

Using python, sqlite, and DBeaver with EBSCO HLM downloads

Melissa Belvadi
mbelvadi@upei.ca

https://github.com/Melissa-Belvadi-Library-Projects/hlm_download_management

hlm_download_management Public Edit Pins Unwatch 1

main 1 Branch 0 Tags Go to file Add file <> Code

mbelvadi	Mac instructions for python, sqlite, dbeaver	146fca0 · last week	10 Commits
LICENSE	Initial commit		last month
README.md	Update README.md		last month
database_schema	Update database_schema		2 weeks ago
hlm_python_sqlite_isbns_subjects.py	Update hlm_python_sqlite_isbns_subjects.py		2 weeks ago
mac_python	Mac instructions for python, sqlite, dbeaver		last week
sql_scripts	Update and rename sql_scripts_subjects to sql_scripts		2 weeks ago
windows_python	Windows users: python, sqlite, and dbeaver installation		last week

The actual python script

README MIT license

hlm_download_management

For EBSCO HLM customers to work with downloaded collection files

Many EBSCO customers (library staff) use EBSCO's holdings management system (known as HLM) to keep

Current Group: Main User Group (main) ▼

Customize Services

Authentication

Holdings Management

Local Collections

Reports & Statistics

Database Title Lists

Customize EBSCO services ?

dit profiles for your library that determine which search options, databases, and collections are available to your patrons.

Choose Profile: EBSCOhost Databases (smab6n) (ehost) - ehost ▼

[Profile Maintenance](#)

Description (Profile ID) - Interface

Searching

Databases

Viewing Results

Linking

Delivery Options

Multilingual Options

Branding

General Settings

Search Modes

[Modify](#)

Default Keyword Search Screen

Advanced ▼

Show Keyword/Title/Author Radio Buttons

☒ On ☐ Off

Require uppercase Boolean operators (AND, OR, NOT)

☐ On ☒ Off

Diacritic Search

☒ On ☐ Off

Display Text-to-Speech

☒ Yes ☐ No

Configure Text-to-Speech

[Modify](#)

Spellchecker

☒ On ☐ Off

Show Did-You-Mean Suggestion

☒ Yes ☐ No

Auto Correct Search Terms

☒ Yes ☐ No

Autocomplete Keyword

☒ On ☐ Off



Download Holdings

Download Holdings Integration Changes

Collection Assessment

View All Downloads

Type ?

- Standard
- KBART

Format ?

- CSV
- Tab Delimited

[Customize Services](#)[Authentication](#)[Holdings Management](#)[Local Collections](#)[Reports & Statistics](#)[Database Title Lists](#)[Overview](#)[Titles](#)[Packages](#)[Vendors](#)[Links](#)[Notes](#)[Upload](#)[Download](#)[Languages](#) ▼[Download Holdings](#)[Download Holdings Integration Changes](#)[Collection Assessment](#)[View All Downloads](#)

◀◀ ◀ 1 ▶ ▶▶

Number of Results 20 ▼

File Name	Type	Title List Summary	Date	Size	Status	Download
hlm_20250110_rl	Standard Holdings List	Title List Summary	Fri 01/10/2025 10:15 AM	149 MB	Completed	Download

Get to Know the Data!

Download tsv - UTF-16

KBID, Resource record, Package record - combination of all three for custom settings

Custom Package can include Managed or Custom resources but custom URL

KBID + PackageName = unique

PrintISBN,OnlineISBN - can have multiple values separated by pipe |

Hyphens are inconsistent - best to strip them out - my script does that

Subject - can have multiple values separated by pipe

Some are BISAC, some are LCSH, you can't tell from the download data

UserDefinedField1-5 - you can relabel, make public visible or not

You can insert "returns" (/n) when doing manual data entry but don't

Python Code

Open with a plain text editor (not Word!)

