



The State University of New York

How to Identify and Clean-up Bad Diacritics for IZ Only Physical Bib Records

Maggie McGee
October 12, 2021

www.suny.edu



Identify and Clean-up Bad Diacritics for IZ Only Bib Records

In this session we will cover:

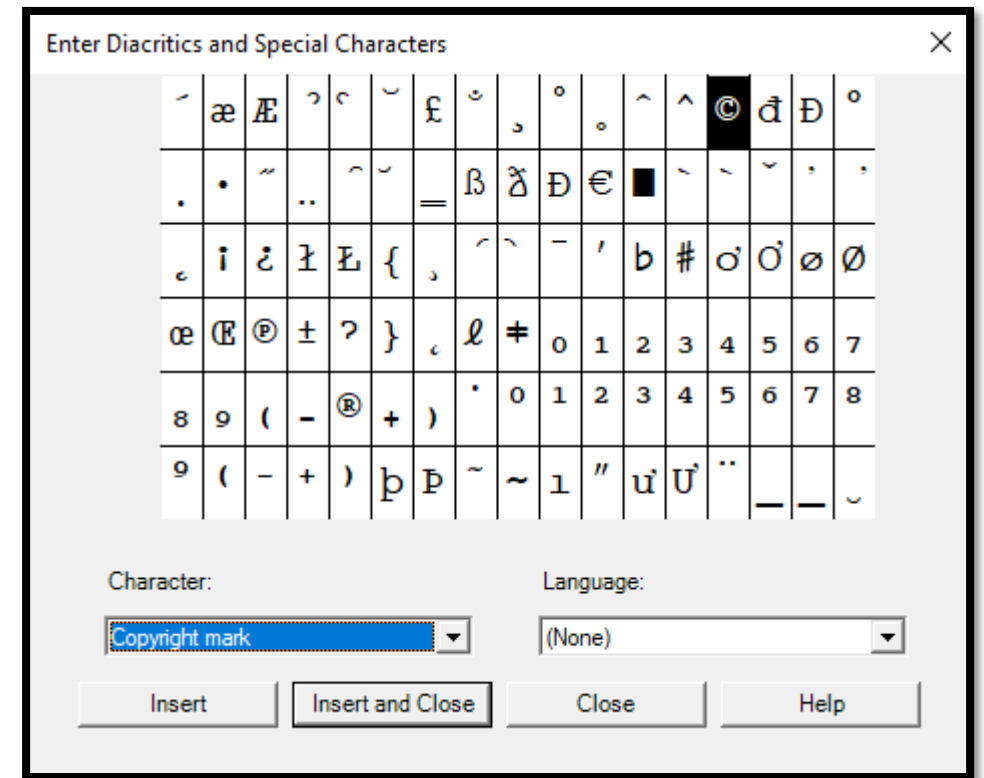
1. Define bad/corrupted diacritics
2. Causes of bad/corrupted diacritics
3. How bad/corrupted diacritics are being updated in the NZ
4. How to identify bad/corrupted diacritics for IZ only bib records
5. How to update identified bib records with bad/corrupted diacritics



Bad/Corrupted Diacritics Defined

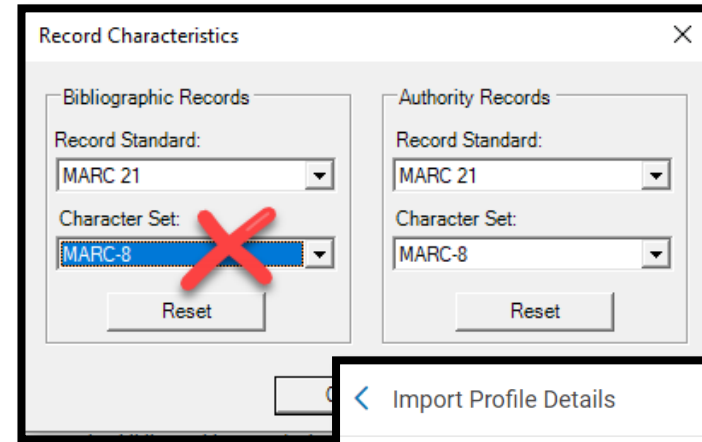
- Diacritics are symbols that are above or below a character to indicate the character has a different pronunciation
- Using the term diacritics in this session also includes special characters, such as the copyright symbol ©
- Bad/corrupted diacritics refer to characters containing symbols or special characters that do not display properly in Alma and/or Primo VE

