Drupal Feeds

Bill Severe

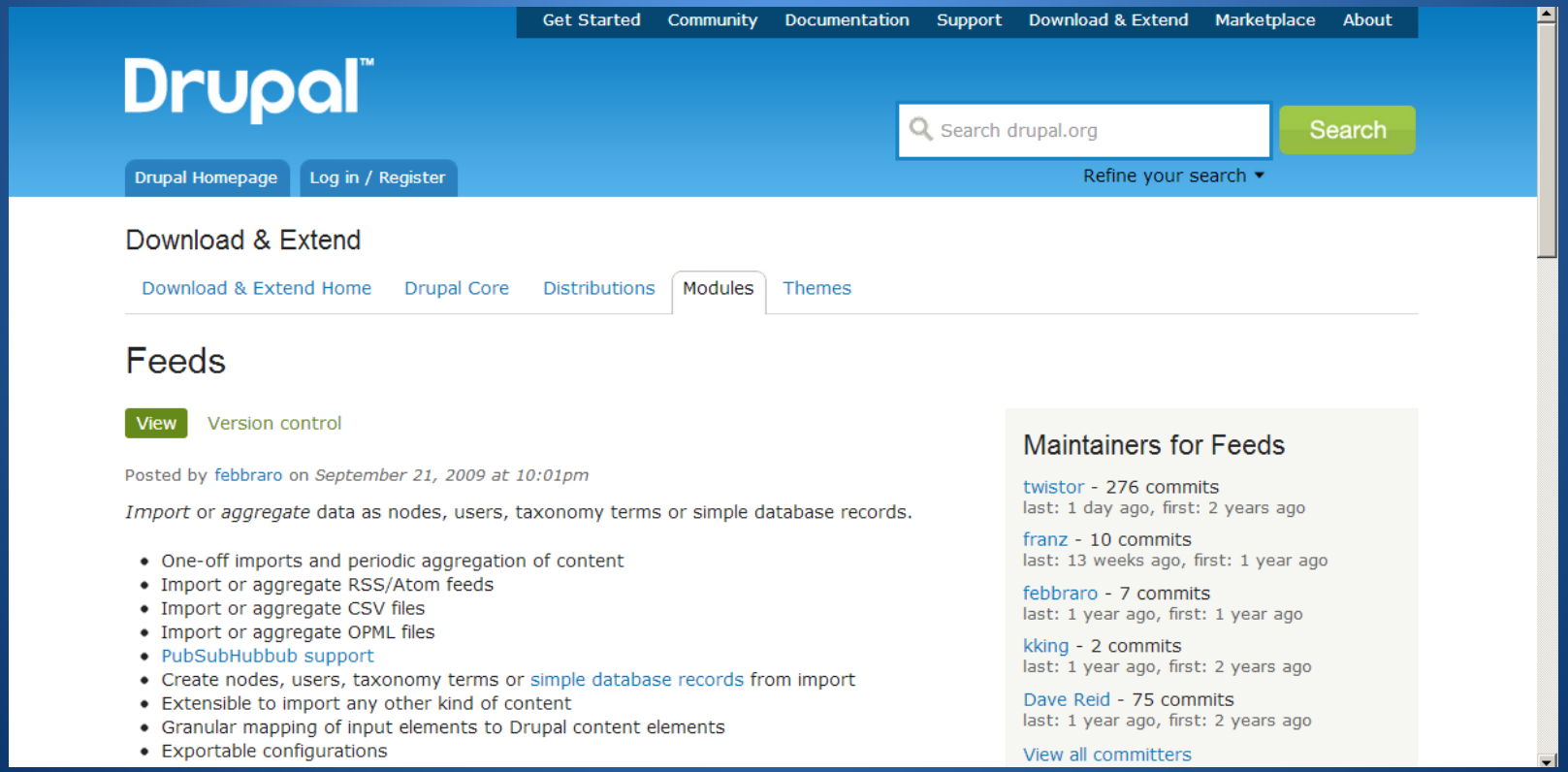
Special Thanks to Sean Wilson [QUOTIENT]

Topics to be discussed

- Using the feeds modules to import/export content to/from Drupal 7 sites, via an XML feed

