

FEEDS UI + MIGRATE ENGINE = DREAM MIGRATIONS AND IMPORTS

Work in progress

Irina Zaks izaks@stanford.edu

<http://fibonaccwebstudio.com/>

In this presentation

- Content migration tools for Drupal 8 in 2018
- Comparison of Feeds and Migrate
- Feeds_Migrate – new module in development
- How to contribute

WARNING!



- This demo uses dev tools
- Things might and will break
- Interface and features can and will change without notification

STAY CALM and DRUPAL
ON



Website is supporting mission and operation of your organization

- Websites are living creatures
- Websites change, grow, and shrink as do organizations that these sites support
- New content is generated inside and outside of your organization and needs to be added or imported into your site

Moving content into your website:

What

- Importing structured data (.csv, xml, json)
- Importing from
 - Other websites
 - Documents
 - Database (s)
- One time import / Periodic import
- Update, replace or preserve previously imported content items
- Monitor import state (last import, # of items, etc.)

**IF YOU COULD GET THOSE
NODES INTO THE SYSTEM BY MONDAY**



THAT'D BE GREAT

Moving content into Drupal 8:

How

- New Drupal 8 Website - "One-click " upgrade via web interface
Configuration > Development->Upgrade
- Set up migrations via drush/yml files and manage in Migrations tab Structure-> Migrations
- Feeds Module – build UI and import data
- Custom scripts



Ready for
Moving

Examples

- D7 site - <https://live-kipac7-clean.pantheonsite.io/>
- New site - <http://test-drupal-8-content-migration.pantheonsite.io>
- Migrate
http://test-drupal-8-content-migration.pantheonsite.io/admin/structure/migrate/manage/migrate_drupal_7/migrations
- Feeds
 - <http://test-drupal-8-content-migration.pantheonsite.io/admin/structure/feeds>
 - <http://test-drupal-8-content-migration.pantheonsite.io/feed/2>

Moving Data

Data Source

- Data structure
- Actual data

Data Processing / Tampering

- Reformat date (to ISO)
- Change relationships (text list to term reference)
- Explode array (from csv list)
- Change text case
- Find/replace
- Other transformations

Destination container

- Entity type
- Fields

D7 Migrations - Migrate / Feeds

Feeds - **configured via web UI**

1 - Create Feed type

- Select Processor (node, user, term, etc)
- Select Parser
- Select Source type (upload file/url)
- Add Feed type specific properties

2 - Create Feed Import

- Add actual data source (upload file or enter url)
- Import data
- Delete data

Migrate - **coded in .yaml files**

- Define data source (file)
- Define processors
- Define destination

If data source is changed new migration has to be set up

Operations

- Run import
- Stop import
- Rollback - delete all imported data

New in Drupal 8 - Migrate Module in Core

Migrate Code is in D8 Core

Migrate in core

- Parsers
- Processors
- Missing - good UI

Feeds

- Great UI
- Feeds-specific features
- Need more parsers and tamperers

Why code and maintain more parsers and processors for feeds if we can use what is in core

Feeds-specific features not handled by Migrate

- periodic import
- unpublish/delete nodes not in feed
- expire/update items that are already on the target site.
- target configuration
- Pubhubsubbub
- etc.

Current status of Feeds

- Most recent release Sept 6, 2018
- CSV import works out of the box , no json and xml
- Tamper plugins can be extended
- Join discussion Feeds weekly meetup

<https://drupal.slack.com/messages/C34CECZAL/details/>

every Thursday 11 am PST

Let's code new module that
will have Feeds UI that
connects to Migrate engine

Feeds Migrate Importer (Type) - Drupal Structure Basic UI

Create new "Feeds Migrate Importer" and define destination of
import - Drupal Structure

https://www.drupal.org/project/feeds_migrate/issues/3002

<http://dev-feeds8.pantheonsite.io/admin/structure/migrate/manage/default/migrations>

▼ BASIC SETTINGS

Name *

A unique label for this feed type. This label will be displayed in the interface.

Description

A description of this feed type.

Fetcher ▼

Parser ▼

Processor ▼

Content type *

 ▼

Fetcher settings

Parser settings

Processor settings

Settings

Migration Group

 ▼

Choose how often a feed should be imported.

