

TBRC DRS Deposit Workflow – Work Plan

Prepared by: Vitaly Zakuta

Last revised: 12/01/2016

I. DRS Registration – done

II. Batch processing and deposit testing – phase 1: PDS Document objects basic workflow test – **need to redo** with new donor agreement and given that DRS has had several releases since 2015

- Create and deposit to the DRS test system a batch with new BDRC donor agreement (Document content model)
- Create and deposit to the DRS test system a batch of two PDS Document objects linked to the donor agreement (use the object URN from deposit of donor agreement for linking).
 - Also link to a test version of project documentation and methodology objects previously deposited into the DRS test system. (The Batch Builder project configuration needs to be updated with new links).
 - Add Harvard Metadata Links to HOLLIS records
 - Add Related Link to BDRC (either general link to the website or related link to the particular record in the digital library)

III. Batch processing and deposit testing – phase 2: PDS Document objects -- measure speed and performance – **need to redo** with new donor agreement and given that DRS has had several releases since 2015.

- Deposit one batch of 30 PDS objects (no more than 25,000 files per batch altogether) – linked to the new donor agreement and test versions of project documentation and methodology.
- Deposit two batches of 30 PDS objects each (no more than 25,000 files per batch altogether) – linked to the new donor agreement and test versions of project documentation and methodology.

IV. Batch processing and deposit testing – phase 3: develop Batch Builder workflow for remaining content models, finalize supporting documentation and finalize DRS and HOLLIS metadata. This includes batch-loading records to HOLLIS and getting reports with HOLLIS numbers that can be used for DRS deposits.

1. Develop workflow for the rest of the content and test processing and deposits of each type of material (Vitaly will review) --*
See Appendix A for table of TBRC content types

- Print Masters (Document)
 - Outlines (Document)
 - TEI/XML (Text)
 - Archive (Opaque) – Raw, Tiff (multipage scans) –
 - XML Metadata (Text) – Work XML, Person XML, Place XML, Outline XML, Schema
 - Java Codebase (Opaque) --
2. Test depositing main content and linking to supporting content in DRS
 - Test depositing content above and linking it to previously deposited initial test versions of the Documentation, Methodology and License items
 3. Finalize DRS metadata
 - for DRS in general (DRS Owner Code, DRS Billing Code, DRS delivery access flags, NRS authority and URN information, success and failure email addresses, etc.) – confirm that all of this metadata is correct and set up in Batch Builder
 - Finalize DRS metadata for PDS Document objects in particular (related links and Harvard metadata links, admin categories, producers, HOLLIS IDs for MODS import, etc).
 4. Develop final versions of project documentation, methodology, copyright statement and donor agreement
 5. Verify with Corinna that TBRC has all the HOLLIS IDs that need to be submitted in Batch Builder to get MODS import from HOLLIS items to be deposited to the DRS and to create Harvard Metadata links
 6. End-to-end workflow test:
 - Produce a small sample batch for each TBRC content type including HOLLIS numbers for MODS import and Harvard Metadata links
 - Deposit batches into the DRS2 test/QA environment
 - Vitaly verifies the batches
 - Fix any issues found and redeposit
 - Final verification by Vitaly

IV. Production ramp-up

1. Once verification is complete, deposit supporting content to production DRS first
 - Methodology (Text)
 - Instructions for finding related copies (masters, print, PDS etc) (Document)
 - Copyright statement (Document)

2. Deposit main content – needs slow ramp up for high volume deposits¹; all of the TBRC content types below need HOLLIS IDs ready
 - Print Masters (Document) to the DRS -- deposit before PDS objects; needs HOLLIS IDs for each work record for HOLLIS MODS import and to create Harvard Metadata links
 - Outlines (Document) – deposit before PDS objects; needs HOLLIS IDs for each work record for HOLLIS MODS import and to create Harvard Metadata links
 - PDS objects (PDS Document) – deposit after Print Masters, and Outlines; add related links to Print Masters and Outlines; add Harvard Metadata links to HOLLIS records and HOLLIS IDs for MODS import from HOLLIS
 - TEI/XML (Text) – needs HOLLIS ID for each work record for HOLLIS MODS import and to create Harvard Metadata links
 - Archive (Opaque) – Raw, Tiff (multipage scans) – needs HOLLIS ID for each work record for HOLLIS MODS import and to create Harvard Metadata links
 - XML Metadata (Text) – Work XML, Person XML, Place XML, Outline XML, Schema -- needs HOLLIS ID for each work record for HOLLIS MODS import and to create Harvard Metadata links
 - Java Codebase (Opaque) -- needs HOLLIS ID for the TBRC electronic resource record in HOLLIS for HOLLIS MODS import and to create Harvard Metadata links
3. Send delivery URNs and HOLLIS IDs to Corinna so she can batch-populate holdings records
4. Volume test with objects that have one file per object (Print Masters, Outlines, TEI XML, XML Metadata)
5. Adjust workflow according to performance test results

¹ First deposit 1 batch of no more than 30 PDS objects each and no more than 25,000 files per batch while LTS monitors DRSs performance; then deposit 10 batches in one day of no more than 30 PDS objects each and no more than 25,000 files per batch while LTS monitors DRSs performance; then ramp up to up 250,000 files per day. For objects that contain one file per object, slowly ramp up to 1,000 objects per batch.

Appendix A: Table of TBRC Content Types

Content type	DRS content model	Supporting content y/n	Needs HOLLIS IDs y/n	Dependencies on other content types
Methodology	Text	y	n	Needs to be deposited first so main content can be linked to it
Instructions	Document	y	n	Needs to be deposited first so main content can be linked to it
Copyright statement	Document	y	n	Needs to be deposited first so main content can be linked to it
Print masters	Document	n	y	Needs to be deposited after supporting content and before PDS Document Objects (in order to create related links from PDS Documents to Print masters)
Outlines	Document	n	y	Needs to be deposited after supporting content and before PDS Document Objects in order to create related links from PDS Documents to Outlines
Page turned objects – one per volume	PDS Document	n	y	Needs to be deposited after supporting content and after Print masters and Outlines have been deposited. Needs to incorporate URNs to Print masters and Outlines in Related Links
TEI/XML	Text	n	y	Needs to be deposited after supporting content
Archive – Raw, Tiff (multipage scans)	Opaque	n	y	Needs to be deposited after supporting content
XML Metadata Work XML, Person XML, Place XML, Outline XML, Schema XML	Text	n	y	Needs to be deposited after supporting content
Java Codebase	Opaque	n	y	Needs to be deposited after supporting content