Topic: Feeds Module

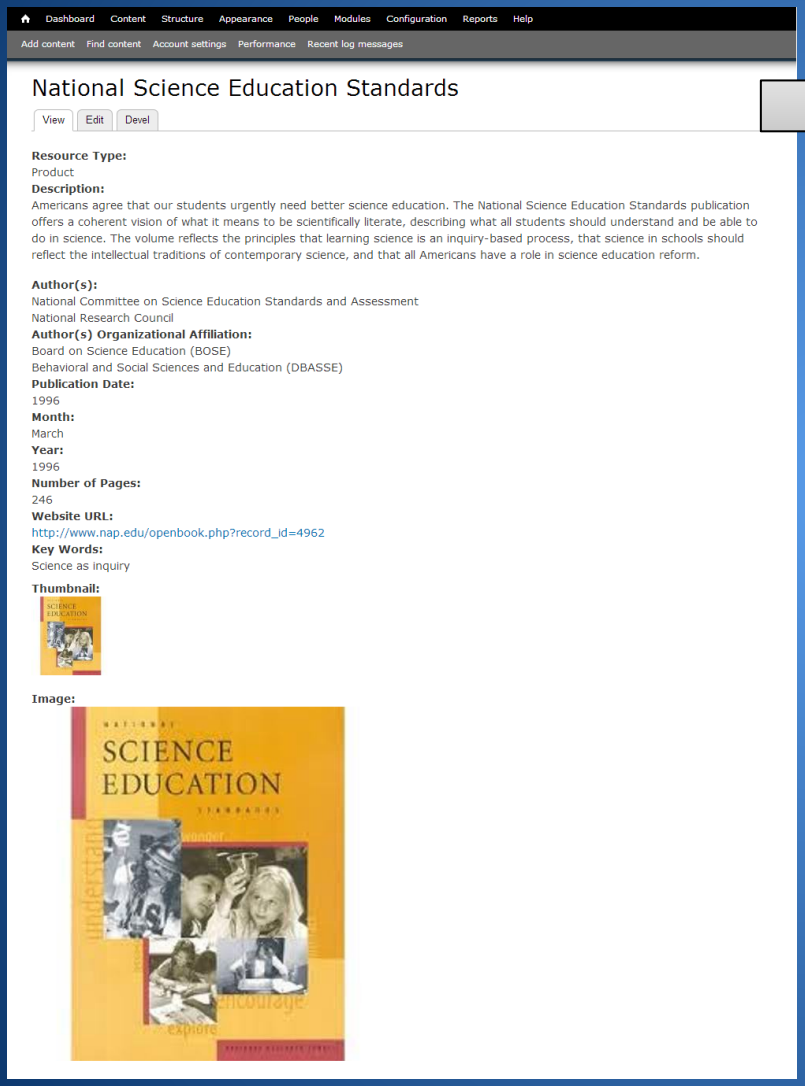
- Pull in contents of XML feed as nodes



What is Feeds useful for? Examples:

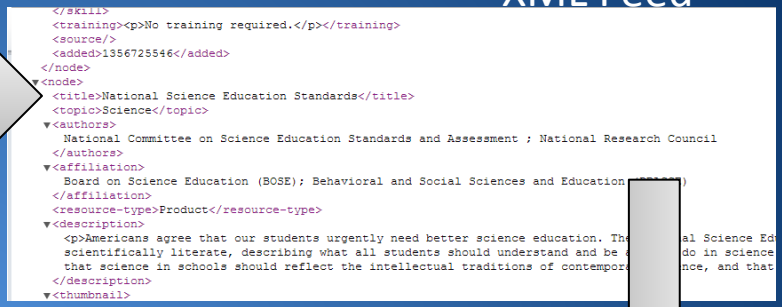
- One-time: migrating content or taxonomy items from one Drupal instance to another
- Creating nodes from a static CSV file's entries
- Providing scheduled content updates from one site to another, our scenario
  - Content is vetted on "exporter site"
  - Once vetted, content is put into an XML feed
  - The XML Feed is ingested by "importer site"
    - new node added/updated on importer site