Save and add mappings

Add and Basic Config for Feed Migration Mapper

<admin/structure/feeds-migrate/sources/add>

- Name/description
- Select fetcher
- Select parser
- Select processor
- Add authentication option
- **Migration group** should be set by default to a group "Default migrations" group and set in Settings tab under all other tabs

Home » Administration » Structure » Feed types

▼ BASIC SETTINGS

Name *

A unique label for this feed type. This label will be displayed in the interface.

Description

A description of this feed type.

Fetcher
Download ▼

RSS/Atom
Sitemap XML
OPML
✓ CSV

Processor
Node ▼

Content type *
Article ▼

Add fetcher settings - fetcher-specific

Fetcher settings	Allowed file extensions <input type="text" value="txt csv tsv xml opml"/> Allowed file extensions for upload.
Parser settings	
Processor settings	Upload directory * <input type="text" value="private://feeds"/> Directory where uploaded files get stored. Prefix the path
Settings	

Parser settings - parser-specific

Settings	
Fetcher settings	
Parser settings	Default delimiter <input type="text" value=","/> Default field delimiter.
Processor settings	<input type="checkbox"/> No headers Check if the imported CSV file does not start with

Bottom Settings tab with Migration Group set by default

Fetcher settings	Migration Group
Parser settings	Default ▼
Processor settings	Choose how often a feed should be imported.
Settings	

Authentication options

If feed is pulling from url it might need authentication
username/password or token fields

- Integration with OAuth?
- Other options? Guzzle in core?

Mapping UI

Configure import structure manually or drop file/point to url - map import sources to fields in destination

https://www.drupal.org/project/feeds_migrate/issues/3002

Mapping sources (current interface)

Manually add name (csv) or json string or xml path. and map to fields for content type

<http://dev-feeds8.pantheonsite.io/admin/>

[structure/feeds/manage/article_import_from_url/mapping](http://dev-feeds8.pantheonsite.io/admin/structure/feeds/manage/article_import_from_url/mapping)

SOURCE	TARGET	SUMMARY	CONFIGURE	UNIQUE	REMOVE
Defendant	Imported Defendant(s) list:			<input type="checkbox"/>	<input type="checkbox"/>
Plaintiff	Imported Plaintiff(s) list:			<input type="checkbox"/>	<input type="checkbox"/>
Case ID	Unique Imported Case ID: This value is used to identify case for import feeds:			<input checked="" type="checkbox"/>	<input type="checkbox"/>
- Select a source -	Feeds item: Item URL:			<input type="checkbox"/>	<input type="checkbox"/>
Case ID	Feeds item: Item GUID:			<input type="checkbox"/>	<input type="checkbox"/>
Included in Random Sample	Include in Random Sample: Included in random sample (we have released a 20% random sample of the dataset and we do want this indicator included to show that the case is or isn't part of that sample):			<input type="checkbox"/>	<input type="checkbox"/>
✓ - Select a source - Case ID Case Title Civil Action Civil Action # Custom Source Defendant Filing Date Included in Random Sample New source... Plaintiff Termination Date	Alleged infringer: - Select a target -	Reference by: Title Autocreate terms: No	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Mapping sources - wish list

Similar to migrate-upgrade config only command, does not import data, but creates import configuration

Pre-populate existing source drop down

Upload file or enter link to predefine source

⚠ Your changes will not be saved until you click the Save button at the bottom of the page.

SOURCE	TARGET	SUMMARY	CONFIGURE	UNIQUE	REMOVE
Title	Title:			<input checked="" type="checkbox"/>	<input type="checkbox"/>
Email	Body: Text	Format: Plain text			<input type="checkbox"/>
Email	Body: Summary				<input type="checkbox"/>
- Select a source -	Authored by: The username of the content author.	Please select a field to reference by. Autocreate terms: No			<input type="checkbox"/>

