

# Building flexible workflows with Fedora at the University of York

Julie Allinson and Frank Feng

The 5<sup>th</sup> International Conference on Open  
Repositories

Madrid, Spain, 8<sup>th</sup> 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

# Components

- Fedora Commons
- Muradora, moving towards FeSL
- Workflow components
  - orbeon xforms, file processing factory, messaging service
- SWORD deposit proxy
- Deposit interface(s)

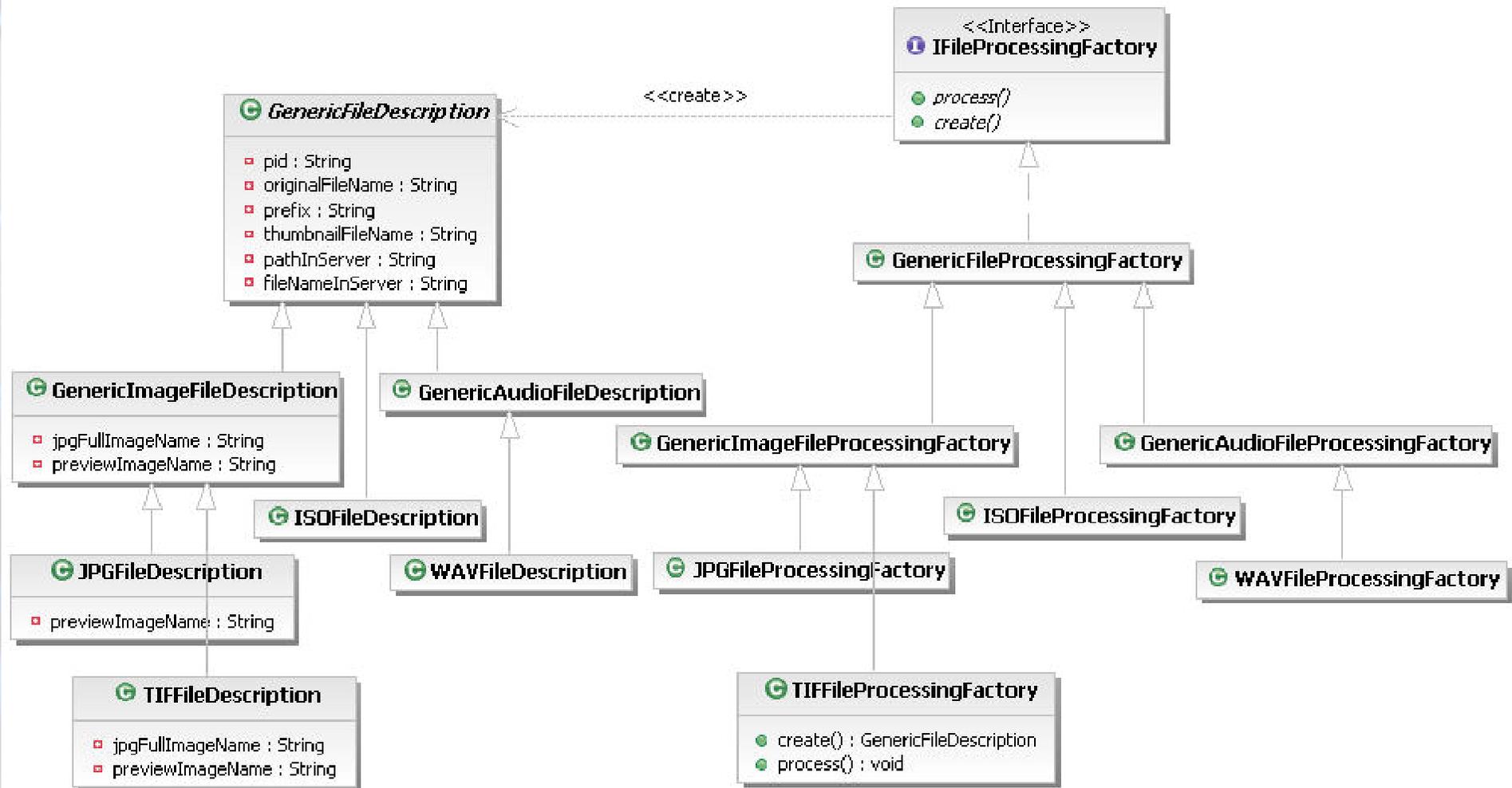
# Three requirements

- 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]

