Building flexible workflows with Fedora at the University of York

Julie Allinson and Frank Feng

The 5th International Conference on Open Repositories

Madrid, Spain, 8th July 2010

THE UNIVERSITY *of York*

What are we doing at York?

- We're making a …
  - Multi-media
  - Multi-disciplinary
  - Multi-access control
- Digital Library
- out of Fedora Commons
- with JISC funding
- and other components

THE UNIVERSITY *of York*

Components

- Fedora Commons

- Muradora, moving towards FeSL

- Workflow components

  – orbeon xforms, file processing factory, messaging service

- SWORD deposit proxy

- Deposit interface(s)

THE UNIVERSITY *of York*

Three requirements

- A workflow for administrators
  - CRUD, plus publish and move
  - rich metadata creation
  - data dictionary management
  - policy application
- A simple user deposit workflow
  - deposit to multiple repositories
- Something for the bulk and batch

Three layers of support

- Supported, e.g. TIFF, JPEG
- Partly Supported, e.g. 'known' image types
- Unsupported, anything else

  - Things can move layer, i.e. if we identify a format that needs partial or full support [with bulk processing of existing resources]
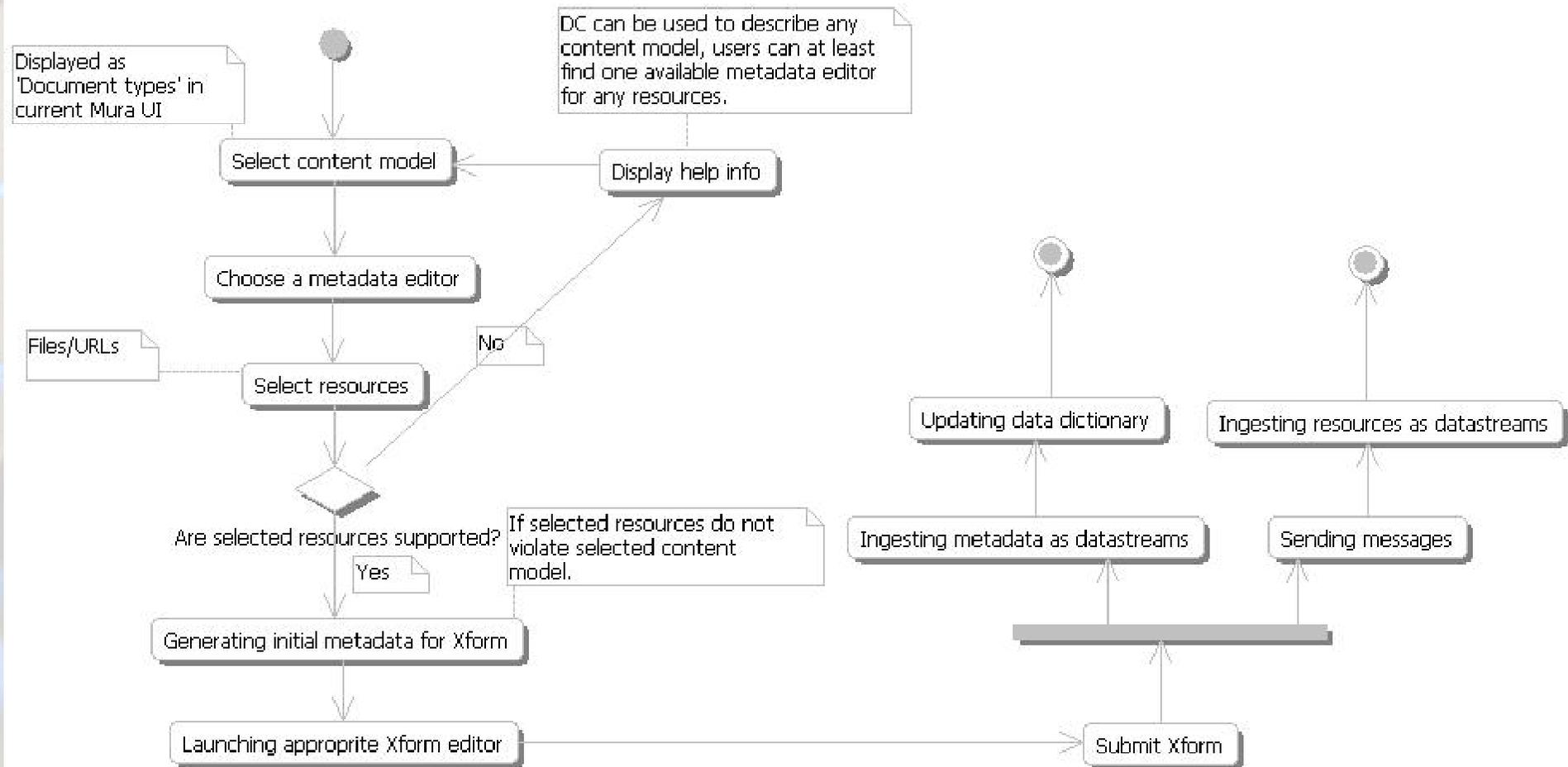
Factory patterns
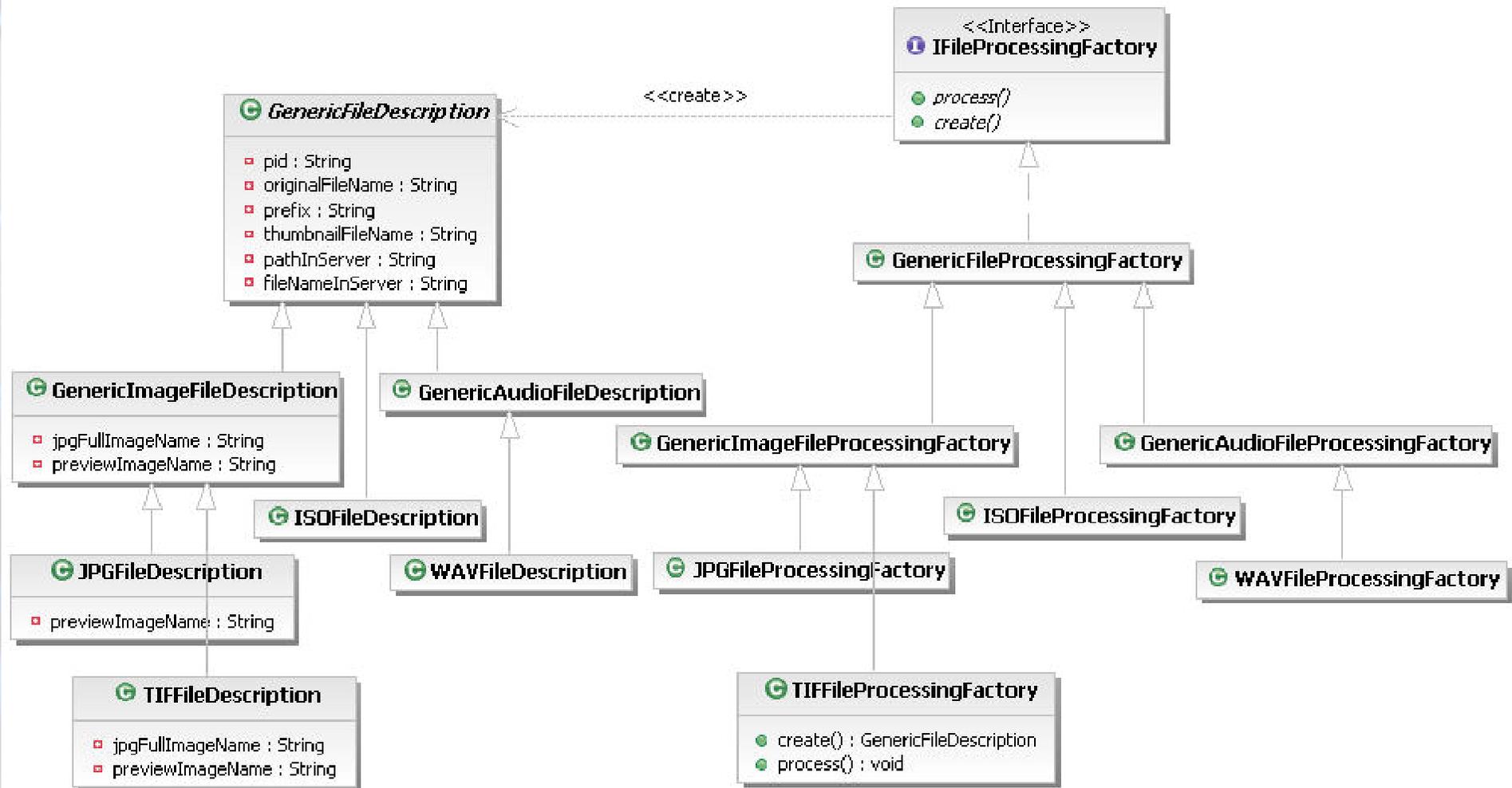
Three step workflow