- Select a target -

Save

Tampering sources

https://www.drupal.org/project/feeds_migrate/issues/3002362

List of Tamperers / processors

Search event

Reset

Show row weights

DESCRIPTION	PLUGIN	STATUS	OPERATIONS	ENABLED
↳ Substring	Substring	Default	Override	<input checked="" type="checkbox"/>
+ Add plugin				

xpathparser:10 -> Section Information:Day(s) of Week

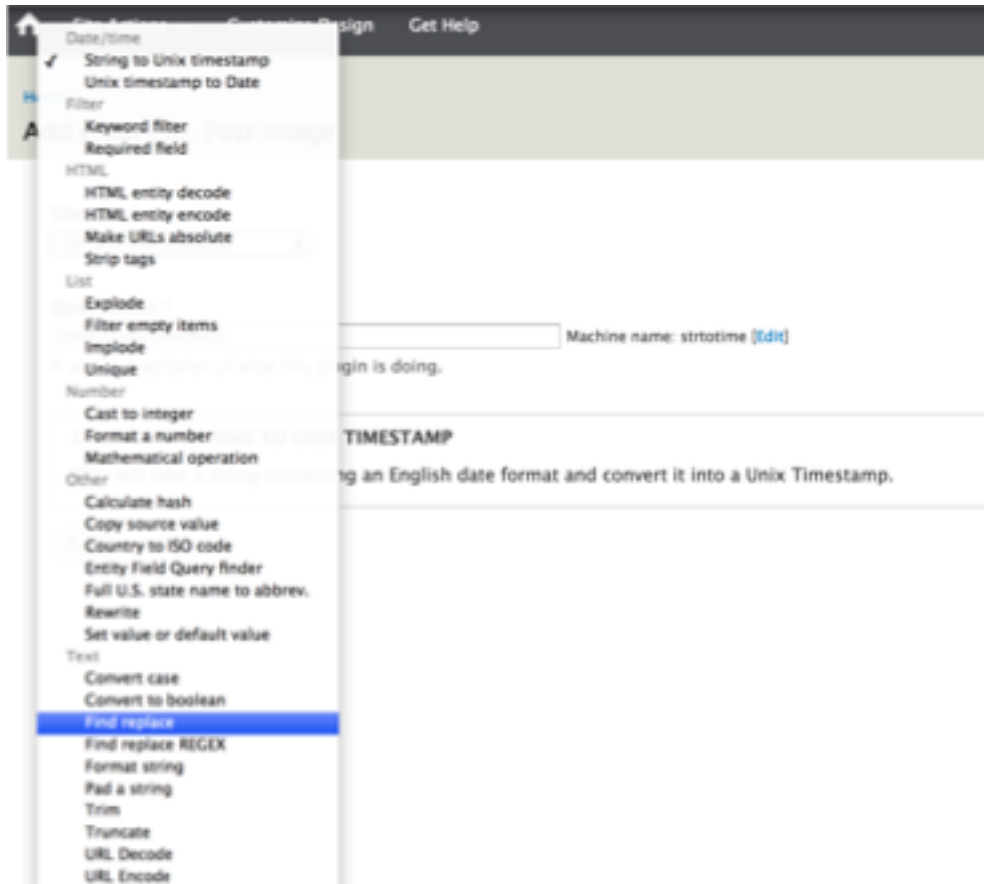
Show row weights

DESCRIPTION	PLUGIN	STATUS	OPERATIONS	ENABLED
↳ Trim whitespace from beginning and end	Trim	Default	Override	<input checked="" type="checkbox"/>
↳ Find replace whitespace via regex	Find replace REGEX	Default	Override	<input checked="" type="checkbox"/>
+ Add plugin				