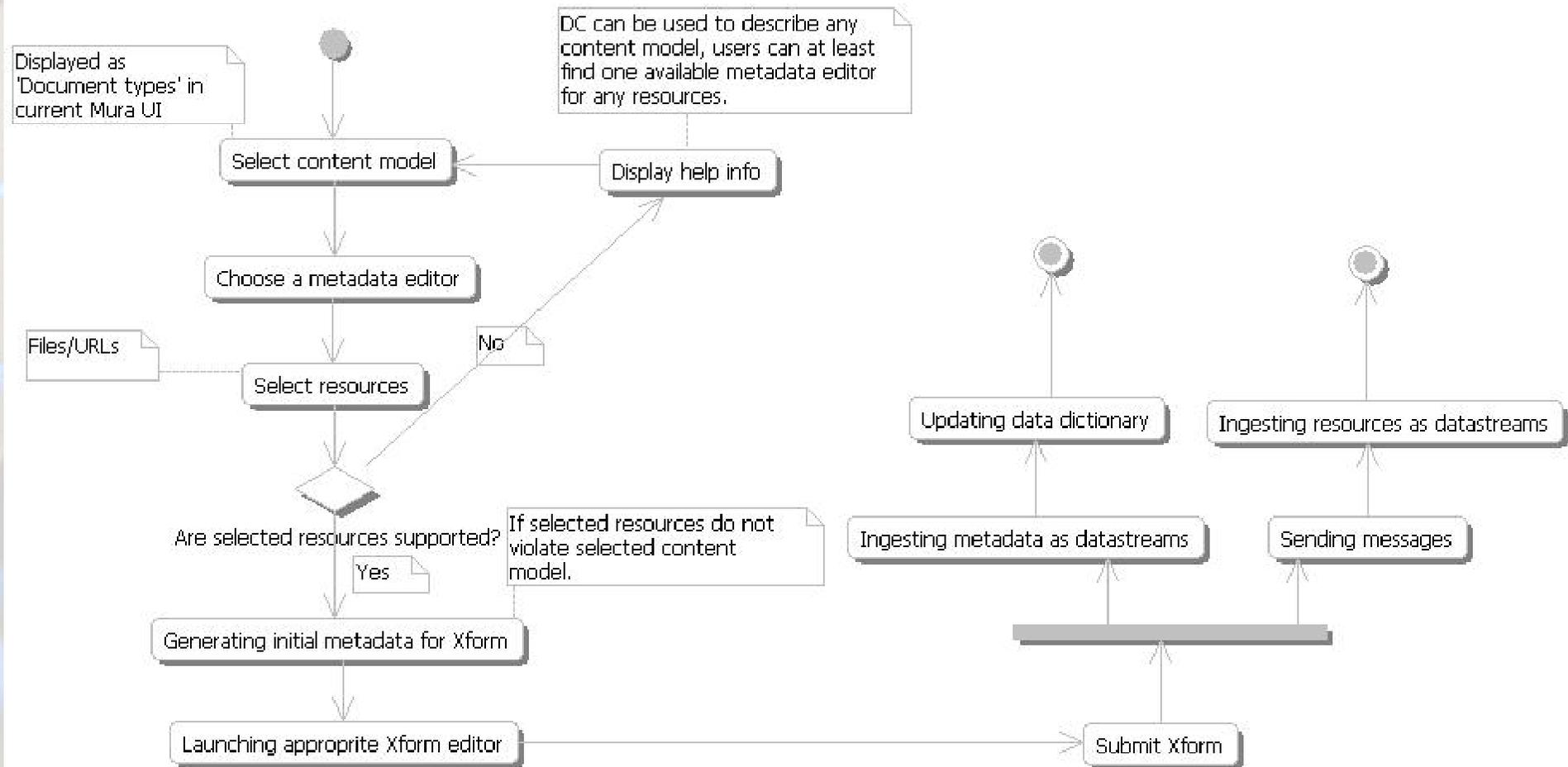
# Factory patterns



# 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

## Submission Wizard

①

### Step 1: Select collection and document type

Collection : [Select collection](#)

Document Types: Image

Metadata type : VRA metadata editor

Next

Home

History of Art [Select](#)

Library and Archives [Select](#)

Painton Cowen Images [Select](#)

Page 1 1

②

Collection : [Select collection](#)

Document Types: Image

Metadata type :  
Image  
Audio  
DiskImage  
Unsupported

③

Collection : [Select collection](#)

Document Types: Image

Metadata type :  
VRA metadata editor  
VRA metadata editor  
Dublin Core metadata editor

# Submission workflow - step 2

## Submission Wizard

### Step 2: Resource selection

#### Select resource from mapped drive

Z:\f1v\_opening\_initial.jpg

Hide other resource select options

#### Resource upload

#### Upload resource from an URL

#### External referenced / Redirected content resources

#### Selected resource list

[Default]uploaded://upload/t9f07e8d3-6313-494d-ae3-10cdc9f...  
urlupload://http://www.york.ac.uk/np/tp-images/logo\_top\_100.gif  
Redirected://http://eprints.whiterose.ac.uk/images/wro.jpg  
mapped://f1v\_opening\_initial.jpg

Add similar

York

# Submission workflow - step 3

Work

Image(s)

### Type of artwork

+ Add a worktype

x Type



### Title of artwork

+ Add a title

x Title



### Name(s) of artist(s) and other persons/groups responsible for the artwork

+ Add a name

Type  corporate  personal  other ?