050	00	a PZ7.C8153573 b In 2020
082	00	a [Fic] 2 23
100	1_	a C<U+fffd>ordova Zoraida, e author.
245	10	a Incendiary / c Zoraida C<U+fffd>ordova.
250	__	a First edition.
264	_1	a Los Angeles ; a New York : b Hyperion, c 2020.
264	_4	c C<U+fffd>2020
300	__	a 450 pages ; c 24 cm



Causes of Bad/Corrupted Diacritics

- Alma cannot convert diacritics and special characters in bib records that are imported into Alma as MARC-8
- Connexion Client Settings:
 - Record Characteristics for “*Character Set*” is not configured to UTF-8 (*Tools>Options>Export tab>Record Characteristics*)
- Connexion Web Version:
 - Records can only be imported into Alma as MARC-8
- Import Profiles:
 - Encoding format is not configured to UTF-8



Record Characteristics

Bibliographic Records

Record Standard: MARC 21

Character Set: MARC-8

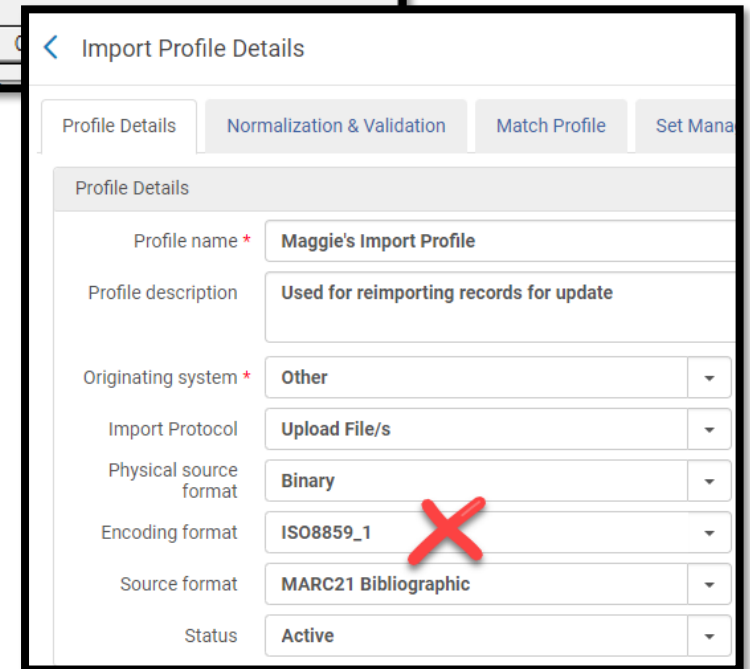
Reset

Authority Records

Record Standard: MARC 21

Character Set: MARC-8

Reset



Import Profile Details

Profile Details

Profile name * Maggie's Import Profile

Profile description Used for reimporting records for update

Originating system * Other

Import Protocol Upload File/s

Physical source format Binary

Encoding format ISO8859_1

Source format MARC21 Bibliographic

Status Active

How Bad/Corrupted Diacritics Are Being Updated for NZ Linked Records

- NZ linked physical bib records are monitored by the Network Zone Coordinator
 - OCLC WorldShare files are monitored daily for corrupted diacritics/special characters
 - All NZ physical bib records are filtered monthly using an indication rule to identify corrupted diacritics/special characters
- NZ Physical bib record with corrupted diacritics/special characters are reimported into Alma daily:
 - OCLC numbers (ocns) are batch imported into Connexion Client
 - OCLC MARC records saved in batch to a local save file
 - OCLC MARC records in the local save file are batch exported to a file
 - OCLC file (.dat) is change to a .mrc file
 - OCLC .mrc file is imported into Alma using an import profile without inventory

