a blank line.

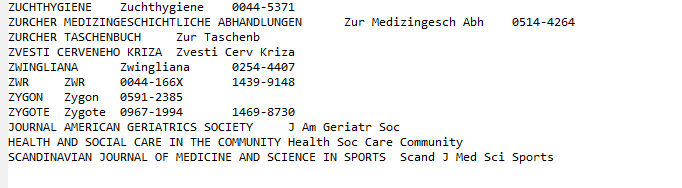


Figure 2: Extract from AMED journal title lookup table.txt showing new journal titles added

The database amed\_citations.db should be checked periodically, using software such as [DB Browser for SQLite](https://sqlitebrowser.org/).

### Processing AMED records exported from Library Master

#### Processing

From the command line, run:

amed\_post\_LM.exe –i <input\_file> --month=<month\_number> --LM2

where <input\_file> is the name of the input file, and <month\_number> is the number of the month to which the file applies (e.g.--month=04 to process the file for April). The input file should be a text (.txt) file exported from Library Master. If <month\_number> is not specified, the program assumes that the file applies to the current month.

The following output files will be produced:

* amdmonthYY.txt
* amedMMYY for hosts.txt
* AMEDmonth.spl
* F164MMDD.dat
* F164.end
* AMED stats YYYY-MM-DD.txt

where YYYY MM DD are components of the date, and month is a three-letter abbreviation for the name of the month (lower-case).

The file AMED stats YYYY-MM-DD.txt contains a brief summary of the numbers of records processed, and range of accession numbers used.

The fields contained in other files are summarised in the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | **Present in … ?** | | | |
| **amdmonthYY.txt** | **amedMMYY for hosts.txt** | **AMEDmonth.spl** | **F164MMDD.dat** |
| ***Output file for*** | | SilverPlatter and Ovid |  | Datastar, Dimdi and EBSCO | Dialog and Dialog2 (also with F164.end) |
| ***Long fields wrapped?*** | | Y | Y | N | Y |
| **Field** | **Description** |  |  |  |  |
| AN | Accession number | Y | Y | Y | Y |
| UD | Date | Y | N | N | N |
| AU | Authors | Y | Y | Y | Y |
| TI | Title | Y | Y | Y | Y |
| SO | Source | Y | Y | Y | Y |
| ET | Entry terms | Y | Y | Y | Y |
| KW | Keywords | Y | Y | Y | Y |
| MT | Minor terms | Y | Y | Y | Y |
| TY | Publication type | Y | Y | Y | Y |
| LA | Language | Y | Y | Y | Y |
| ES | English summary indicator | Y | Y | Y | Y |
| IS | ISSN | Y | Y | Y | Y |
| MD | Abstract indicator | Y | Y | Y | Y |
| AB | Abstract | Y | Y | Y | Y |

### Source code

Code is written in Python, version 3.7, and source code is saved in the folder [K:\Utilities\VM tools\amed\Source code](file:///K:\Utilities\VM%20tools\amed\Source%20code).