- Content model
  - Image, Theses, Audio, Unsupported
- Resources(s)
  - from local drive, mapped drive, by URI or external reference
- Metadata
  - VRA (images), MODS (music), UKETD (theses), DC (anything else)
  - all mapped to simple DC for base interoperability

Submission workflow : overview

Submission workflow – step 1

THE UNIVERSITY of York

Submission workflow – step 2

Submission workflow – step 3

Submission workflow – step 3

THE UNIVERSITY *of York*

Submission workflow – step 3

Location/Citation

Name type   ○ corporate   ○ personal   ● other   ?

Reference no./URI

⊕ Add source information

✗ ISBN or ISSN   ?   Type: ● ISBN ○ ISSN ?

⊕ Add other location information

Rights

Type of rights   copyrighted ▼ ?

License   Please choose ▼ ?

| Please choose |
| CLA Photocopying and scanning HE licence |
| Creative Commons (by-nc-sa) |
| Creative Commons (by-nc-nd) |
| Permitted Acts |

Rights holder

⊕ Add rights statement

⊕ Add copyright clearance method

⊕ Add copyright clearance duration

Submission workflow – step 3

Location/Citation

Name type

Reference no./URI

Add source infor

ISBN or ISSN

Add other locati

Type of rights

License

Rights holder

Add rights state

Add copyright cl

Add copyright cl

Rights

Type of rights   copyrighted

License   CLA Photocopying and scanning HE licence

Rights holder

Add rights statement

Rights statement   Scanned under the terms of a CLA Photocopying and Scanning HE Licence. Access is restricted to Course Users associated with Courses of Study for which this resource is designated. Please refer to the Copyright Notice and University of York Digital Library Rights Policy for further information.

Add copyright clearance method

Clearance method CLA

Add copyright clearance duration

Switch to work   Save

Asynchronized back-end program for file processing

Step 2: Resource selection

Select resource from mapped drive

"Z:\Copy of york_88 [Choose...]

▼ Show other resource select options

Selected resource list

```
[Default]mapped://Copy of york_8812.tif
```

[Remove selected resource(s)] [Make selected resource as default]

☐ Add similar

[Back] [Next]

```
Launching FileMapping message processor...
Loading configuration file...Done.
Connecting to Message server...Done.
Retrieving messages...

        For pid: york:7
        Moving file: Copy of york_8812.tif ... Done.
        Adding data stream: ARCHIVAL_IMAGE ...Done.
        Generating jpg image from tif ... Done.
        Generating preview image ... Done.
        Generating thumbnail image ... Done.
        Adding data stream: FULL_IMAGE_JPEG ...Done.
        Adding data stream: PREVIEW_IMAGE ...Done.
        Adding data stream: Thumbnail ...Done.
```

Access Control

Data dictionaries

- Data dictionary management tool
  - Managing existing DD items

  - Approve / deny
  entered DD items

Low barrier self-deposit

- Tool for **anyone** at York (and WRRO colleageus) to deposit (just about) **anything**