Change the UPEI customization for UserDefinedField1-5 easily:

```
column_renames = {  
    'UserDefinedField1': 'StaffOnlyNotes',  
    'UserDefinedField2': 'OtherInfoAboutThisTitle',  
    'UserDefinedField3': 'SpecialNotesAboutParticularVolumes',  
    'UserDefinedField4': 'UPEIAccessNote',  
    'UserDefinedField5': 'CallNumber'  
}
```


mbeivadi@apex:~/Downloads\$ python3 hlm_python_sqlite_isbns_subjects.py hlm_20250110_all.tsv

Step 1: Read the UTF-16 file directly and split lines into columns

Header columns: ['KBID', 'Title', 'AlternateTitle', 'PackageName', 'URL', 'ProxiedURL', 'Publisher', 'Edition', 'Author', 'Editor', 'OnlineISSN', 'PrintISBN', 'OnlineISBN', 'DOI', 'PeerReviewed', 'ManagedCoverageBegin', 'ManagedCoverageEnd', 'CustomCoverageBegin', 'CoverageStatement', 'Embargo', 'CustomEmbargo', 'Description', 'Subject', 'ResourceType', 'PackageContentType', 'CreateCustom', 'HideCover', 'FullTextFinder', 'ExcludeFromMARC', 'Delete', 'OrderedThroughEBSCO', 'IsCustom', 'UserDefinedField1', 'UserDefinedField2', 'UserDefinedField3', 'UserDefinedField4', 'UserDefinedField5', 'SmartLinks+ Eligible', 'PackageType', 'AllowEBSCOtoSelectNewTitles', 'PackageID', 'VendorName', 'VendorID', 'VendorPart', 'Merged', 'Split']

Number of header columns: 51

Step 2: Rename specific columns

Step 3: Replace spaces with underscores and plus sign with "Plus"

Processed header columns: ['KBID', 'Title', 'AlternateTitle', 'PackageName', 'URL', 'ProxiedURL', 'Publisher', 'Edition', 'Author', 'Editor', 'OnlineISSN', 'PrintISBN', 'OnlineISBN', 'DOI', 'PeerReviewed', 'ManagedCoverageBegin', 'ManagedCoverageEnd', 'CustomCoverageBegin', 'CoverageStatement', 'Embargo', 'CustomEmbargo', 'Description', 'Subject', 'ResourceType', 'PackageContentType', 'CreateCustom', 'FullTextFinder', 'ExcludeFromMARC', 'Delete', 'OrderedThroughEBSCO', 'IsCustom', 'StaffOnlyNotes', 'OtherInfoAboutThisTitle', 'UPEIAccessNote', 'CallNumber', 'SmartLinksPlus_Eligible', 'PackageType', 'AllowEBSCOtoSelectNewTitles', 'PackageID', 'VendorName', 'VendorID', 'VendorPart', 'Continued', 'Continued_in_part', 'Merged', 'Split']

Number of processed header columns: 51

Step 4: Create SQLite database and table

Step 5: Create unique index on KBID and PackageName

Step 6: Create non-unique indexes

Step 7: Create kbid_isbns table

Step 8: Create unique index on kbid_isbns

Step 9: Create kbid_subjects table

Step 10: Create unique index on kbid_subjects

Step 11: Populate kbid_isbns and kbid_subjects tables

Step 12: Create non-unique indexes for kbid_isbns table

Step 13: Create non-unique indexes for kbid_subjects table

Processing complete.

```
-rw-r--r-- 1 mbelvadi mbelvadi 4062105600 Jan 14 08:39 hlm_20250110_all.db
```



- SQLite
- Db2 for LUW
- MariaDB
- MySQL
- Oracle
- PostgreSQL
- SQL Server
- Altibase
- Apache Calcite Avatica
- Apache Drill
- Other
- From JDBC URL ...
- New Database Connection Shift+Ctrl+N

New Database Connection Shift+Ctrl+N
Last edited SQL script Ctrl+Enter
Show Key Assist Shift+Ctrl+L
Find Actions Ctrl+3

Project - General X



Name

DataSource

- Bookmarks
- Dashboards
- Diagrams
- Scripts

- 5.1 merlin search.db
- hlm_20240503_all.db
- hlm_lcc.db

- | Name | DataSource |
|------------|------------|
| Bookmarks | |
| Dashboards | |
| Diagrams | |
| Scripts | |

Connect to a database

Generic JDBC Connection Settings

SQLite connection settings

Main Driver properties

General

Connect by: ☒ Host ☐ URL

JDBC URL: jdbc:sqlite:/home/mbelvadi/Downloads/hlm_20250110_all.db

Path: /home/mbelvadi/Downloads/hlm_20250110_all.db Open ... Create ...