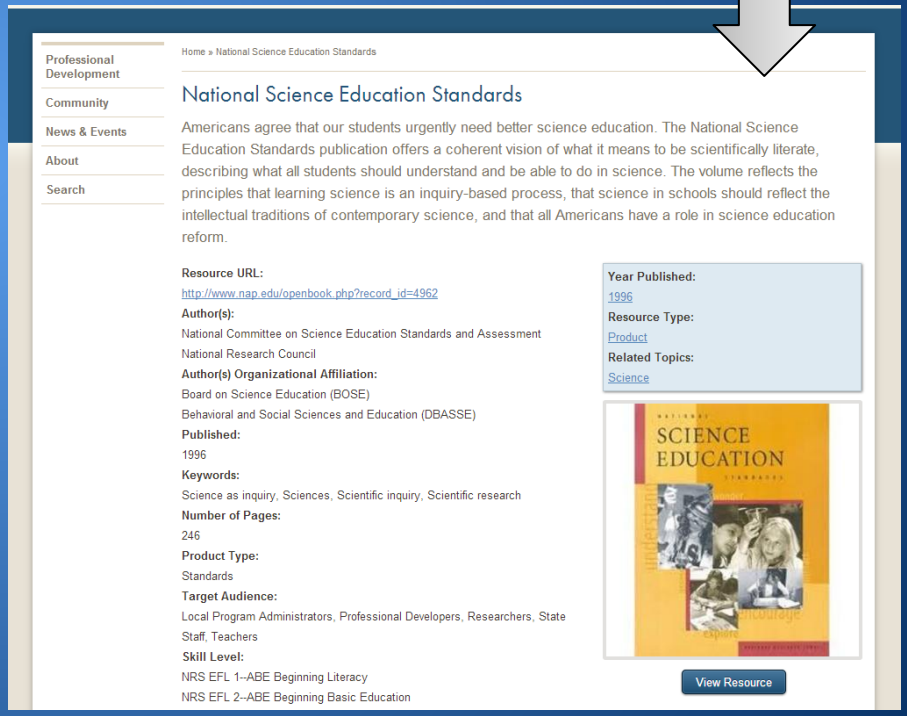
Exporter-Importer model



Exporter Site

XML Feed





Importer Site

Modules Needed for Exporter-Importer model

- On Exporter Site
  - **Views**
  - **Views Data Export**
- On Importer Site
  - **Feeds**
  - **Feeds XPath Parser**
  - **Job Scheduler**
  - **Feeds Tamper!**

Steps to implementation
Exporter Site

Exporter Site

- Set up Content Type on Exporter Site ex. – Article
- Configure Views Data Export view display of Article nodes (Remember - filter for Published only!)
- Test viewing of the feed URL as an anonymous user

Note: Every time you make a change to the data export view config, remember to clear your cache!!!

Even if you have caching disabled on the view, I have still experienced caching problems.

Steps to implementation Importer Site

Importer Site

- Set up Content Type on Importer Site - ex. Article
- Create feed importer
- Add settings to fetch XML feed via HTTP and create Article Nodes based on it
  - Important Options:
    - Periodic Import
    - Update Existing Nodes
    - Expire Nodes
  - Ignore XPath XML parser settings for now

Steps to implementation Importer Site – Cont.

Node Processor Mapping

- Fill in XPath parser settings
- Set XPath targets (AKA Node fields) for retrieved data
- Set a field as unique in order that nodes can be updated later – otherwise they get duplicated! ☹
  - Good choice for unique ID – Node ID on exporter site

Steps to implementation Importer Site – Cont.

At http(s)://ROOT_URL/import:

- Input feed URL

XPath Parser settings

- Set context to: /nodes/node
- Enumerate XML feed tags (source) to corresponding Node fields (target)
- Click Import and go!

Trouble spots

- Importing data for List fields
  - List options on Exporter and Importer fields should be the same
  - If they change on Exporter site, must be changed manually on Importer site too
  - Can get around this by using taxonomy fields instead of list fields
    - Then exporting and importing via a separate, taxonomy XML feed

Trouble spots cont.

- Multiple List items?
  - Use Feeds Tamper!
  - Explode on item separator
  - Trim whitespace on both sides
  - Ex. value “red, round” gets split into two values, “red” and “round”
- Need to find/replace section of content before it is saved to a node’s field?
  - Use feeds tamper!

Disadvantages of Exporter-Importer model?

- The Feed can get quite large
  - On both exporter and importer sites:
    - update /etc/my.cnf - set max_allowed_packet to be larger than your feed size, and watch over time
  - Import delays due to feed size: Not such a big issue if both sites are on same internal network
    - Currently 500 nodes in our feed
    - Total feed size: Currently ~1.6 MB
    - Total Time to import: 10 – 20 seconds if very few node changes?
- Tip to reduce feed size: Just set filter on exporter site to only display items modified today in feed

Security Concerns

- DDoS attack on exporter site
- Man in the middle insertion attack on feed

Concerns can be mitigated by Apache directives – restrict access to feed to certain IPs only, including IP of importer site

DIY!

- http://www.ostraining.com/blog/drupal/feeds
- http://www.atendesigngroup.com/blog/drupal-imports-using-feeds-module
- DrupalCon Portland 2013: GET YOUR DATA INTO DRUPAL WITH FEEDS http://www.youtube.com/watch?v=isF04Aswf7w

DIY!

- https://github.com/bsevere/feedsdemo
  - This presentation
  - Views_data_exporter configuration
  - Article content type export
  - Article feed importer configuration

Questions?

bsevere4@gmail.com