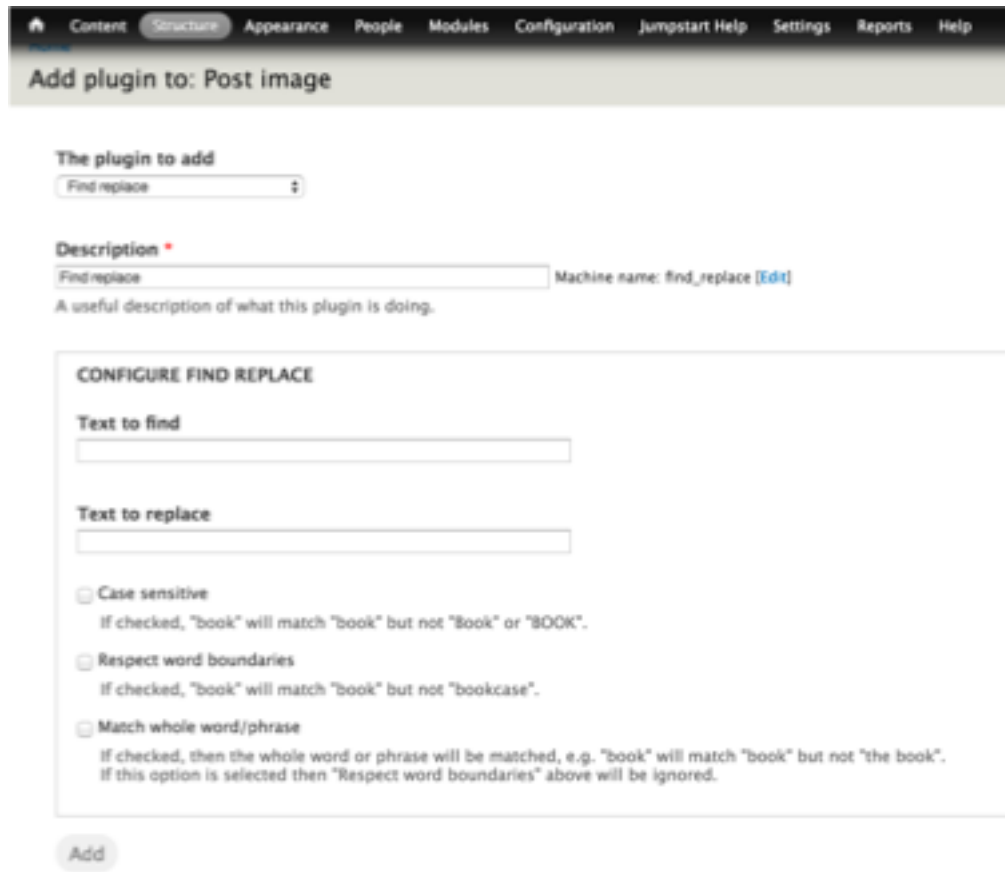
xpathparser:11 -> Section Information:Location(s)

DESCRIPTION	PLUGIN	STATUS	OPERATIONS	ENABLED
No plugins defined.				
+ Add plugin				

Adding tamper/process plugin



Configure each plugin with specific settings



The screenshot shows the Drupal administration interface for configuring a plugin. At the top, a navigation bar includes links for Content, Structure, Appearance, People, Modules, Configuration, Jumpstart Help, Settings, Reports, and Help. Below this, a header indicates the current page is for adding a plugin to the 'Post image' field.

The plugin to add
A dropdown menu is set to 'Find replace'.

Description *
The description field contains 'Find replace'. To the right, the machine name 'find_replace' is shown with an [Edit] link. Below the field, a note states: 'A useful description of what this plugin is doing.'

CONFIGURE FIND REPLACE

Text to find
An empty text input field.

Text to replace
An empty text input field.

Case sensitive
If checked, "book" will match "book" but not "Book" or "BOOK".

Respect word boundaries
If checked, "book" will match "book" but not "bookcase".

Match whole word/phrase
If checked, then the whole word or phrase will be matched, e.g. "book" will match "book" but not "the book".
If this option is selected then "Respect word boundaries" above will be ignored.

Add

Save Feeds Migrate Type -

Now we have configuration with

- Fetcher (upload/download/..)
- Processor type (json/xml/csv)
- Named sources mapped to fields in a content type
- Processors