Name



x Attribution



Role



### Date(s) of the artwork

k

# Submission workflow - step 3

The screenshot displays a web interface for submitting artwork. At the top, there are two tabs: "Work" and "Image(s)". Below the tabs is a large cyan-colored area with a yellow header "Type of artwork". On the left side of this area, there is a vertical list of artwork types, each preceded by a small icon (either a plus sign in a green circle or a red X). The "Type" dropdown menu is open, showing a list of options: "p", "painting", "pamphlet", "performance", "performance art", "periodical", "photocollage", "photograph", "photomechanical print", "photomontage", "photomural", "plate", "postcard", "poster", "print", "program", "pulpit", "pyx", and "pyxis". The "painting" option is currently selected and highlighted in blue. To the right of the dropdown menu, there is a red rectangular button with a white question mark icon. Below the dropdown menu, there is a white input field.

# 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**  ?

**License**  ?

**Rights holder**  ?

**Add rights statement**

**Add copyright clearance method**

**Add copyright clearance duration**

**License dropdown menu:**

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

# Submission workflow - step 3

Location/Citation

**Rights**

Type of rights  ?

License  ?

Rights holder  ?

**Add rights statement**

Rights statement  ?

**Add copyright clearance method**

Clearance method  ?

**Add copyright clearance duration**

Name type

Reference no./URI

**Add source information**

ISBN or ISSN

**Add other locations**

Type of rights

License

Rights holder

**Add rights statement**

**Add copyright clearance method**

**Add copyright clearance duration**

Switch to work  ?

# Asynchronized back-end program for file processing

## Step 2: Resource selection

### Select resource from mapped drive

"Z:\Copy of york\_88"

