

Moving metadata: batch ingesting from Sirsi WorkFlows to the DSpace workspace

Ling He, Digital Services Librarian
York University

OR2013, Friday July 12th, 2013.

Sheet Music Collection digitization project

-
- Digitization efforts are focused on the extensive sheet music collection (approximately 150,000 items) of the late pianist John Arpin (1936-2007). The collection includes examples of Canadian, Broadway, American Standard, Pop music, Jazz and Ragtime from the late nineteenth century to the present day.
- The project is a collaborative effort across different departments in York University library. Key members include a part time music cataloguing librarian, and several part time student digital project assistants.
- The collection is harvested by Sheet Music Consortium through OAI-PMH.

Sheet Music Collection in YorkSpace

[YorkSpace Home](#) → [Clara Thomas Archives and Special Collections](#) → [Sheet Music Collections](#) → [Browsing Sheet Music Collections by Title](#)

Browsing Sheet Music Collections by Title

[0-9](#) [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

Or enter first few letters:

Sort by: Order: Results:

Now showing items 1-20 of 3257

[Next Page](#)



Title: [12th Street rag.](#)
Author: Wheeler, Clarence E. [arranger]; Bowman, Euday L. (Euday Louis) [composer]
Date: 1944



Title: [12th Street rag.](#)
Author: Carle, Frankie [arranger]; Bowman, Euday L. (Euday Louis) [composer]
Date: 1946



Title: [12th Street rag.](#)
Author: Hunt, Pee Wee [performer]; Wheeler, Clarence E. [arranger]; Bowman, Euday L. (Euday Louis) [composer]
Date: 1944



Title: [12th Street rag.](#)
Author: Wheeler, C. E. [arranger]; Bowman, Euday L. [composer]
Date: 1919

Search YorkSpace

- Search YorkSpace
 This Collection

[Advanced Search](#)

Browse

All of YorkSpace

[Communities & Collections](#)

[By Issue Date](#)

[Authors](#)

[Titles](#)

[Subjects](#)

[Series](#)

[Fonds](#)

[Format](#)

[City](#)

[County](#)

[Region](#)

This Collection

[By Issue Date](#)

[Authors](#)

[Titles](#)

[Subjects](#)

[Series](#)

[Fonds](#)

[Format](#)

[City](#)

[County](#)