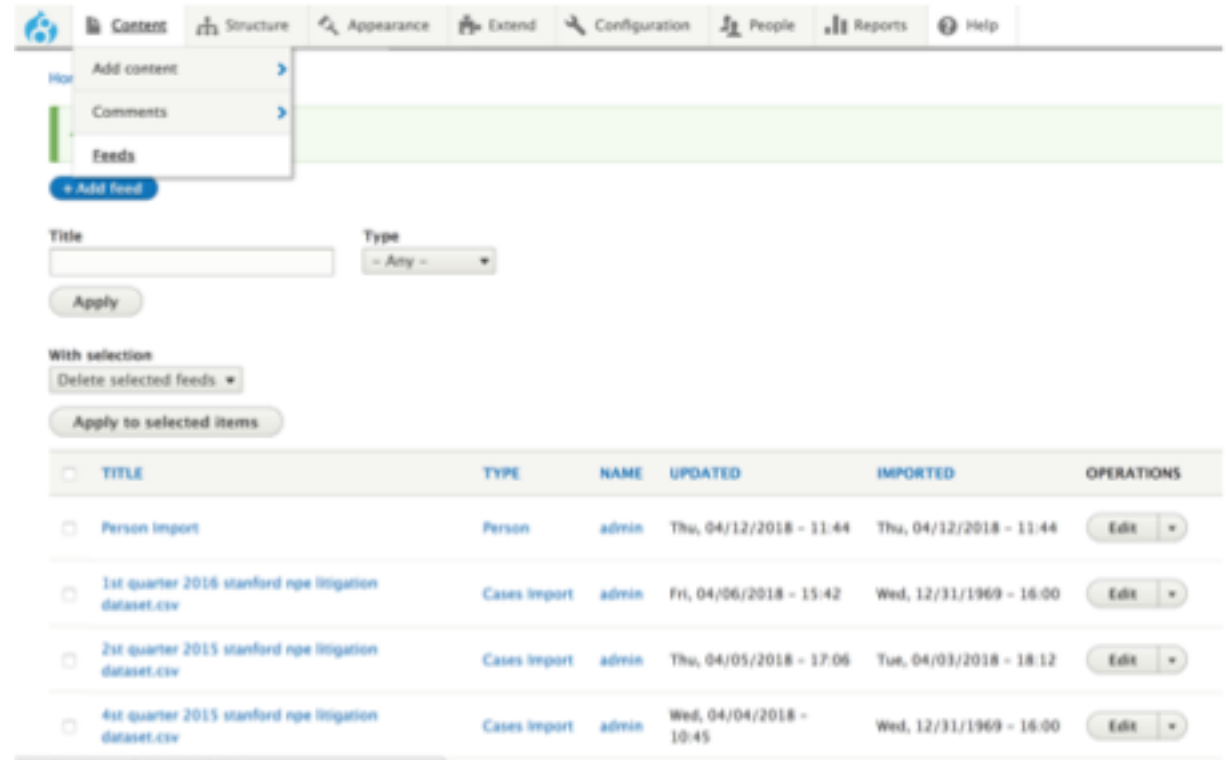
Import actual content

Add Feed – Drupal Content

Run actual import

Add Feed and upload actual import file

Set unlimited number of different imports/feeds with the same structure



The screenshot shows the Drupal administration interface. The top navigation bar includes 'Content', 'Structure', 'Appearance', 'Extend', 'Configuration', 'People', 'Reports', and 'Help'. A dropdown menu is open under 'Content', showing 'Add content', 'Comments', and 'Feeds'. Below the menu is a '+ Add feed' button. The 'Title' field is empty, and the 'Type' dropdown is set to '- Any -'. There is an 'Apply' button. Below that, the 'With selection' section shows 'Delete selected feeds' and 'Apply to selected items' buttons. The main content area displays a table of existing feeds.

<input type="checkbox"/>	TITLE	TYPE	NAME	UPDATED	IMPORTED	OPERATIONS
<input type="checkbox"/>	Person Import	Person	admin	Thu, 04/12/2018 - 11:44	Thu, 04/12/2018 - 11:44	Edit ▾
<input type="checkbox"/>	1st quarter 2016 stanford rpe litigation dataset.csv	Cases Import	admin	Fri, 04/06/2018 - 15:42	Wed, 12/31/1969 - 16:00	Edit ▾
<input type="checkbox"/>	2nd quarter 2015 stanford rpe litigation dataset.csv	Cases Import	admin	Thu, 04/05/2018 - 17:06	Tue, 04/03/2018 - 18:12	Edit ▾
<input type="checkbox"/>	4th quarter 2015 stanford rpe litigation dataset.csv	Cases Import	admin	Wed, 04/04/2018 - 10:45	Wed, 12/31/1969 - 16:00	Edit ▾