- EPrints to WRRO (backed-up in YODL) via SWORD

- Other stuff to YODL

- Web-based (at present) and uses Adobe Flex

THE UNIVERSITY *of York*

SWORD deposit proxy design
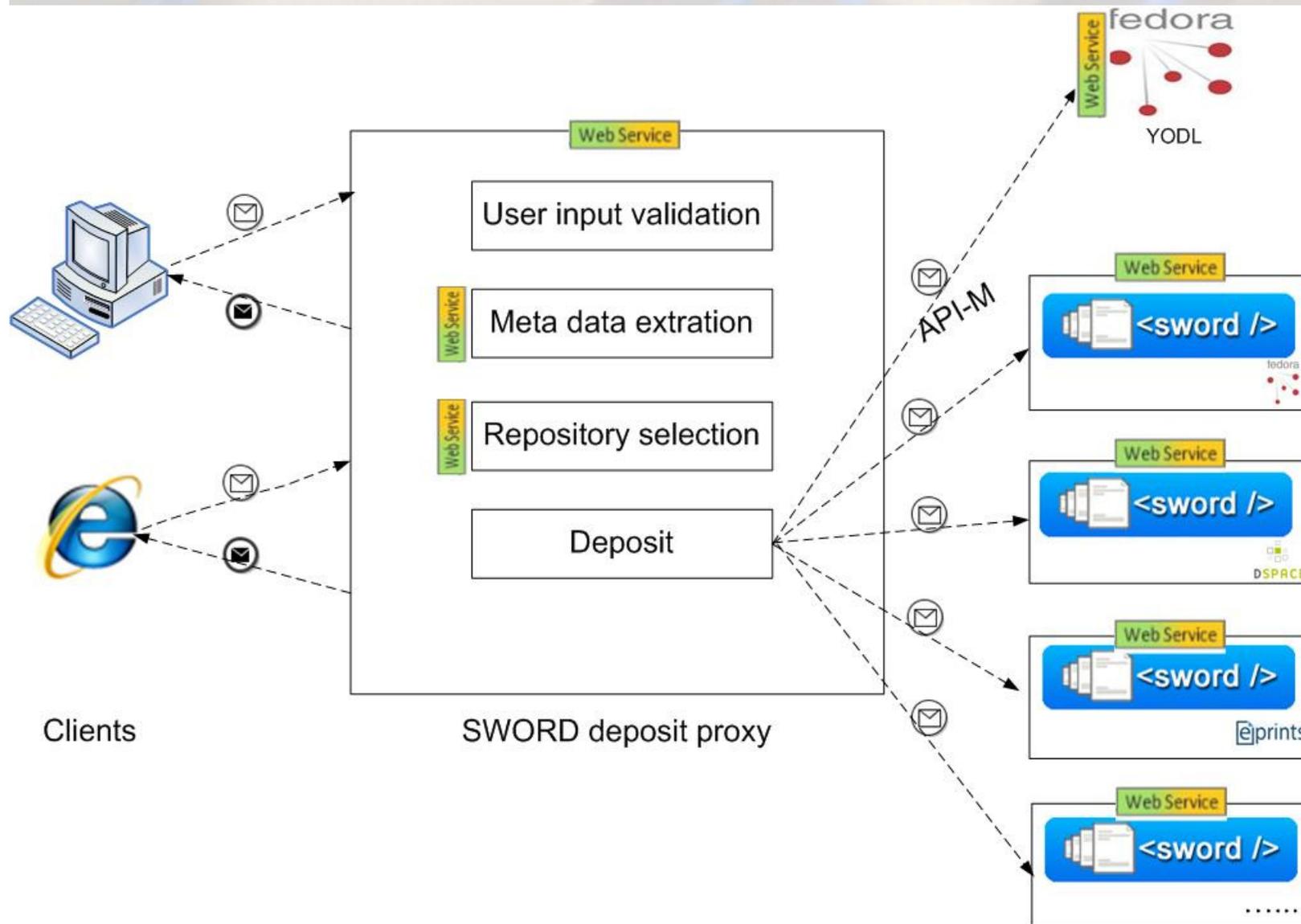
Prototype deposit tool

Resources | **Metadata**

Please select file(s) to deposit, the top file is the main file and will be used to determine suitable repository(repositories)

File	File Type	File Size

Browse For Files | Upload | Remove | Clear All

Please choose a repository (repositories)

Next

Bulk and batch

- Currently by a series of one-off scripts and programs
  - Batch ingest, loading batches of objects/datastreams
  - Batch update, updating existing batches
- These need turning into a suite of admin tools that either stand-alone or integrate with current admin or self-submission tools, e.g. a bulk ingest desktop tool

Summary

- Accept anything
- Deal with multiples: resources, access requirements, repositories, collections
- Three types of 'workflow'
- Re-use of components, eg. file processing factories will be used to process self-deposits

Questions?

ja546@york.ac.uk

THE UNIVERSITY *of York*