

Drupal 8 Migrate API

Bulk-Loading Content into Islandora 8

IslandoraCon 2019

Seth Shaw

The Drupal 8 Migrate API...

"...provides services for migrating data from a source system to Drupal 8."

"Source Systems" include things like...

- Islandora 7  islandora

- A pile of scans and a CSV



- A REST API 

A Migration Configuration

.. defines an Extract, Transform, Load (ETL) process.

- Source plugins extract data from a source
- Process plugins transform the data
- Destination plugins load the data (create new entities)

Examples can be found at

<https://github.com/seth-shaw-unlv/claw-migrate-files-poc>

Migration Metadata

```
id: auth_person
label: Import MADS PersonalName Authorities
migration_group: authorities
```

Source

```
source:
  plugin: url
  data_fetcher_plugin: file
  urls: 'data/apollo-auths.rdf' # Path relative to Drupal site root
  data_parser_plugin: xml
  item_selector: /rdf:RDF/madsrdf:PersonalName
  fields:
    -
      name: uri
      selector: '@rdf:about'
    -
      name: pref_label
      selector: 'madsrdf:authoritativeLabel'
  ids:
    uri:
      type: string
  constants:
    source: lcnaf
```

Process

```
process:
  vid:
    plugin: default_value
    default_value: person
  name: pref_label
  field_authority_link/uri: uri
  field_authority_link/source: constants/source
```

Destination

```
destination:
  plugin: entity:taxonomy_term
```

Easy to run once configured!

```
$ drush migrate:status --group=authorities
```

Group	Migration ID	Status	Total	Imported	Unprocessed	Last Imported
authorities (authorities)	auth_complex	Idle	1	0	1	
authorities (authorities)	auth_geographic	Idle	1	0	1	
authorities (authorities)	auth_person	Idle	2	0	2	
authorities (authorities)	auth_topic	Idle	1	0	1	

```
$ drush migrate:import --group=authorities
```

```
[notice] Processed 1 item (1 created, 0 updated, 0 failed, 0 ignored) - done with 'auth_complex'  
[notice] Processed 1 item (1 created, 0 updated, 0 failed, 0 ignored) - done with 'auth_geographic'  
[notice] Processed 2 items (2 created, 0 updated, 0 failed, 0 ignored) - done with 'auth_person'  
[notice] Processed 1 item (1 created, 0 updated, 0 failed, 0 ignored) - done with 'auth_topic'
```

```
$ drush migrate:status --group=authorities
```

Group	Migration ID	Status	Total	Imported	Unprocessed	Last Imported
authorities (authorities)	auth_complex	Idle	1	1	0	2019-09-18 15:49:42
authorities (authorities)	auth_geographic	Idle	1	1	0	2019-09-18 15:49:42
authorities (authorities)	auth_person	Idle	2	2	0	2019-09-18 15:49:42
authorities (authorities)	auth_topic	Idle	1	1	0	2019-09-18 15:49:42

Creating Islandora 8 Repository Items

Repository Items are composed of three parts:

1. File (the actual file)
2. Node (descriptive metadata)
3. Media (links Nodes and Files; stores technical metadata)

Group	Migration ID	Status	Total	Imported	Unprocessed	Last Imported
apollo (apollo)	islandora8_file	Idle	7	0	7	
apollo (apollo)	islandora8_metadata	Idle	7	0	7	
apollo (apollo)	islandora8_media	Idle	7	0	7	

```
## islandora8_media snippet ##
```

```
migration_dependencies:
```

```
  required:
```

- islandora8_files # so we can lookup the value for field_media_file
- islandora8_metadata # so we can lookup the value for field_media_of

Beyond simple mappings...

```
## snippet from islandora8_metadata ##
```

```
temp_subjects_person: # Temporary array of person entity reference identifiers
-
  plugin: skip_on_empty # Don't bother if the source field is blank
  source: subject_person
  method: process # Only skip this field, not the whole record (CSV row)
-
  plugin: explode # Account for multiple entries in a cell delimited by ;
  delimiter: ';'
-
  plugin: entity_generate # Make a new entity if you can't find a match
  value_key: name
  bundle_key: vid
  bundle: person
  entity_type: taxonomy_term
  default_values:
    vid: person
```

continued next slide...


```
temp_subjects: # Temporary field of subjects
```

- - plugin: skip_on_empty
 - source: subjects
 - method: process
- - plugin: explode
 - delimiter: ';'
- - plugin: entity_generate
 - value_key: name
 - bundle: subject
 - bundle_key: vid
 - entity_type: taxonomy_term
 - default_values:
 - vid: subject

```
field_subjects: # Gather temp arrays into the destination field
```

- - plugin: get
 - source:
 - '@temp_subjects_person'
 - '@temp_subjects'
- # an array of arrays to a flat array of entity refs
 - plugin: flatten

Make your own plugins!

ListToRange: e.g. “1991; 1992; 1993;” -> “1991/1993”

```
public function transform($value, MigrateExecutableInterface $migrate_executable,
    Row $row, $destination_property){

    // Config and sanity checking trimmed for brevity...

    $list = explode($this->configuration['source_delimiter'], $value, PHP_INT_MAX);
    if ( count($list) < 2 ){
        return $list[0];
    } else {
        return implode( $this->configuration['range_delimiter'],
            [trim($list[0]), trim($list[count($list)-1])]
        );
    }
}
```

Why use the Migrate API?

- You can (potentially) do everything with configs!
- Leverage contrib module plugins.
- Making plugins for more complex sources and processes is (relatively) simple.
- Updating metadata is as simple as:

```
`drush mim islandora8_metadata --update`
```