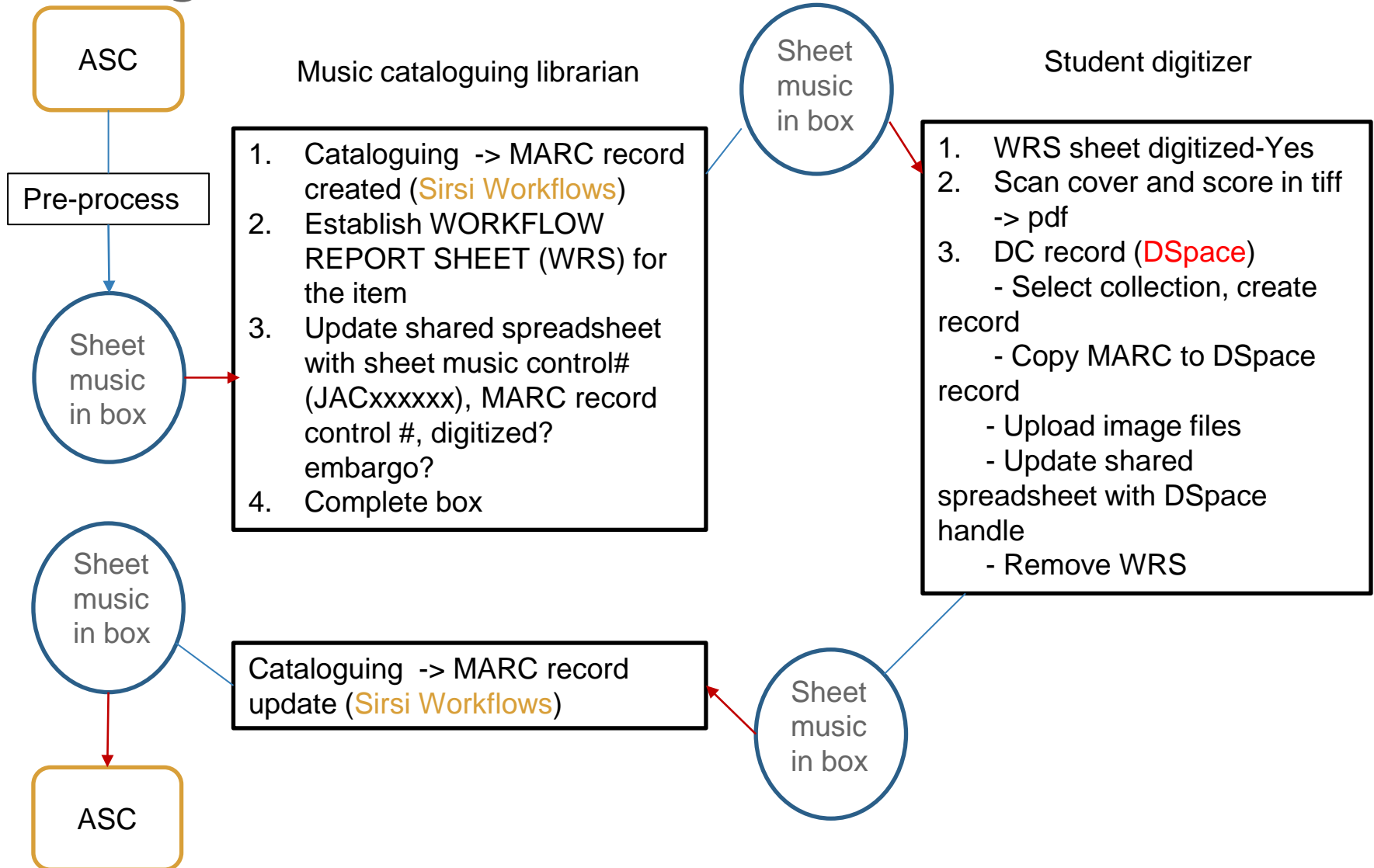
[Region](#)

My Account

[Login](#)


[Register](#)

Original workflow



York Library Online Catalogue record example

Until I say good-bye / lyric by Walter Hirsch ; music by May Hill.

Main Author:	Hill, May.
Corporate Author:	John Arpin Collection - York University.
Other Authors:	Hill, May, composer. Hirsch, Walter, 1891-1967, lyricist.
Format:	 Digitized Score
Language:	English
Published:	Chicago : Roger Graham, 1920
Subjects:	Popular music » United States » 1911-1920. Songs with piano. Sheet music.
Variant Title:	Until I say good-bye.
Language Notes:	English.
Related Resource:	
Digital Bookplate:	Gift of Dr. Mary Jane Esplen; click here to view the bookplate
Item Description:	Gift of Dr. Mary Jane Esplen. Piano vocal [instrumentation] There's a word that I heard from the blue bird [first line] Let your loving eyes shine on sweetheart [first line of chorus] G major [key] Valse lento [tempo] Popular song [form/genre] Publisher's advertisement on back cover [note]
Physical Description:	2 p. of music ; 31 X 24 cm.
Technical Details:	Available on the Internet. MODE OF ACCESS via web browser by entering the following URL: http://hdl.handle.net/10315/22050
Classic Catalogue:	View this record in the classic catalogue

LEADER	01831ccm a2200409 a 4500
008	110518s1920 xxuppz eng d
035	a (Sirsi) a3124715
040	a CaOTY b eng d CaOTY
090	a JAC003240
100	1 a Hill, May.
245	1 a Until I say good-bye / c lyric by Walter Hirsch ; music by May Hill.
246	3 a Until I say good-bye.
260	a Chicago : b Roger Graham, c 1920.
300	a 2 p. of music ; c 31 X 24 cm.
500	a Gift of Dr. Mary Jane Esplen.
500	a Piano vocal [instrumentation]
500	a There's a word that I heard from the blue bird [first line]
500	a Let your loving eyes shine on sweetheart [first line of chorus]
500	a G major [key]
500	a Valse lento [tempo]
500	a Popular song [form/genre]
500	a Publisher's advertisement on back cover [note]
546	a English.
538	a Available on the Internet. MODE OF ACCESS via web browser by entering the following URL: http://hdl.handle.net/10315/22050
592	* * a dom
653	a Love, separation.
650	0 a Popular music z United States y 1911-1920.
650	0 a Songs with piano.
655	a Sheet music.
700	1 a Hill, May, e composer.