IDENTIFYING BAD/CORRUPTED DIACRITICS FOR IZ ONLY BIB RECORDS



The State University
of New York

Identifying Bad/Corrupted Diacritics for IZ Only Bib Records

Bad/Corrupted Diacritics for IZ only records can be identified by:

1. Indication rule
2. Analytics report

Using an indication Rule	Using an Analytics Report
<ul style="list-style-type: none">• Requires a “Cataloger” user role• Can search all MARC fields• Requires a set of records• Requires a filter to be run on the set of records	<ul style="list-style-type: none">• Requires a “Design Analytics” user role• Limited ability to search all MARC fields<ul style="list-style-type: none">○ 5XXs and 9XXs cannot be searched unless they have been configured to Local Params by Ex Libris○ May result in bib records with corrupted characters remaining in the repository• Can be added to the Alma dashboard as a report and/or widget

Identifying Bad/Corrupted Diacritics for IZ Only Bib Records

Using an indication rule:

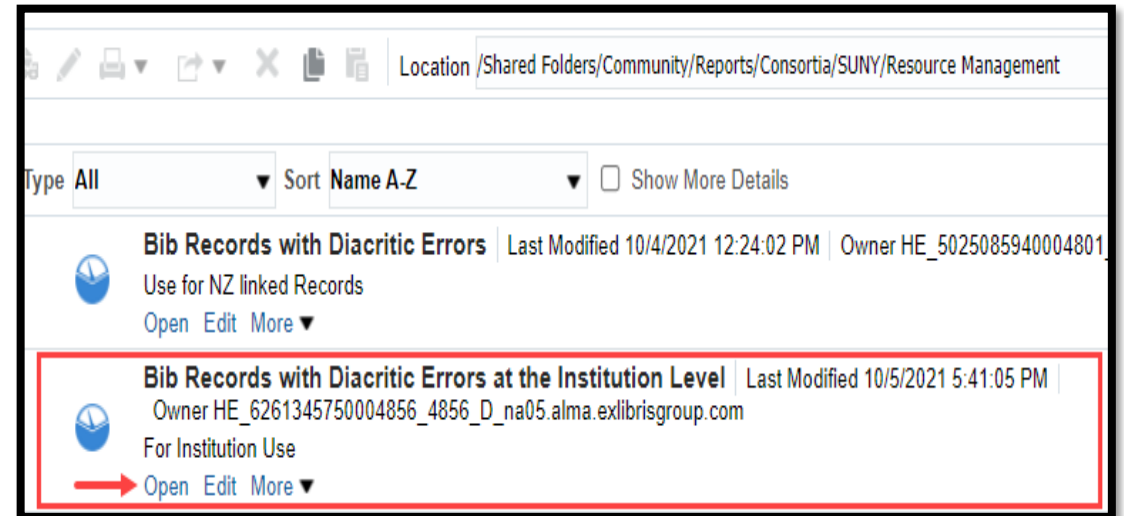
1. Open the MD Editor
2. Check the “placement of new rules” is set to local
 - Click on **Records or Templates tab>New>Placement options**
3. Click on **Rule>Indication**
4. Search for the SUNY indication rule called: *SUNY -Indication Rule for Corrupted Diacritics & Special Characters*
5. Right-click on the indication rule and select **Duplicate**
 1. Edit the Name
 2. Click **Save**
7. Go to **Admin>Manage Jobs and Sets>Manage Sets**
8. Create a set of IZ only physical bib records using a repository search
9. Click **Save and Filter Query**
 1. Name: Name the set
 2. Description: [Optional]
 3. Click **Save and Filter**
 4. Filter: [Select the duplicated rule]
 5. Name: [Edit to identify better]
 6. Click **Filter this set**
 7. Click **Confirm**
10. Review the records
 1. Go to **Admin>Manage Sets and Jobs>Manage Sets**
 2. Click **Members** from the ellipsis for the set

Identifying Bad/Corrupted Diacritics for IZ Only Records

Using an analytics report:

1. Go ***Analytics>Analytics>Design Analytics***
2. Go to the following path: ***/Shared Folders/Community/Reports/Consortia /SUNY/Resource Management/Bib Records with Diacritic Errors at the Institution Level***
3. Click ***Open*** to run the report directly from the path
4. Export the results for review
 - Click ***Export>Formatted>Excel***

NOTE: DO NOT make changes to the analytics reports. If an institution wants to make changes, they need to save the file using “Save as” to the Institutions Reports path.