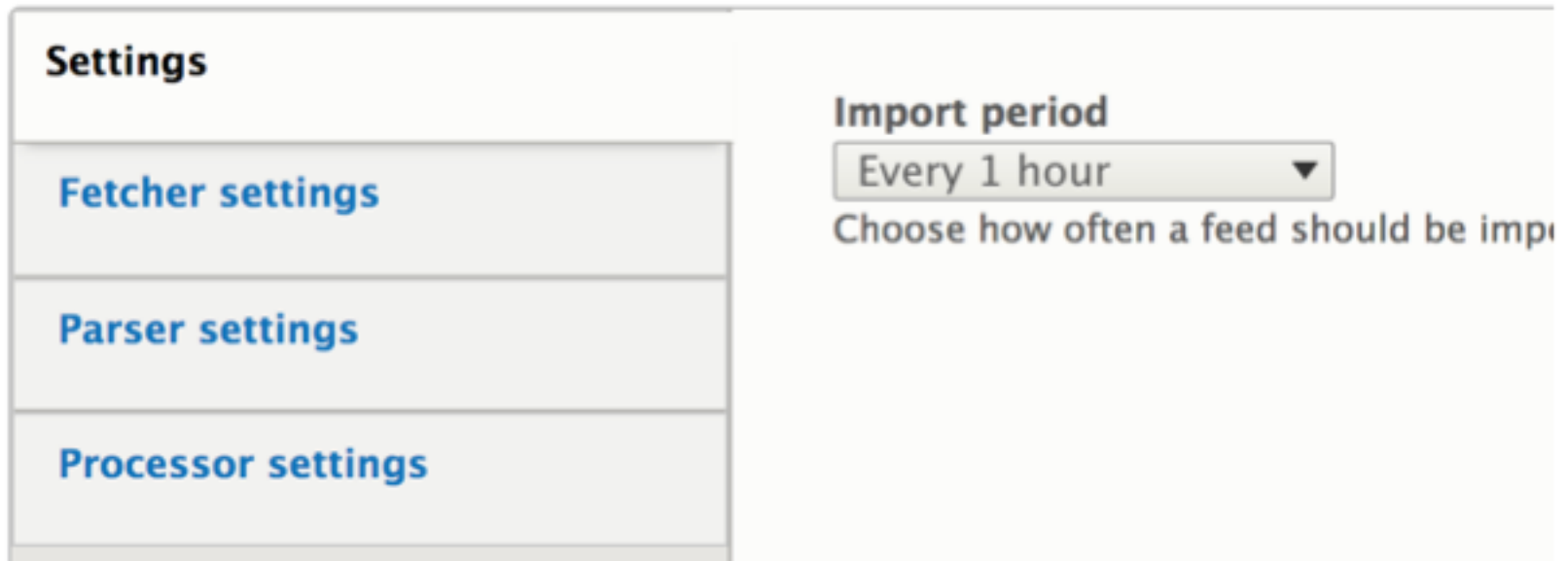
<http://dev-feeds8.pantheonsite.io/admin/content/feeds-migrate/importer/add>

<http://dev-feeds8.pantheonsite.io/admin/content/feeds-migrate>

Processor settings

Settings	
Fetcher settings	
Parser settings	
Processor settings	<p>Update existing contents</p> <p><input checked="" type="radio"/> Do not update existing contents</p> <p><input type="radio"/> Replace existing contents</p> <p><input type="radio"/> Update existing contents</p> <p>Existing contents will be determined using mappings that are unique.</p> <p>Expire contents</p> <p><input type="text" value="Never"/></p> <p>Select after how much time contents should be deleted.</p> <p><input type="checkbox"/> Owner: Feed author Use the feed author as the owner of the entities to be created.</p> <p>Owner</p> <p><input type="text" value="Anonymous (0)"/></p> <p>Select the owner of the entities to be created. Leave blank for <i>Anonymous</i>.</p> <p>▶ ADVANCED SETTINGS</p>

Set periodic import



The image shows a settings interface with a sidebar on the left and a main content area on the right. The sidebar is titled "Settings" and contains three menu items: "Fetcher settings", "Parser settings", and "Processor settings". The main content area is titled "Import period" and features a dropdown menu currently set to "Every 1 hour". Below the dropdown, there is a descriptive text: "Choose how often a feed should be imported".

Settings
Fetcher settings
Parser settings
Processor settings

Import period

Every 1 hour ▼

Choose how often a feed should be imported

Delete / rollback imports



1st quarter 2016 stanford npe litigation

View Edit Preview Delete Items Devel

- [Import](#)
- **Delete items**

Source: private://feeds/1st quarter 2016 stanford npe litigation dataset_2.csv

Items imported: 0

Last import: never

Next import: Wed, 12/31/1969 - 15:59

Unlock (feeds) = stop+reset (migrate)

Processors for Feeds_Migrate

https://www.drupal.org/project/feeds_migrate/issues/299199

Why it takes so long?