YorkSpace Sheet Music Item Submission Form

Item submission



Describe Item

SMC/JAC/ASC Number:

Please obtain your identifier from the folder or item slip. Example: SMC0014

Link to VuFind record:

Please copy and paste the VuFind URL for this record here.

Add

Title (short):

This is the title without the author's name added. Enter the Variant Title or Cover Title as it appears in VuFind. Do not include any square brackets.

Personal author, Performer:

Enter each author in their own box. Do not include comma after first name. Do not include date of birth. If qualifier (ie. lyricist) is provided, enclose in square brackets with one space after name. Example: Bayly, Thomas Haynes [lyricist]

Add

Published:

Copy publisher information exactly as it appears from catalogue record. Leave date intact.

Date:

Get from Publication Info. Only enter date if it exists and is not in brackets. Enter yyyy only. Example: 1889

Physical Description:

Copy directly from catalogue record.

Add

General Note:

Make sure each note is entered separately (in its own box). Do not include anything that has [first line] otherwise, enter everything else as it appears.

Add

Language:

Select the language of the main content of the item. If the language does not appear in the list below, please select "Other". If the content does not really have a language (for example, if it is a dataset or an image) please select "N/A".

#653 in Staff View:

Open "Staff view" in the VuFind record, copy from #653. Put each description (in between commas) into its own box. Capitalize the first word in each box. Do not put a period at the end of the word.

Add

Series name:

If SMC: select Sheet Music Collection If JAC: select John Arpin Sheet Music Collection

- Sheet Music Collection
- John Arpin Sheet Music Collection
- Allan Robb Fleming Digitization Project
- Toronto Telegram
- York 50 Photograph Collection
- From Act to Activism : York50

Type:

Select "Sheet music".

- Animation
- Article
- Book
- Book chapter
- Dataset
- Learning Object

Rights information:

Please select the rights information so that this field is populated.

N/A

Add

Issues

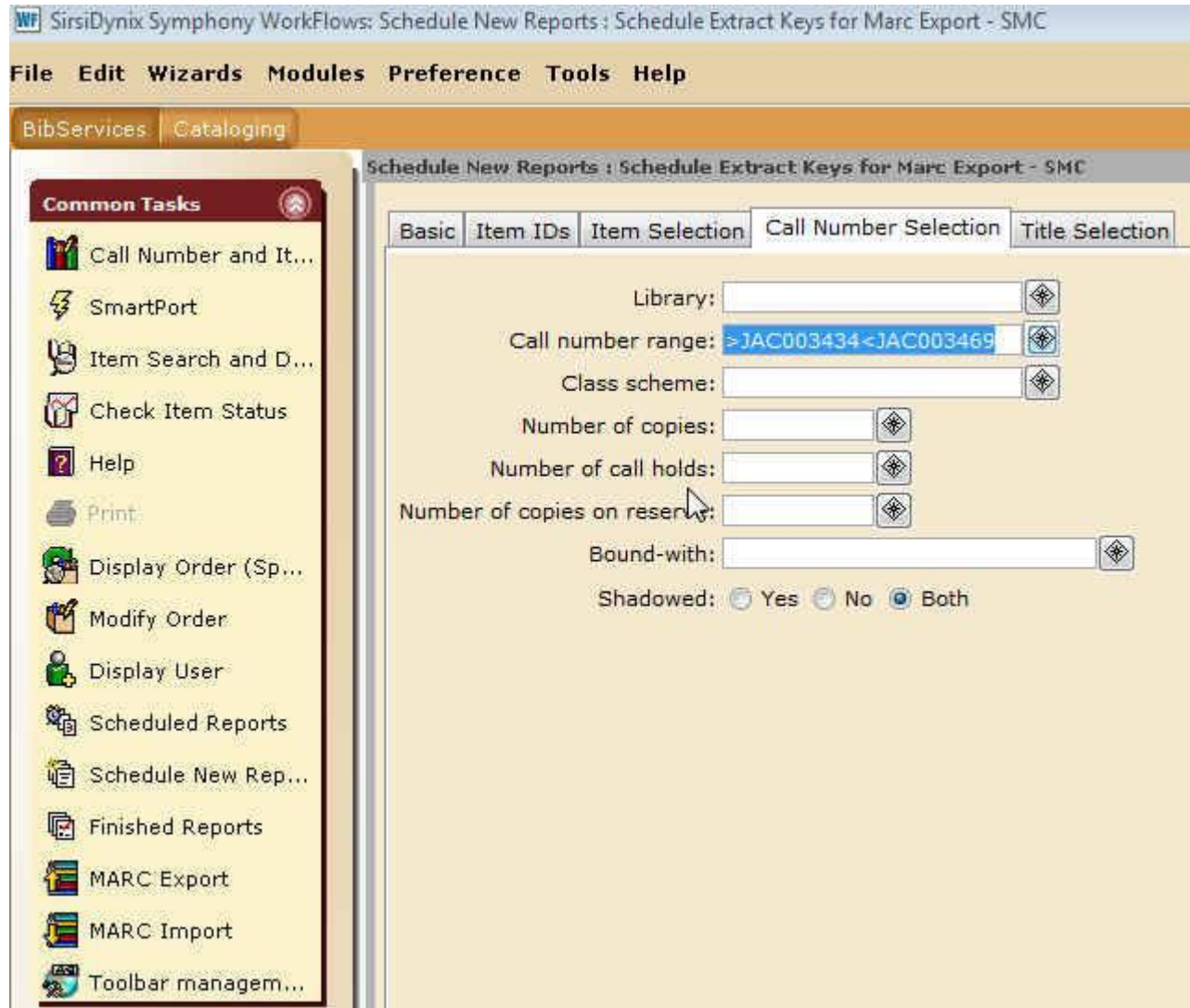
- Inefficient workflow
- Inconsistent DSpace record quality
- Not easy to train new student digital project assistants