		996193149704806	65667718	Concerto lyrique : pour clarinet et orchestre : partition d'orchestre /	Gotkovsky, Ida.	G Billaudot	<U+fffd>1979.	3/10/2020 10:49:35 AM	WorldCat	7/27/2021
		996193149804806	732055916	Piezas caracter<U+fffd>isticas [op. 92] /	Alb<U+fffd>eniz, Isaac Komposit, Spanien, 1860 - 1909.	Union Musical Ediciones	[1996]	3/10/2020 10:47:15 AM	WorldCat	7/27/2021
Potsdam Crumb Library	Crumb ON ORDER	996262341904806		ncendary /	C<U+fffd>ordova, Zoraida, author.	Hyperion	2020.	10/1/2021 9:58:05 AM	WorldCat	10/1/2021
		996262342104806		Christmas at Long Pond /	George, William T.	Greenwillow Books	<U+fffd>1992.	10/1/2021 9:46:26 AM	WorldCat	10/1/2021
				The colors we share /	Dass, Ang<U+fffd>elica, 1979- author, photographer.	Aperture	2021.	10/1/2021 9:41:41 AM	WorldCat	10/1/2021

PDF

Excel

Powerpoint

Formatted

Data

Web Archive

Edit - Refresh - Print - Export - Add to Briefing Book - Copy

UPDATING BAD/CORRUPTED DIACRITICS FOR IZ ONLY BIB RECORDS

Updating Bad/Corrupted Diacritics in IZ Only Bib Records

Updating corrupted characters in IZ only bib records:

1. With an OCLC number
2. Without an OCLC number

Updating IZ Only bib records without OCLC Numbers	Updating IZ Only bib records with OCLC Numbers
<ul style="list-style-type: none">• Manually:<ul style="list-style-type: none">○ Make changes manually in the MD Editor	<ul style="list-style-type: none">• Reimport the bib records into the IZ:<ul style="list-style-type: none">○ Export records one by one from OCLC○ Export a set of records to file and import them into Alma using a Repository Import Profile without inventory

Updating IZ Only bib records without OCLC Numbers

To update IZ Only bib records in the MD Editor:

1. Open the bib record in the MD Editor
2. Correct the bad/corrupted diacritics
 - Copy and paste
 - Using the Windows built in character map
3. Go to **Save>Save and Release the Record**

Christ's nativity : an early choral suite by Benjamin Britten / by Edward Lundergan. (990004251540204844)

Record 01 MARC21 Bib

LDR 00687nam##2200193###4500

001 990004251540204844

005 20190818144729.0

008 051214s2004#####000###eng#d

035 \$\$a NEW01000425154

035 \$\$a (Aleph)000425154NEW01

099 \$\$a FacPubs \$\$a Lundergan \$\$a no. 2

100 1 0 \$\$a Lundergan, Edward Joseph, \$\$d 1955-

245 1 0 \$\$a Christ's nativity : \$\$b an early choral suite by Benjamin Britten / \$\$c by Edward Lundergan.

300 \$\$a Pp. 19-30.

440 0 \$\$a SUNY New Paltz faculty publications ; \$\$v Lundergan, no. 2

500 \$\$a Reprinted from: Choral journal, June/July 2008

600 1 \$\$a Britten, Benjamin #d 1913-1976.

