

# Exporting/Importing Fedora 3 Data for Testing

This is the process that was used to export a set of test fedora 3 data from Media Collections Online and import it on mallorn. This process currently ignores the actual derivative files, so imported items will not be streamable.

## Export dataset from fedora on IU Media Collection Online (palm)

create export folder that includes the folders collection, mediaobject, masterfile, and derivative

check variable values within and run:

```
/srv/avalon/avalon_r5/current/script/export_test_data.rb
```

This will generate a bunch of foxml files withing those four directories

## Backup fedora and solr on mallorn

Zip up /srv/fedora and copy to safe place

Zip up /user/local/solr/avalon and copy to a safe place

## Import dataset into mallorn

copy exported data to a local place on drive (ie /var/www/avalon/export). It needs to include the folders collection, mediaobject, masterfile, and derivative.

check variable values within and run:

```
/var/www/avalon/current/script/import_test_data.rb
```

reindex when import is complete

```
<avalon> bundle exec rake avalon:reindex
```

## Export script

```
collection_pids = ['avalon:19844', 'avalon:778', 'avalon:966', 'avalon:19816']
export_base_dir = '/srv/avalon/avalon_r5/export' #must contain a directory for each model imported (collection,
mediaobject, masterfile, derivative)
media_object_pids = []
master_file_pids = []
derivative_pids = []
collection_pids.each do |collection_pid|
  puts collection_pid
  collection = Admin::Collection.find(collection_pid)
  media_objects = MediaObject.where({is_member_of_collection_ssim: "info:fedora/#{collection_pid}"})
  media_objects.each do |media_object|
    puts " "+media_object.pid
    media_object_pids << media_object.pid
    media_object.parts.each do |masterfile|
      puts " "+masterfile.pid
      master_file_pids << masterfile.pid
      masterfile.derivatives.each do |derivative|
        puts " "+derivative.pid
        derivative_pids << derivative.pid
      end
    end
  end
end
end
collection_pids.each do |pid|
  export_file = "#{export_base_dir}/collection/#{pid}.xml"
  command = "curl http://localhost:8983/fedora/objects/#{pid}/export?context=archive --user fedoraAdmin:
fedoraAdmin > #{export_file}"
  system(command)
end
media_object_pids.each do |pid|
  export_file = "#{export_base_dir}/mediaobject/#{pid}.xml"
  command = "curl http://localhost:8983/fedora/objects/#{pid}/export?context=archive --user fedoraAdmin:
fedoraAdmin > #{export_file}"
  system(command)
end
master_file_pids.each do |pid|
  export_file = "#{export_base_dir}/masterfile/#{pid}.xml"
  command = "curl http://localhost:8983/fedora/objects/#{pid}/export?context=archive --user fedoraAdmin:
fedoraAdmin > #{export_file}"
  system(command)
end
derivative_pids.each do |pid|
  export_file = "#{export_base_dir}/derivative/#{pid}.xml"
  command = "curl http://localhost:8983/fedora/objects/#{pid}/export?context=archive --user fedoraAdmin:
fedoraAdmin > #{export_file}"
  system(command)
end
```

## Import script

```
fedora_url = 'http://localhost:8983/fedora/objects'
creds = '--user fedoraAdmin:fedoraAdmin'
import_base_dir = '/var/www/avalon/export'
params = '-H "Content-type:text/xml" -d "format=info:fedora/fedora-system:FOXML-1.1" -X POST --upload-file'
models = ['collection', 'mediaobject', 'masterfile', 'derivative']
models.each do |model|
  Dir["#{import_base_dir}/#{model}/*.xml"].each do |filename|
    pid = File.basename( filename, ".*" )
    command = "curl -vv #{creds} #{params} #{filename} #{fedora_url}/#{pid}"
    #puts command
    system(command)
  end
end
```