Fedora and Django

at Memorial Sloan-Kettering Cancer Center

Thursday, May 21, 2009
First Use of Fedora

Surgery

- Fedora 2.2.3 + Muradora
- 2DC, MELCOE, MediaShelf

Thursday, May 21, 2009
Second Use of Fedora

Public Affairs

- Fedora 3.1
- Python
- Django
Use cases

- Allow MSKCC employees to find and re-use graphics assets
- Allow Public Affairs employees with better knowledge to find and re-use assets
Use cases

• Allow MSKCC employees to find and re-use graphics assets

• Allow Public Affairs employees with better knowledge to find and re-use assets
Use cases

- Allow MSKCC employees to find and re-use graphics assets
- Allow Public Affairs employees with better knowledge to find and re-use assets
Constraints

- Roughly 4 month deadline
- Must look at least vaguely like Surgery's Muradora instance
- Internal resistance to a new language & platform
(by me)
using Python + Django
Why internal?
Experience with Muradora

- slow turn-around
- not enough expertise in-house
- unproductive language
Experience with Muradora

- slow turn-around
- not enough expertise in-house
- unproductive language
Experience with Muradora

- slow turn-around
- not enough expertise in-house
- unproductive language
Experience with Muradora

- slow turn-around
- not enough expertise in-house
- unproductive language
How do you steer the Titanic left?
How do you steer the Titanic left?

Very, very slowly.
Agile language, Python

http://www.python.org
Django
"The framework for perfectionists with deadlines"
Pros

- Experience writing Synapse (http://synapse.mskcc.org)
- Availability of plug-in apps for various needs
- Rapid development
- Control
- Ease of learning for new programmers
- 80% done for you, just concentrate on domain
Pros

- Experience writing Synapse (http://synapse.mskcc.org)
Pros

- Availability of plug-in apps for various needs
  - http://www.djangosnippets.org/
  - http://djangoplugables.com/
  - http://github.com/search?q=django&type=Repositories&x=0&y=0
  - http://bitbucket.org/repo/all/?name=django
  - https://launchpad.net/+search?field.text=django
Pros

- Rapid development
- Control
- Ease of learning for new programmers
- 80% done for you, just concentrate on domain
Pros

• Rapid development
• Control
• Ease of learning for new programmers
• 80% done for you, just concentrate on domain
Pros

- Rapid development
- Control
- Ease of learning for new programmers
- 80% done for you, just concentrate on domain
Pros

- Rapid development
- Control
- Ease of learning for new programmers
- 80% done for you, just concentrate on domain
Cons

• ORM -- interfacing the Django relational perspective with Fedora

• Python performance v. Java
Cons

• ORM -- interfacing the Django relational perspective with Fedora

• Python performance v. Java
Cons

- ORM -- interfacing the Django relational perspective with Fedora
- Python performance v. Java
Custom ingest script

- Uses Python Imaging Library, Exempi
- Extracts EXIF/XMP metadata from image
- Watermark, screensize, thumbnail
- Easily extensible for new kinds of content, new types of datastream
Ingest How-To
ingest --help

Usage: python ingest [OPTIONS] DIRECTORY [DIRECTORY DIRECTORY...]

Ingest the contents of the supplied directory or directories into a FEDORA repository

Options:
--version show program's version number and exit
-h, --help show this help message and exit
-H HOST_URL, --host=HOST_URL URL of Fedora repository [default: http://localhost:8080/fedora]
-u USER, --user=USER User account for ingest [default: fedoraAdmin]
-p PASSWORD, --password=PASSWORD Password for ingest
-c COLLECTION, --collection=COLLECTION Sets the top-level collection that will contain all ingested files [default: phfed31:1]
-w, --watermark Overlay screen-size images with contents of WATERMARK_FILE
--watermarkfile=WATERMARK_FILE Set path to WATERMARK_FILE [default: /Users/herndonp/Projects/Work/fedora31/watermark3.png]
--no-recurse Turns off recursive consumption of directories, sets RECURSE=False [default: RECURSE=True]
Django for the web UI

Security

LDAP authN/authZ against Active Directory
Group permissions
ACLs per object using django-granular-permissions
Not using XACML
Easy permissions management
Django for the web UI

Browse View

Thursday, May 21, 2009

Lightbox effect using Ajax
Metadata details
Keywords linked to search
Django for the web UI

Search

Solr integrated via django-solr-search, aka Solango
Django for the web UI

Tagging with controlled vocabularies
Django for the web UI

Bookmarks and bulk download
Django for the web UI

Django custom management command
To-Do

• New Object Submission

• Object Editing

• Better LDAP <-> Group mapping UI

• Finish tagging, batch tagging

• Handle additional file formats
  • PDF, PPT, Word, Excel

• Improve ingest efficiency

• Test with Fedora 3.2

Thursday, May 21, 2009
Available Now, GPLv3

- http://bitbucket.org/tpherndon/django-fedora/
- http://bitbucket.org/tpherndon/django-vocabulary/
- http://bitbucket.org/tpherndon/pyfedora/
Q & A

• Peter Herndon

• Sr. Unix Systems Administrator

• Memorial Sloan-Kettering Cancer Center

• herndonp@mskcc.org

• tpherndon@gmail.com

• @tpherndon