Show other resource select options

### Selected resource list

[Default]mapped://Copy of york\_8812.tif

Add similar

```
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

Create policy for the selected resource

## Select users or groups

administrator (role)  
public (role)  
york (role)  
staff (role)  
student (role)  
teacher (role)  
support (role)  
**hoaStaff (role)**  
hoaPG (role)  
hoaUG (role)

Add users or groups

## Basic permissions

## Advanced permissions

Permission	Allow	Deny	N/A	Error
create	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
read	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
delete	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
update	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
admin	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
publish	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

## Select users or groups

administrator (role)  
public (role)  
york (role)  
staff (role)  
student (role)  
teacher (role)  
support (role)  
hoaStaff (role)  
hoaPG (role)  
hoaUG (role)

Add users or groups

## Basic permissions

## Advanced permissions

### Rule effect

Permit  Deny

### Criteria combination

AND  OR

### Attribute

hasDatastreamLabel

equals

### Value

ARCHIVAL\_IMAGE

Add criterion



# Data dictionaries

- Data dictionary management tool
  - Managing existing DD items
  - Approve / deny entered DD items

**Data dictionary management**

Agent name | Agent role | Location name | Location place | Material | Subject topic | Technique | Work type  
| Photographers | Image rights holder |

709 items found, displaying 1 to 20.  
[First/Prev] 1, 2, 3, 4, 5, 6, 7, 8 [Next/Last]

ID	File	Element	Text	New value	Delete
1	work.agent.name.xml	name	Abacco, Antonio dall'	<input type="text"/>	<input type="checkbox"/>
2	work.agent.name.xml	name	Acconci, Vito	<input type="text"/>	<input type="checkbox"/>
3	work.agent.name.xml	name	Ader, Bas Jan	<input type="text"/>	<input type="checkbox"/>
4	work.agent.name.xml	name	Albers, Josef	<input type="text"/>	<input type="checkbox"/>
5	work.agent.name.xml	name	Alberti, Cherubino	<input type="text"/>	<input type="checkbox"/>
6	work.agent.name.xml	name	Alberti, Giovanni	<input type="text"/>	<input type="checkbox"/>
-	name.xml	name	Alberti, Leon Battista	<input type="text"/>	<input type="checkbox"/>
-	name.xml	name	Alexander VII	<input type="text"/>	<input type="checkbox"/>
-	name.xml	name	Algardi, Allessandro	<input type="text"/>	<input type="checkbox"/>
-	name.xml	name	Allori, Cristofano	<input type="text"/>	<input type="checkbox"/>

5 items found, displaying all items.  
1

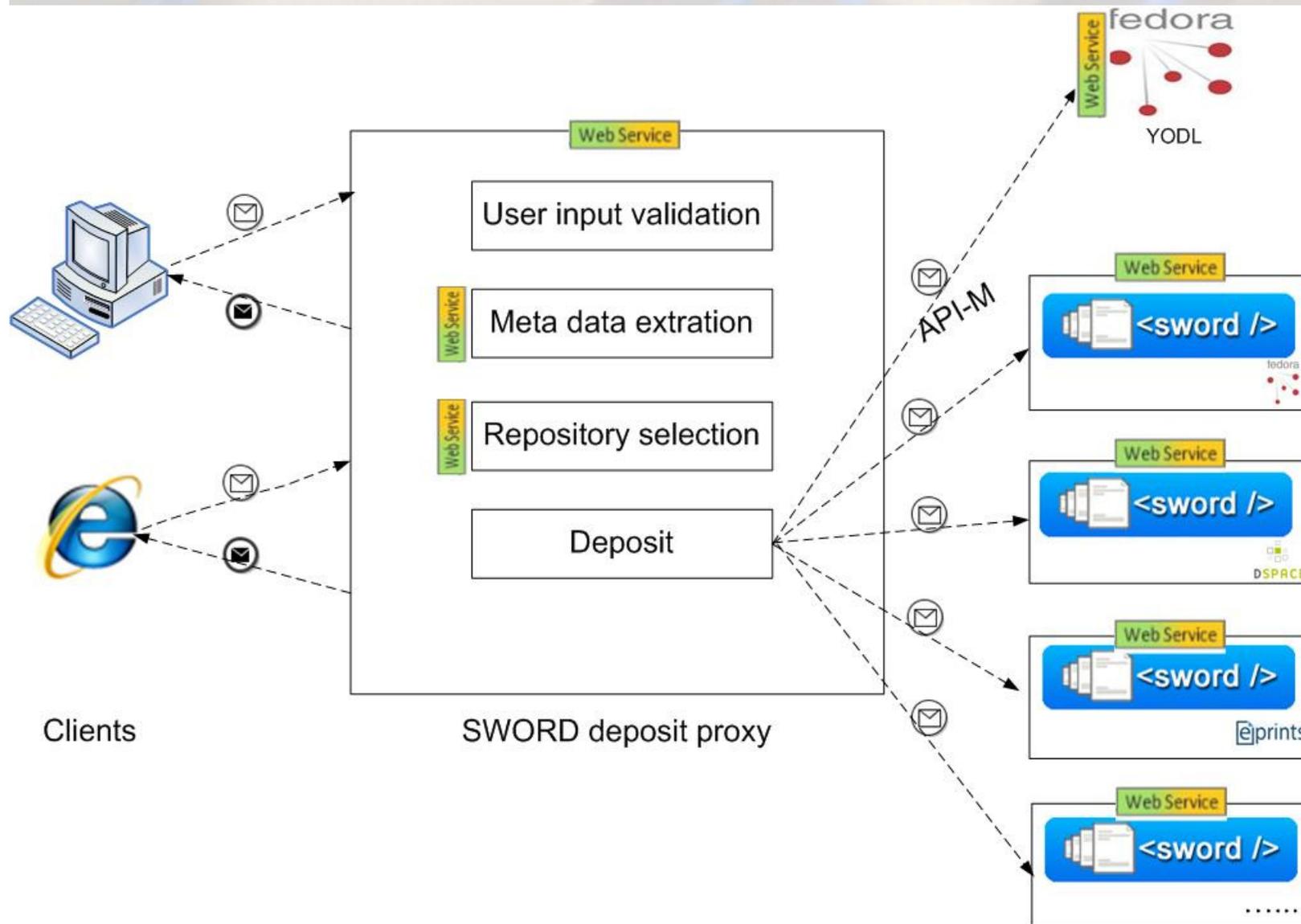
ID	File	Element	Text	Pending	Approve	Deny
1	pre_work.agent.name.xml	name	Anarchitecture	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	pre_work.agent.name.xml	name	Broodthaers, Marcel	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	pre_work.agent.name.xml	name	Man Ray	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	pre_work.agent.name.xml	name	Matta-Clark, Gordon	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	pre_work.agent.name.xml	name	Varo, Remedios	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

Approve/Deny selected item(s)

# Low barrier self-deposit

- Tool for **anyone** at York (and WRRO colleagues) 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

# 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*