935 \$\$a pbook

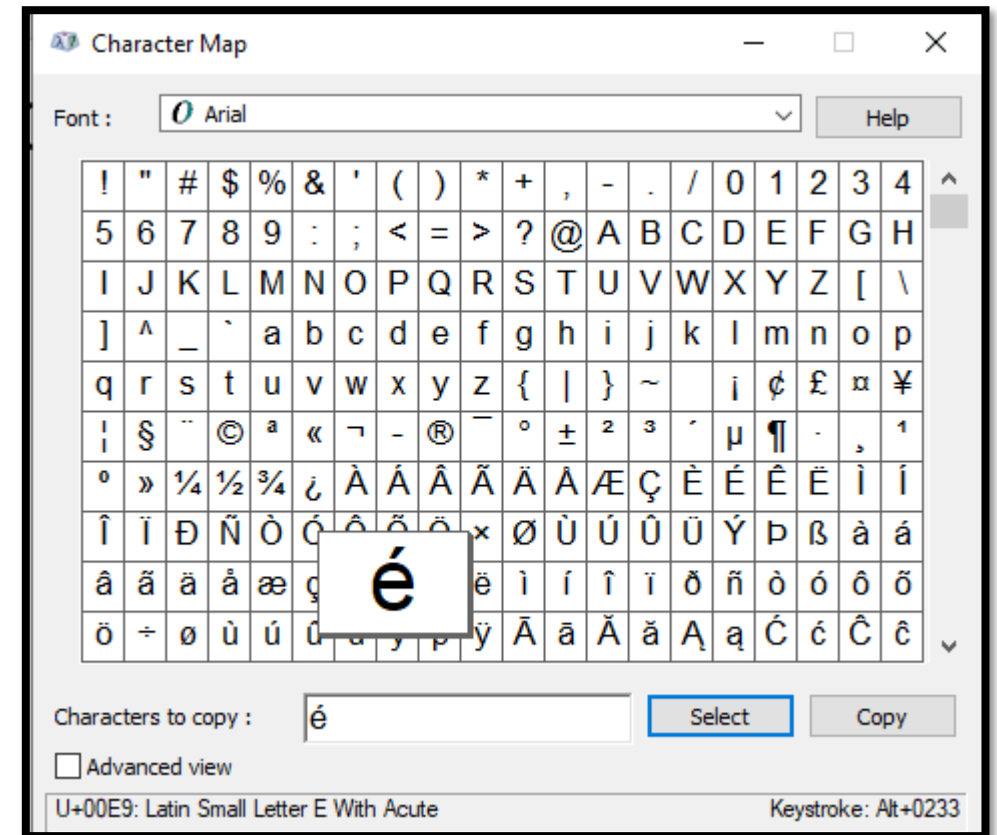
986 \$\$a NEW01000425154

Corrupted Character

Updating IZ Only bib records without OCLC Numbers

Using the Windows built in character map:

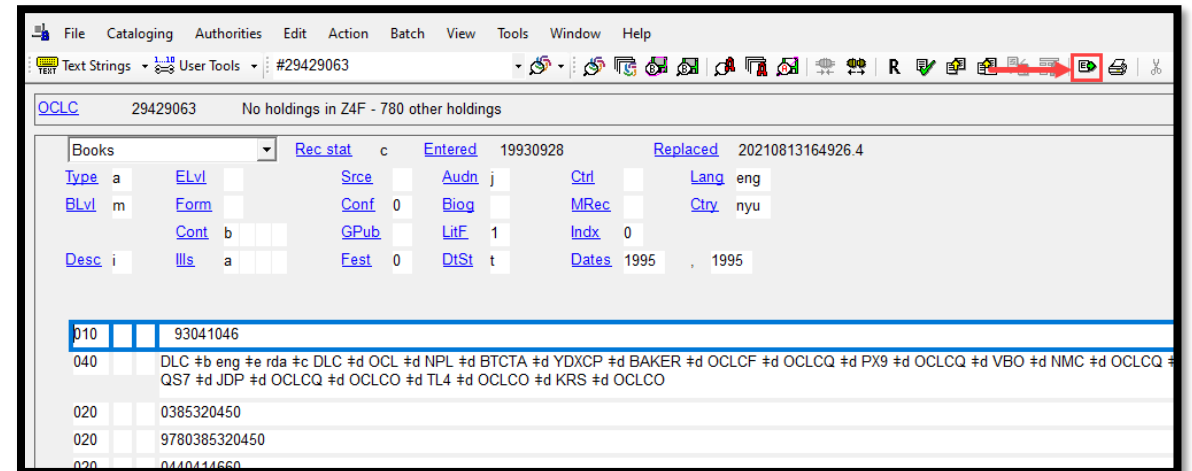
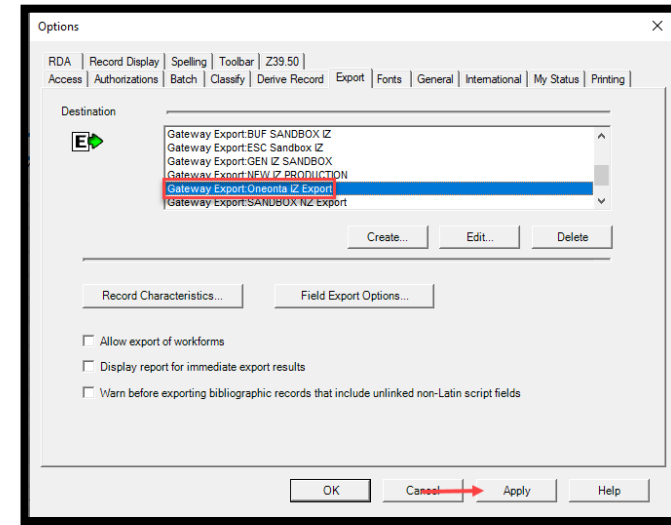
1. Search for Character Map from the Windows search menu
 - Or go to **All Programs>Accessories>System Tools>Character Map** from the Windows start menu
2. Select a font from the Font drop-down list
3. Click on the special character
4. Click **Select**
5. Click **Copy**
6. Paste the character in the MD Editor



Updating IZ Only Bib Records with OCLC Numbers Manually

Reimporting IZ Only bib records manually:

1. Open Connexion Client
 - **Do not** use the Web version as it will not correct the corrupted characters
2. Verify the Character Set is configured to UTF-8
 - Go to **Tools>Options>Export tab>Record Characteristics**
3. Configure the export to the Alma IZ export for your institution
 1. Go to **Tools>Options>Export tab**
 2. Select the Alma IZ export
 3. Click **Apply**
 4. Click **Close**
4. Search for the bib record by OCLC number
5. Export the bib record into the IZ
6. Review the record was updated properly in Alma



Updating IZ Only Bib Records with OCLC Numbers Using an Import Profile

Steps to update IZ only bib record using an Import Profile

1. Create an IZ Repository Import Profile without Inventory
2. Copy the OCLC numbers to a .txt file
3. Batch import the OCLC number to Alma
4. Batch load the MARC records to a local save file
5. Export the MARC records in the local save file to a file
6. Rename the file from .dat to .mrc
7. Import the .mrc file into Alma using the IZ Repository Import Profile



1. Create a Repository Import Profile Without Inventory

1. Go to **Resources>Import>Manage Import Profiles**
2. Click **Add New Profile**
3. Select the radio button next to Repository
4. Click **Next**
5. **Uncheck “Use Network Zone”**
6. Name: [Name the Import Profile]
7. Profile Description: [Add description]
8. Originating System: [Select Other]
9. File Name Patterns: [Type .mrc]
10. Import Protocol: [Select Upload File/s]
11. Physical Source Format: [Select Binary]
12. Encoding Format: [Select UTF-8]
13. Source Format: [Select MARC21 Bibliographic]
14. Status: [Select Active]
15. Click **Next**

The screenshot displays the 'Import Profile Details' form, which is part of a multi-step process indicated by a progress bar at the top (steps 1-6). The form is titled 'Import Profile Details' and includes a 'Save Draft' button and a 'Next' button. The form fields are as follows:

- Use Network Zone:** A checkbox that is unchecked.
- Profile name:** A text input field containing 'Update IZ Only Bib Records'.
- Profile description:** A text input field containing 'Used to update IZ Only bib records'.
- Originating system:** A dropdown menu with 'Other' selected.
- File name patterns:** A text input field containing '.mrc'.
- Import Protocol:** A dropdown menu with 'Upload File/s' selected.
- Physical source format:** A dropdown menu with 'Binary' selected.
- Encoding format:** A dropdown menu with 'UTF-8' selected.
- Source format:** A dropdown menu with 'MARC21 Bibliographic' selected.
- Target format:** A dropdown menu with 'MARC21 Bibliographic' selected.
- Status:** A dropdown menu with 'Active' selected.
- Cross walk:** Radio buttons for 'No' (selected) and 'Yes'.

1. Creating a Repository Import Profile Without Inventory

16. Filter: [Leave blank]
17. Correct the Data Using: [Select SUNY OCLC Normalization Rule]
18. Handle Invalid Data Using: [Select MarcXML Bib Import]
19. Click **Next**

The screenshot shows the 'Import Profile Details' form. At the top right, there are buttons for 'Cancel', 'Save Draft', 'Back', and 'Next'. A progress indicator shows six steps, with step 3 highlighted. The form has three main sections: 'Filter' with a 'Filter out the data using' dropdown, 'Normalization' with a 'Correct the data using' dropdown set to 'SUNY OCLC Normalization Rule', and 'Validation Exception Profile' with a 'Handle invalid data using' dropdown set to 'MarcXML Bib Import'. A red arrow points from the 'Back' button to the 'Next' button.

1. Creating a Repository Import Profile Without Inventory

- 20. Match by Serial Match Method: [Select No]
- 21. Match Method: [Select 035 (Other System Identifier) Match Method]
- 22. System Identifier Prefix: [Type (OCoLC)]
- 23. Handling Method: [Select Automatic]
- 24. Upon Match: [Select Overlay]
- 25. Single match - match only record with the same inventory type (electronic/physical): [Check box]

The screenshot shows a web-based configuration form for a 'Match Profile'. It is divided into two main sections: 'Match Profile' and 'Match Actions', each with a dropdown arrow on the right. The 'Match Profile' section contains: 'Match by Serial / Non Serial' with radio buttons for 'Yes' (unselected) and 'No' (selected); 'Match Method' with a dropdown menu showing '035 (Other System Identifier) Match'; and 'System Identifier Prefix' with a text input field containing '(OCoLC)'. The 'Match Actions' section contains: 'Handling method' with radio buttons for 'Automatic' (selected) and 'Manual' (unselected); 'Upon match' with radio buttons for 'Merge' (unselected), 'Overlay' (selected), 'Do Not Import' (unselected), and 'Import New Record' (unselected); and 'Single match - match only record with the same inventory type (electronic/physical)' with a checked checkbox.



1. Creating a Repository Import Profile Without Inventory

26. Merge Method: [Select SUNY - Overlay All But Local Fields]
27. Select Action: [Check the box for Prefer record with the same inventory type (electronic/physical)]
28. Merge and Combine: [Select Manually]
29. Preferred Record: [Select Record with the most items]
30. Secondary Record: [Select Delete]
31. Merge Method: [Select SUNY - Overlay All But Local Fields]
32. Upon No Match: [Select Do Not Import]
33. Click **Next**

The screenshot displays the 'Automatic Multi-Match Handling' configuration window, which is organized into several sections with expandable headers. The 'Automatic Multi-Match Handling' section is currently expanded, showing three options: 'Disregard matches for bibliographic CZ linked records' (unchecked), 'Disregard invalid/canceled system control number identifiers' (unchecked), and 'Prefer record with the same inventory type (electronic/physical)' (checked). Below this, the 'Merge Records and Combine Inventory for Multi-Match' section is expanded, showing 'Merge and Combine' set to 'Manually', 'Preferred record' set to 'Record with the most items', 'Secondary record' set to 'Delete', 'Merge method' set to 'Overlay all fields but local', and 'Update holdings call number' unchecked. The 'Handle Record Redirection' section is also expanded, showing 'Canceled record field' and 'Canceled record subfield' as empty text boxes, 'Canceled record' set to 'Delete', 'Merge method' set to 'Overlay all fields but local', and 'Update holdings call number' unchecked. Finally, the 'No Match' section is expanded, showing 'Upon no match' set to 'Do Not Import'. A blue circular button with a white upward arrow is located at the bottom right of the form.