The first time you use sqlite you may be asked to download a driver and it will find and install the correct driver for you

[Connection variables information](#) [Database documentation](#) Connection details (name, type, ...)

Driver name: SQLite Driver Settings Driver license

Test Connection ... < Back Next > Cancel Finish

Right-click -
View
Diagram

Enter a part of

- 5.1 merlin
- hlm_20240503_...db

- hlm_20250110_all.db

- Tables
 - kbid_isbns
 - kbid_subjects
 - main
- Views
- Indexes
- Sequences
- Table Triggers
- Data Types
- hlm_1cc.db

Project - General

Name

DataSource

- Bookmarks
- Dashboards
- Diagrams
- Scripts

*hlm_20250110_all.db x

Properties ER Diagram

hlm_20250110_all.db

main

A-Z KBID
A-Z Title
A-Z AlternateTitle
A-Z PackageName
A-Z URL
A-Z ProxiedURL
A-Z Publisher
A-Z Edition
A-Z Author
A-Z Editor
A-Z Illustrator
A-Z PrintISSN
A-Z OnlineISSN
A-Z PrintISBN
A-Z OnlineISBN
A-Z DOI
A-Z PeerReviewed
A-Z ManagedCoverageBegin
A-Z ManagedCoverageEnd
A-Z CustomCoverageBegin
A-Z CustomCoverageEnd
A-Z CoverageStatement

kbid_subjects

A-Z KBID
A-Z Subject

kbid_isbns

A-Z KBID
A-Z ISBN
A-Z PrintOrE

Refresh Save ... Revert Keep layout

100%

AST en_CA



Database Navigator ×

Projects

Enter a part of object name here

5.1 merlin search.db

hlm_20240503_all.db

hlm_20250110_all.db

Tables

kbid_isbns

kbid_subjects

main

Views

Indexes

Sequences

Table Triggers

Data Types

hlm_1cc.db

Project - General ×

Name

Bookmarks

Dashboards

Diagrams

Scripts

SQL Editor

Create

Edit Connection

F4

Connection view

Browse from here

Filter

Connect

Invalidate/Reconnect

Disconnect

View Diagram

Shift+Ctrl+Enter

Compare/Migrate

Import Data

Tools

Copy

Ctrl+C

Paste

Ctrl+V

Copy Advanced Info

Shift+Ctrl+C

Delete

Rename

F2

Refresh

F5

Open SQL script

F3

Last edited SQL script

Ctrl+Enter

New SQL script

Ctrl+]

Open SQL console

Ctrl+Alt+Enter

New Database Connection Shift+Ctrl+N
Last edited SQL script Ctrl+Enter
Show Key Assist Shift+Ctrl+L
Find Actions Ctrl+3

DBEaver 24.3.1 - <hlm_20250110_all.db> Script

File Edit Navigate Search SQL Editor Database Window Help

SQL Commit Rollback Auto hlm_20250110_all.db < N/A >

Database Navigator x Projects x *<hlm_20250110_all.db> Script x

Enter a part of object name here

- 5.1 merlin search.db
- hlm_20240503_all.db
- hlm_20250110_all.db**
 - Tables
 - kbid_isbns
 - kbid_subjects
 - main
 - Views
 - Indexes
 - Sequences
 - Table Triggers
 - Data Types
- hlm_1cc.db

Execute this query or right-click to choose Export from Query

```
select VendorName , count(vendorname) VendorCount
from main group by VendorName order by VendorCount desc;
```

main 1 x

select VendorName , count(vendorname) VendorCount

	A-Z VendorName	123 VendorCount
1	EBSCO	641,053
2	Coherent Digital	282,464
3	ProQuest Info & Learning Co	278,025
4	HeinOnline	255,156
5	Gale Cengage	248,738
6	JSTOR	160,905
7	UNIVERSITY OF PRINCE EDW	158,169
8	OAPEN	133,762
9	Alexander Street	128,283

Refresh Save Cancel Export data 200

197 197 row(s) fetched - 0.246s (0.003s fetch), on 2025-01-14 at 09:24:38

AST en CA Writable Smart Insert 2 : 1 : 51 Sel: 0 | 0