

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>

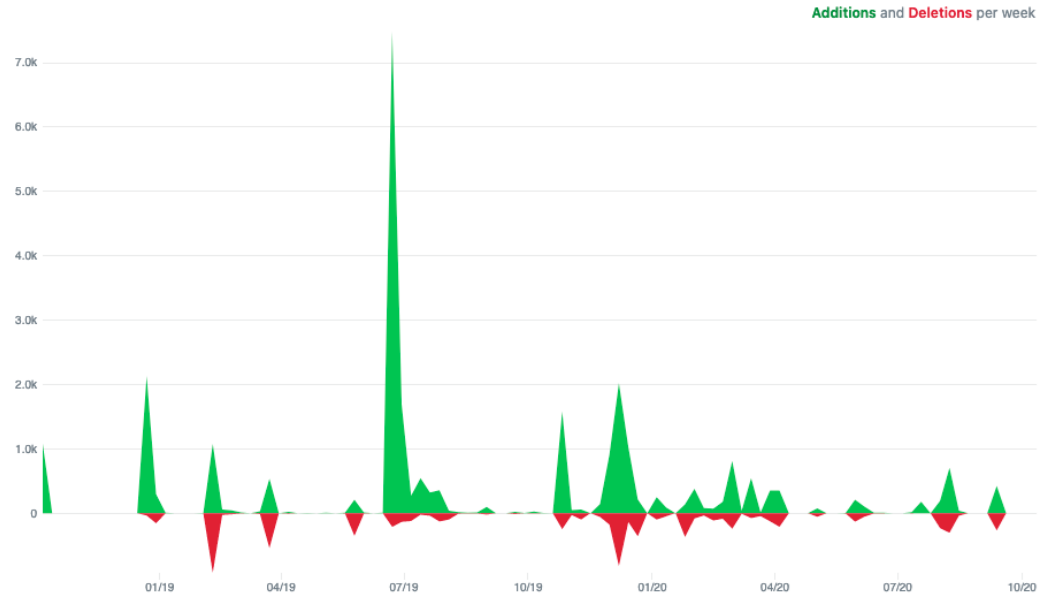
Statistics

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

Export

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

master bulkrax / app / parsers / bulkrax / Go to file Add file

smithjp Add row param to solr queries for exports (#253) ✓ bb45f90 3 days ago History

..		
application_parser.rb	Status overhaul (#237)	last month
bagit_parser.rb	Status overhaul (#237)	last month
csv_parser.rb	Add row param to solr queries for exports (#253)	3 days ago
oai_dc_parser.rb	Status overhaul (#237)	last month
oai_qualified_dc_parser.rb	result of rubocop auto-correct	11 months ago
xml_parser.rb	Status overhaul (#237)	last month

Improvements and Bug Fixes

support@notch8.com

[Home](#) / [Dashboard](#) / Importers

Importers

[New](#)

Show 10 entries

 Search:

Name	Status	Last Run	Next Run	Entries Enqueued	Entries Processed	Entries Failed	Entries Deleted Upstream	Total Collection Entries	Total Work Entries				
Bulkrax Demo	Complete	Oct 21, 2020		0	6	0	0	1	25504	i	p	x	
CSV Demo	Complete	Jul 21, 2020		0	0	2	0	0	2	i	p	x	
CSV Demo 2	Complete	Jul 21, 2020		0	0	2	0	0	2	i	p	x	
CSV Demo 3	Complete	Jul 27, 2020		0	0	2	0	0	2	i	p	x	
CSV Importer Test	Complete	Jul 27, 2020		0	0	3	0	0	3	i	p	x	
Demo	Complete	Jul 28, 2020		1	2	5	0	2	25456	i	p	x	
ETD	Complete	Jul 20, 2020		0	1	2	0	1	6568	i	p	x	
ETD New	Complete	Jul 20, 2020		0	1	2	0	1	2	i	p	x	
I192 Deploy	Complete	Aug 27, 2020		0	2	5	0	2	5	i	p	x	
OER Full	Complete	Jul 20, 2020		0	1	1	0	1	3304	i	p	x	

Showing 1 to 10 of 18 entries

 Previous [1](#) [2](#) Next

ACTIVITY

[Repository Activity](#) >

[Your activity](#) >

[Reports](#)

REPOSITORY CONTENTS

[Collections](#)
[Works](#)
[Importers](#)
[Exporters](#)

TASKS

[Review Submissions](#)
[Manage Users](#)
[Manage Groups](#)
[Manage Embargoes](#)

ACTIVITY

[Repository Activity](#) >

[Your activity](#) >

[Reports](#)

REPOSITORY CONTENTS

[Collections](#)
[Works](#)
[Importers](#)
[Exporters](#)

TASKS

[Review Submissions](#)
[Manage Users](#)
[Manage Groups](#)
[Manage Embargoes](#)
[Manage Leases](#)

CONFIGURATION

[Settings](#) >

[Workflow Roles](#)

Importer: Bulkrax Demo

Name: Bulkrax Demo

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

User: support@notch8.com

Frequency: PT0S

Parser class: Bulkrax::CsvParser

Limit:
Succeeded At: Item has not yet been imported successfully

 Parser fields:
 [+](#)

 Field mapping:
 [+](#)
Total Works: 25504

Total Collections: 1

[Work Entries](#)
[Collection Entries](#)

Identifier	Collection	Entry ID	Status	Errors	Status Set At	Actions
oer_3	854405f0-456a-40d4-b81b-67f734e7feaf	102	✓ Complete		2020-10-21 00:05:50 UTC	i
oer_1	854405f0-456a-40d4-b81b-67f734e7feaf	100	✓ Complete		2020-10-21 00:05:50 UTC	i
oer_5	854405f0-456a-40d4-b81b-67f734e7feaf	104	✓ Complete		2020-10-21 00:05:50 UTC	i
oer_4	854405f0-456a-40d4-b81b-67f734e7feaf	103	✓ Complete		2020-10-21 00:05:50 UTC	i
oer_2	854405f0-456a-40d4-b81b-67f734e7feaf	101	✓ Complete		2020-10-21 00:05:50 UTC	i

1 - 5 of 5

[Edit](#) | [Back](#)

Succeeded At: Item has not yet been imported successfully

Parser fields:



Field mapping:



Full

Raw

title:

```
{"from"=>["title"], "split"=>true}
```

creator:

```
{"from"=>["creator"], "split"=>true}
```

keyword:

```
{"from"=>["keyword"], "split"=>true}
```

description:

```
{"from"=>["description"], "split"=>true}
```

subject:

```
{"from"=>["subject"], "split"=>true}
```

license:

```
{"from"=>["license"], "split"=>"\\|"}
```

contributor:

```
{"from"=>["contributor"], "split"=>true}
```

publisher:

```
{"from"=>["publisher"], "split"=>true}
```

date_created:

```
{"from"=>["date_created"], "split"=>true}
```


support@notch8.com

[Home](#) / [Dashboard](#) / [Importers](#) / [i192 Deploy](#) / 49

ACTIVITY

[Repository Activity](#) >[Your activity](#) >[Reports](#)

REPOSITORY CONTENTS

[Collections](#)[Works](#)[Importers](#)[Exporters](#)**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