We wanted to reuse MARC records and batch ingest digital objects in Dspace!

Challenges

- Limited access to Sirsi WorkFlows
- *Problem:* Can't integrate our tools with Sirsi WorkFlows
- *Solution:* Use MARC Export Utility from Sirsi WorkFlows Client

Export MARC records from Sirsi Workflows



Challenges (con't)

- Catalogue record URL not in MARC record
- *Problem:* DSpace records need editing, can't be ingested into DSpace archive and generate handles directly
- *Solution:* Use SWORD v2 In-Progress HTTP header to enable the item to be deposited into DSpace workspace to add catalogue record URL

MARC to MARCXML Conversion Software

- Use existing open source software
- File_MARC: PHP package, parse or modify existing MARC records read from different sources, and create new MARC records
 - MARC-8 to UTF-8 conversion issue (incorrect display for é, á in a test record Itzés, Tamás [composer])
- MARC4J: JAVA API, read and write MARC and MARCXML, support MARC-8 to UTF-8 conversion

Customized MARCXML to QDC Stylesheet

- Based on The Library of Congress MARCXML to DC Stylesheet

```
<dcterms:type>
  <xsl:text>Sheet Music</xsl:text>
</dcterms:type>

<dcterms:rights>
  <xsl:text>For further copyright information contact : ascproj@yorku.ca</xsl:text>
</dcterms:rights>

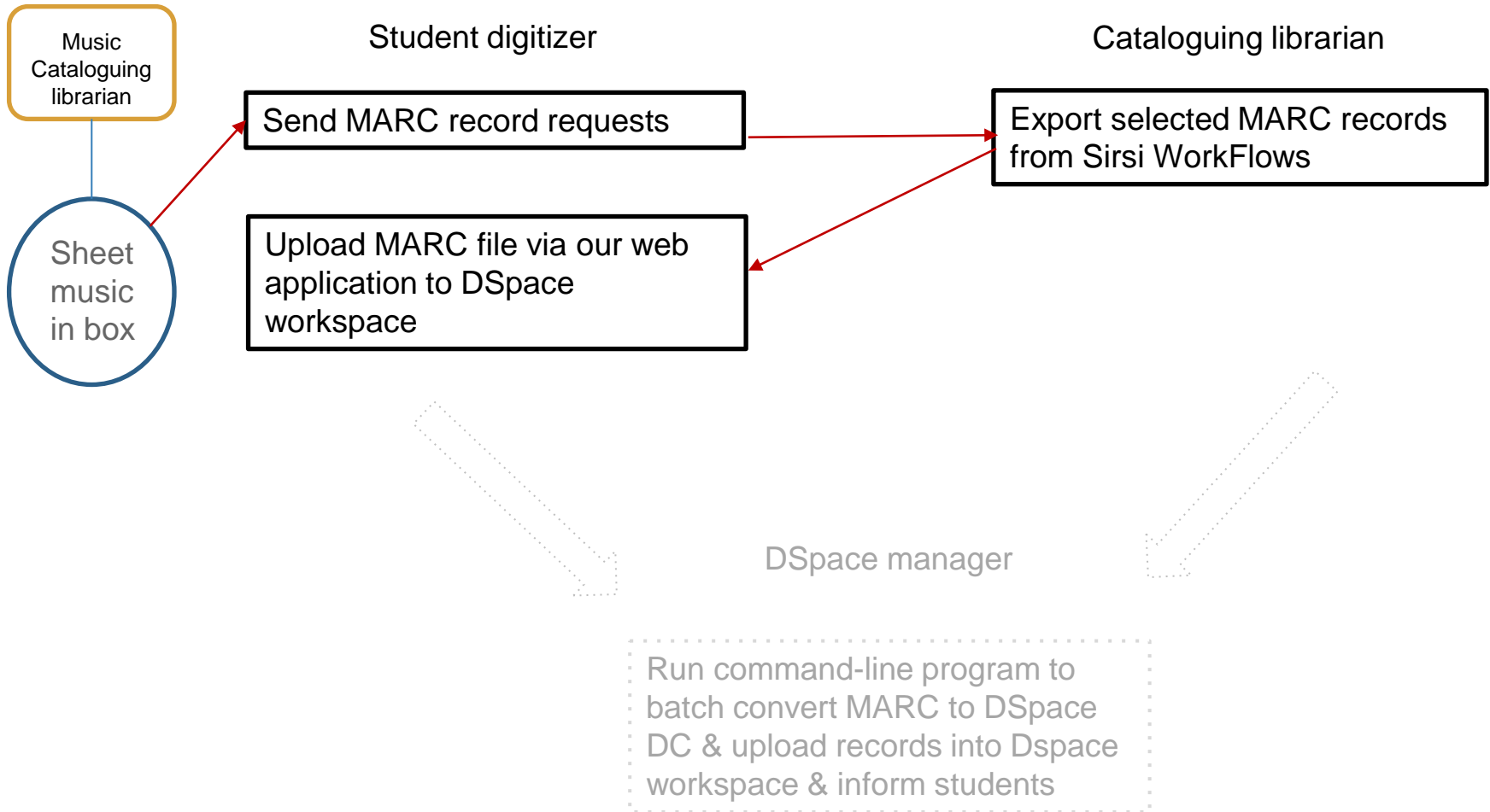
<dcterms:isPartOfSeries>
  <xsl:variable name="identifier">
    <xsl:value-of select="marc:datafield[@tag=090]/marc:subfield[@code='a']"/>
  </xsl:variable>
  <xsl:choose>
    <xsl:when test="substring($identifier, 1, 3)='JAC'">
      <xsl:text>John Arpin Sheet Music Collection</xsl:text>
    </xsl:when>
    <xsl:otherwise>
      <xsl:text>Sheet Music Collection</xsl:text>
    </xsl:otherwise>
  </xsl:choose>
</dcterms:isPartOfSeries>

<xsl:for-each select="marc:datafield[@tag=260]">
  <dcterms:publisher>
    <xsl:call-template name="subfieldSelect">
      <xsl:with-param name="codes">abc</xsl:with-param>
    </xsl:call-template>
  </dcterms:publisher>
</xsl:for-each>
```

Developed program

- Based on The PHP SWORD v2 client library
- A command-line tool to run by the DSpace manager to convert MARC to DSpace Sheet Music QDC and deposit into DSpace workspace
- A web application to allow student digital project assistants to upload the MARC file to deposit into DSpace workspace

