Bulkrax: An Update

Samvera Connect 2020 Lightning Talk
Kiah Stroud
What is Bulkrax?

- Samvera Labs gem
- Batteries-included importer for Samvera applications
- Designed to be extensible
- Provides a full admin interface for imports
  - Creating
  - Editing
  - Scheduling
  - Reviewing

https://github.com/samvera-labs/bulkrax
Statistics

Oct 2019 - Oct 2020
Statistics*

- 187 (non-merge) commits
- +11,240 additions
- -4,556 deletions
- 7 new contributors
- 140 PRs merged

*Oct 2019 - Oct 2020
New Features
● Metadata and files
● Export from:
  ○ Importer
  ○ Collection
  ○ Work Type
● Supported formats:
  ○ CSV
  ○ Others coming soon
● Export errored / upload corrected Entries
Other New Features

- Newly supported file formats for import:
  - XML
  - Bagit
- BrowseEverything integration
- Command Line Interface (CLI)
  - Creating and updating Importers
  - Utilizes JSON API
- Install generator
Improvements and Bug Fixes
<table>
<thead>
<tr>
<th>Name</th>
<th>Status</th>
<th>Last Run</th>
<th>Next Run</th>
<th>Entries Enqueued</th>
<th>Entries Processed</th>
<th>Entries Failed</th>
<th>Entries Deleted Upstream</th>
<th>Total Collection Entries</th>
<th>Total Work Entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>BulkRax Demo</td>
<td>Complete</td>
<td>Oct 21, 2020</td>
<td></td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>25504</td>
</tr>
<tr>
<td>CSV Demo</td>
<td>Complete</td>
<td>Jul 21, 2020</td>
<td></td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>CSV Demo 2</td>
<td>Complete</td>
<td>Jul 21, 2020</td>
<td></td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>CSV Demo 3</td>
<td>Complete</td>
<td>Jul 27, 2020</td>
<td></td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>CSV Importer Test</td>
<td>Complete</td>
<td>Jul 27, 2020</td>
<td></td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Demo</td>
<td>Complete</td>
<td>Jul 28, 2020</td>
<td></td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>25456</td>
</tr>
<tr>
<td>ETD</td>
<td>Complete</td>
<td>Jul 20, 2020</td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>6568</td>
</tr>
<tr>
<td>ETD New</td>
<td>Complete</td>
<td>Jul 20, 2020</td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>I192 Deploy</td>
<td>Complete</td>
<td>Aug 27, 2020</td>
<td></td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>OER Full</td>
<td>Complete</td>
<td>Jul 20, 2020</td>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3304</td>
</tr>
</tbody>
</table>

Showing 1 to 10 of 18 entries
**Importer: Bulkrax Demo**

**Name:** Bulkrax Demo

**Admin set:** 40626b12-36d3-4116-991b-7e042a3c83fc

**User:** support@notch8.com

**Frequency:** PT05

**Parser klass:** Bulkrax::CsvParser

**Limit:**

**Successed At:** Item has not yet been imported successfully

**Parser fields:**

**Field mapping:**

**Total Works:** 25504

**Total Collections:** 1

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Collection</th>
<th>Entry ID</th>
<th>Status</th>
<th>Errors</th>
<th>Status Set At</th>
</tr>
</thead>
<tbody>
<tr>
<td>oer_3</td>
<td>85440590-456a-403d-b8b1-b7f73e77eaf</td>
<td>102</td>
<td>✓ Complete</td>
<td></td>
<td>2020-10-21 00:05:50 UTC</td>
</tr>
<tr>
<td>oer_1</td>
<td>85440590-456a-403d-b8b1-b7f73e77eaf</td>
<td>100</td>
<td>✓ Complete</td>
<td></td>
<td>2020-10-21 00:05:50 UTC</td>
</tr>
<tr>
<td>oer_5</td>
<td>85440590-456a-403d-b8b1-b7f73e77eaf</td>
<td>104</td>
<td>✓ Complete</td>
<td></td>
<td>2020-10-21 00:05:50 UTC</td>
</tr>
<tr>
<td>oer_4</td>
<td>85440590-456a-403d-b8b1-b7f73e77eaf</td>
<td>103</td>
<td>✓ Complete</td>
<td></td>
<td>2020-10-21 00:05:50 UTC</td>
</tr>
<tr>
<td>oer_2</td>
<td>85440590-456a-403d-b8b1-b7f73e77eaf</td>
<td>101</td>
<td>✓ Complete</td>
<td></td>
<td>2020-10-21 00:05:50 UTC</td>
</tr>
</tbody>
</table>

**1 - 5 of 5**

[Edit | Back]
Identifier: oer_1
Entry ID: 49
Type: Oer

Raw Metadata:

Parsed Metadata:

Succeeded At: 2020-08-27 21:57:33 UTC
Importer: i192 Deploy
Oer Link: Oer
Improvements and Bug Fixes

- Delete Works on import if their records upstream are deleted
- Import handles child relationships
  - Works in Collections
  - Works in Works
- Continuous Integration (CI)
  - Tests
  - Lint
- Improved documentation
  - GitHub Wiki
    - Configuration and customization guides
    - Specific information for importers
    - CLI / JSON API docs
Coming Soon
Coming Soon

- Imported cloud files follow redirects
- Export to MARC
- Roundtrip export and import
- More documentation
- Continued improvements and bug fixes
Thank you!

Kiah Stroud
Web Developer - Notch8
kiah@notch8.com