

Packaging and Pushing

Experimenting with workflows for APTrust

Nathan Tallman
Digital Content Strategist
University of Cincinnati Libraries

APTrust Fall Meeting
6 October 2014

Sources of Digital Content

- Digital Resource Commons
 - Dspace 1.8.2



- LUNA Image and Media Repository
 - LUNA 6.3.6.2



- Scholar@UC
 - Hydra

scholar@uc

- SAN Filestore

DSpace: Simple Archive Format

```
item_000/  
  dublin_core.xml  
  contents  
  file_1.doc  
  file_2.pdf
```

DSpace: AIP METS Packages

item_000/

mets.xml

bitstream.doc

license

DSpace: Replication Task Suite Bags

item_000/

bagit.txt

manifest-md5

tagmanifest-md5

data/

license/

metadata.xml

object.properties

ORIGINAL/



http://fc03.deviantart.net/fs50/f/2009/326/b/8/Hydra_monster_by_velinov.jpg

Sample Bags

- Call to Dspace forum for sample RTS produced bags
 - Markus Wust, NCSU Libraries
- Make copies, change properties
- XSL transformations to conform to APTrust bagit spec

Sample Bags

- Push bags to APTrust receiving bucket using AWS CLI
- Wait, wait, wait...
 - fail?



<http://editorial.designtaxi.com/news-dog030613/1.jpg>

Replication Task Suite

- Install RTS on test Dspace instance
- Bash script
 - unzip
 - apply XSLT
 - tar
 - push to APTrust
- Wait, wait, wait...
 - partial success!



<http://rwtrainingkennel.com/wp-content/uploads/Happy-Dog-Pictures-3.jpg>

But wait...

- Administrative sets vs. flat bags
- Platform provenance
- File system metadata
 - dates
- Thumbnails



<http://www.prwatch.org/files/images/dogheadtilted.jpg>

Next Steps

- Use AWS SW3 API?
- Test restoration of content and collections
- RTS or not?
- Refactor bash script and put on github, more automating
- Send more content! Other platforms!

Questions

- Are others preserving file system metadata?
- How are others tracking administrative context of collections?
- Has anyone tested restoration?
- Is anyone creating bags containing additional bags? Or do you repackaging existing bags?