- Steep learning curve for developers
- Migrate developers ~~do not know~~ need to learn what is end user flow
- Feeds developers ~~do not know~~ need to learn API's and functions of migrate module

Our approach

- Weekly meetups on Slack channel
- Two workshops for Feed developers with @heddn, Migrate maintainer
- Documented UI for new module

Other presentations

James Dixon @ DUG <https://docs.google.com/presentation/d/1yz2mYlrrrel9q2XysBr6ZKFSuoXXhFVvSHG8YAF7acU/edit#slide=id.p>

Feeds Migrate UI

https://docs.google.com/presentation/d/1Hv8VPh6mD35U2dH2e1HhROKPNkea9GB3BSetX4St5x8/edit#slide=id.g3e2d97c3ca_1_28

Next step – Feeds+Migrate

- Integration of Feeds UI with Migrate engine
- https://www.drupal.org/project/feeds_migrate
 - [@MegaChriz](#) + [@heddn](#)
- UI and requirements
<https://docs.google.com/presentation/d/1Hv8VPh6mD35U2dH2e1HhROKPNkeaoGB3BSetX4St5x8/edit#slide=id.p>
- Session @BADcamp on Saturday, Oct 28

DevOps

Thanks to Pantheon for providing all tools

- Dev / test / multidev instances
- Backups
- Db / files export/import
- Drush



Contribute to Drupal ☺

<https://contribkanban.com/board/feeds>

The screenshot displays a Kanban board for the 'Feeds Migrate' project. At the top, there is a blue header with the title 'Feeds Migrate'. Below the header, there are two dropdown menus: 'Task' and 'Any priority'. The board is divided into four columns representing different stages of task completion:

- Postponed (0):** This column is currently empty.
- Active (4):** This column contains four task cards:
 - [META] Build a Feeds migrate UI (mockups) #2986955:** A green card with labels '8.X-1.X-DEV', 'NORMAL', 'USER INTERFACE', 'TASK', and 'FEEDS MIGRATE'.
 - Feeds migrate Mapping UI #3002361:** A green card with labels '8.X-1.X-DEV', 'NORMAL', 'USER INTERFACE', 'TASK', and 'FEEDS MIGRATE'.
 - Feeds migrate Processors/Tamper UI #3002362:** A green card with labels '8.X-1.X-DEV', 'NORMAL', 'USER INTERFACE', 'TASK', and 'FEEDS MIGRATE'.
 - Suggest unique selector path based on uploaded file #3006045:** A green card with labels '8.X-1.X-DEV', 'NORMAL', 'CODE', 'TASK', and 'FEEDS MIGRATE'.
- Needs Work (2):** This column contains two orange cards:
 - Use external form classes for configuring process plugins #2991199:** An orange card with labels '8.X-1.X-DEV', 'NORMAL', 'CODE', 'TASK', and 'FEEDS MIGRATE'.
 - Feeds migrate UI Basic settings #3002360:** An orange card with labels '8.X-1.X-DEV', 'NORMAL', 'USER INTERFACE', 'TASK', and 'FEEDS MIGRATE'.
- Needs Review (0):** This column is currently empty.

Feedback? Questions?

#feeds on Drupal slack.

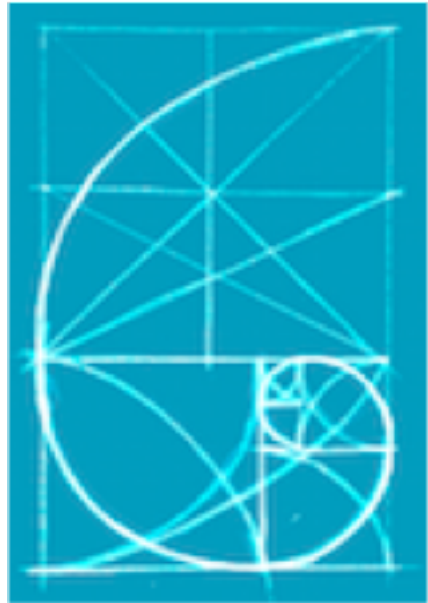
Office hours Thursday @ 11am Pacific

Get involved! Check out the roadmap on project page:

https://www.drupal.org/project/feeds_migrate

James Dixon. james@dialedin.ca





Thank you!

Irina Zaks,
Fibonacci Web Studio
Stanford Open Source Lab
izaks@stanford.edu