1. Creating a Repository Import Profile Without Inventory

34. Synchronize with OCLC: [Select Publish Bib]
35. Condition: [Select Unconditionally]
36. Click **Next**
37. Inventory Operations: [Select None]
38. Click **Save**

Import Profile Details

Cancel Save Draft Back Next

1 2 3 4 5 6

Update IZ Only Bib Records >

Set management tags for all the records imported using this profile

Suppress record/s from publish/delivery ☐ Condition Only for new record

Suppress record/s from external search ☐ Condition Only for new record

Synchronize with OCLC Publish Bibliograph Condition Unconditionally

Synchronize with Libraries Australia Don't publish Condition Only for new record

Import Profile Details

Cancel Save Draft Back Save

1 2 3 4 5 6

Update IZ Only Bib Records >

Inventory Operations

☐ Electronic ☐ Mixed ☐ Physical ☒ None

2. Copy the OCLC Numbers to a .txt File

1. Open a text editor, such as Notepad or Notepad+++
2. Paste the OCLC numbers into the text editor
 - Must not contain the (OCoLC) prefix
3. Click ***File>Save***
4. Select the path where the file will be saved
5. Name the file



3. Batch import the OCLC number to Alma

1. Open Connexion Client
2. Click **Batch>Enter Bibliographic Search Keys**
3. Click **Import**
4. Select the .txt file
5. Click "**No**" when the pop-up screen appears and asks "Do you want to delete your original import profile"
6. Click **Save**
7. Click **Close**



4. Batch Load the MARC Records to a Local Save File

1. Click **Batch>Process Batch**
2. Check the box next to the default bib file path (or to custom path if configured)
3. Check the box next to Online Searches under Process
4. Click **Ok**
 - Files will upload to the save file
5. Batch Search Report will display when the upload is complete
6. Click the "X" to close the Batch Search Report



5. Export the MARC records in the local save file to a file

1. Go to **Tools>Options**
2. Click on the **Export** tab
3. Click on the file export being used
 - Create a file export if one does not already exist
4. Click **Apply**
5. Click **Close**
6. Click **Cataloging>Search>Local Save File** (F3)
7. Use the default file path unless the user selected a customized save file to batch upload
8. Click **OK**
9. The MARC records will display
10. Click the **Export** icon (Actions>Export F5)

6. Rename the File from .dat to .mrc

1. Open the file manager on your computer
2. Go to the path where the file was created based on the file export in OCL
3. Right-click on the file
4. Click ***Rename***
5. Change .dat to .mrc
 - Optional: change the whole name as well as the file extension



7. Import the .mrc File into Alma Using the IZ Repository Import Profile

1. Go to **Resources>Import>Manage Import Profiles**
2. Select **Run** from the ellipses of the import profile created to import OCLC batch files
3. Click **Select File**
4. Click on the file icon to select the .mrc file
5. Click **Add**
6. Click **Submit**
7. The job will run
8. Review the job report after the job is complete



SUPPORTING DOCUMENTATION



The State University
of New York

Supporting Documentation

SUNY:

- [Why are diacritics displaying strangely in Alma and Primo VE from an OCLC record I exported?](#)
- [Our diacritics are not displaying correctly in Primo VE. What should we do?](#)
- [How do batch export bibliographic \(bib\) records from OCLC to Alma?](#)

Ex Libris:

- [Special characters and diacritics appear corrupted in records imported from OCLC Connexion into Alma](#)
- [Alma Indication Rule for Bad Diacritics](#)
- [How to ensure correct display of diacritics in Connexion imports](#)
- [How to Import Special Characters in Alma](#)