Workflow for MARC reuse in DSpace



Developed program

Deposit a MARC file to YorkSpace workspace

Username:

Password:

Upload a MARC file: jac003240_3434.mrc

Deposit completed! View your submissions at [YorkSpace](#)

Click [here](#) to go back to the Upload page

The following records have been deposited:

1. JAC003240
2. JAC003434

YorkSpace Sheet Music Item Submission Form

Item submission



Describe Item

SMC/JAC/ASC Number:

Please obtain your identifier from the folder or item slip. Example: SMC0014

JAC003163

Link to VuFind record:

Please copy and paste the VuFind URL for this record here.

Title (short):

This is the title without the author's name added. Enter the Variant Title or Cover Title as it appears in VuFind. Do not include any square brackets.

Just for old time's sake

Personal author, Performer:

Enter each author in their own box. Do not include comma after first name. Do not include date of birth. If qualifier (ie. lyricist) is provided, enclose in square brackets with one space after name. Example: Bayly, Thomas Haynes [lyricist]

- Scott, S. W. [composer]
- Brock, Anna [composer]

Remove selected

Published:

Copy publisher information exactly as it appears from catalogue record. Leave date intact.

St. Louis, MO : Weile Publishing Co. Inc, 1920.

Date:

Get from Publication Info. Only enter date if it exists and is not in brackets. Enter yyyy only. Example: 1889

1920

Physical Description:

Copy directly from catalogue record.

- 2 p. of music ; 31 X 24 cm.

Remove selected

General Note:

Make sure each note is entered separately (in its own box). Do not include anything that has [first line], otherwise, enter everything else as it appears.

- Gift of Dr. Mary Jane Esplen.
- Piano vocal [instrumentation]
- Secrets of love stars up above shine down on us all the time [first line]
- Just for old time's sake of the love I've known [first line of chorus]
- C major [key]
- Valse Moderato [tempo]
- Popular song [form/genre]
- Man, woman, road, house [illustration]
- Publisher's advertisement on back cover [note]

Remove selected

Language:

Select the language of the main content of the item. If the language does not appear in the list below, please select 'Other'. If the content does not really have a language (for example, if it is a dataset or an image) please select 'N/A'.

N/A

#653 in Staff View:

Open "Staff view" in the VuFind record, copy from #653. Put each description (in between commas) into its own box. Capitalize the first word in each box. Do not put a period at the end of the word.

- Love
- Courtship

Remove selected

Series name:

If SMC: select Sheet Music Collection If JAC: select John Arpin Sheet Music Collection

Sheet Music Collection
John Arpin Sheet Music Collection
Allan Robb Fleming Digitization Project
Toronto Telegram
York 50 Photograph Collection
From Act to Activism : York50

Type:

Select "Sheet music".

Preprint
Presentation
Recording, acoustical
Recording, musical
Recording, oral
Sheet Music

Rights information:

Please select the rights information so that this field is populated.

For further copyright information contact : ascproj@yorku.ca

Benefits & Tradeoffs

Benefits

- Training becomes very easy!
- Efficiency seems to be improved – No more back log!

Tradeoffs

- Extra human resource involved
- Extra program to be maintained

Potential next steps

- Batch ingest digital files in addition to metadata – *Free students from Dspace submission process*
- Integrate with cataloguing system completely
 - *No need to involve extra human resource any more*
- Expand use to other cases or metadata formats

References and resources

The Library of Congress MARCXML framework:

<http://www.loc.gov/standards/marcxml/>

File_MARC: http://pear.php.net/package/File_MARC

MARC4J: <http://marc4j.tigris.org/>

SirsiDynix Symphony WorkFlows: <http://www.sirsidynix.com/symphony>

SWORD v2 PHP client library: <https://github.com/swordapp/swordappv2-php-library/>

York University Libraries online catalogue: <http://www.library.yorku.ca/>

YorkSpace: <http://yorkspace.library.yorku.ca/>

Thank you!

Contact information:

Ling He

Digital Services Librarian

York University Libraries

Scott Library, 4700 Keele St. Room 105D, Toronto ON M3J 1P3

Email: linghe@yorku.ca

Phone: 416-736-2100 x20461