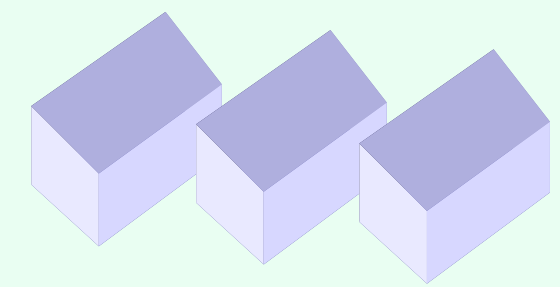# Scanning TRAIL Project Technical Reports
A Workflow for a Large-Scale Collaborative Digitization Effort
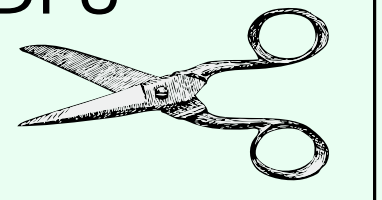
Stage One: Inventory

When boxes arrive at UNT DPU, reports are inventoried and noted in a text file

Bound reports are pre-cut in the DPU and routed to Preservation

Already-cut reports are put in the scanning queue

Reports are cut at Preservation and returned to the DPU

Scanning queue

Stage Two: Scanning

1. Disbound pages are scanned on a duplex scanner

2. Fold-out pages and color/gray-scale pages are flagged and placeholders are scanned

3. Files are renamed; initial quality control (QC) checks for missing pages

4. Flagged pages are scanned on a large flatbed scanner or planetary scanner

5. Excessively large fold-out pages are scanned in pieces and "stitched" in Photoshop
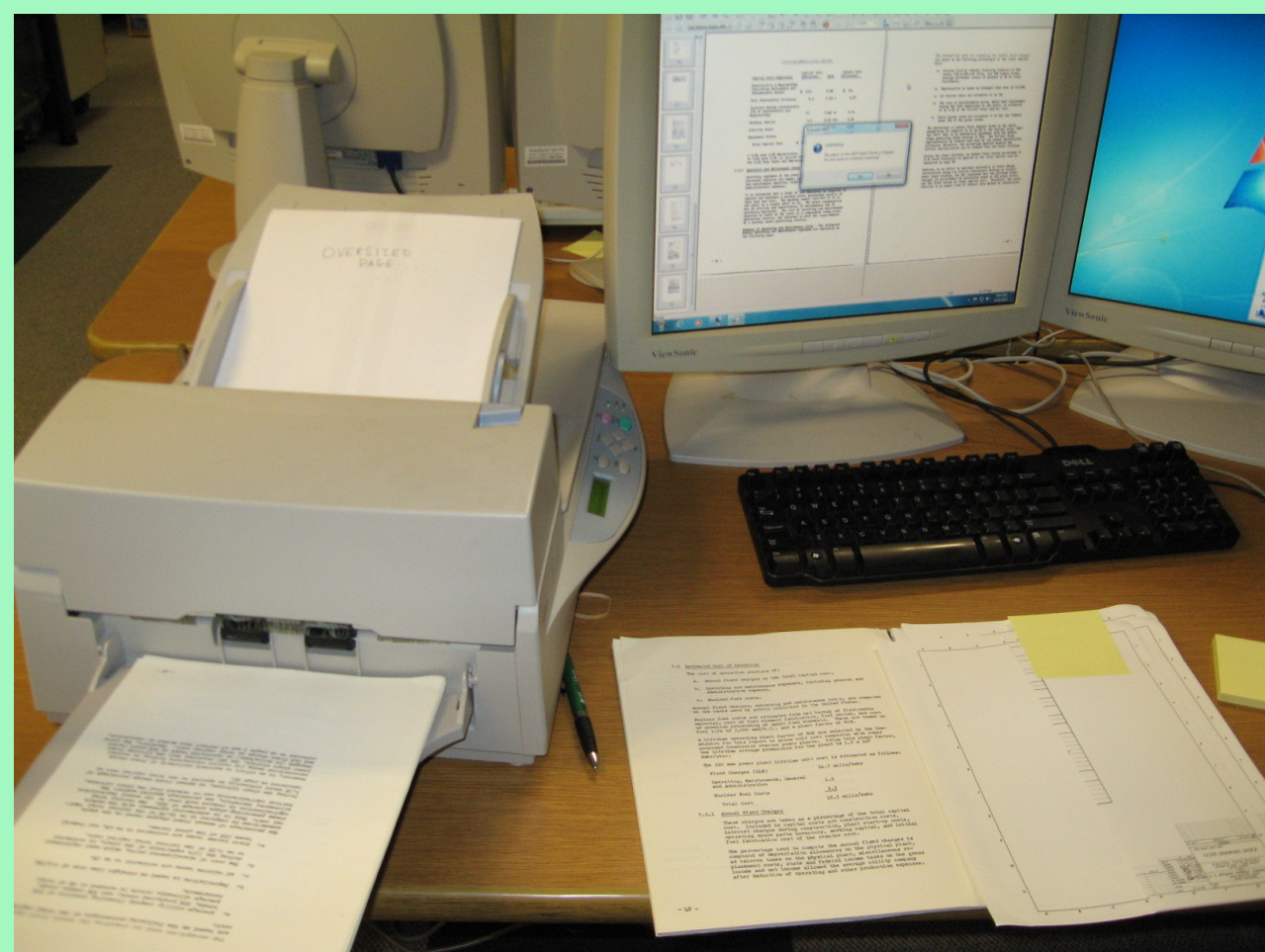
[1-2] Loose pages scanning on a Fujitsu fi-4340C Image Scanner; oversized pages are marked with sticky notes and filler pages are scanned in their place
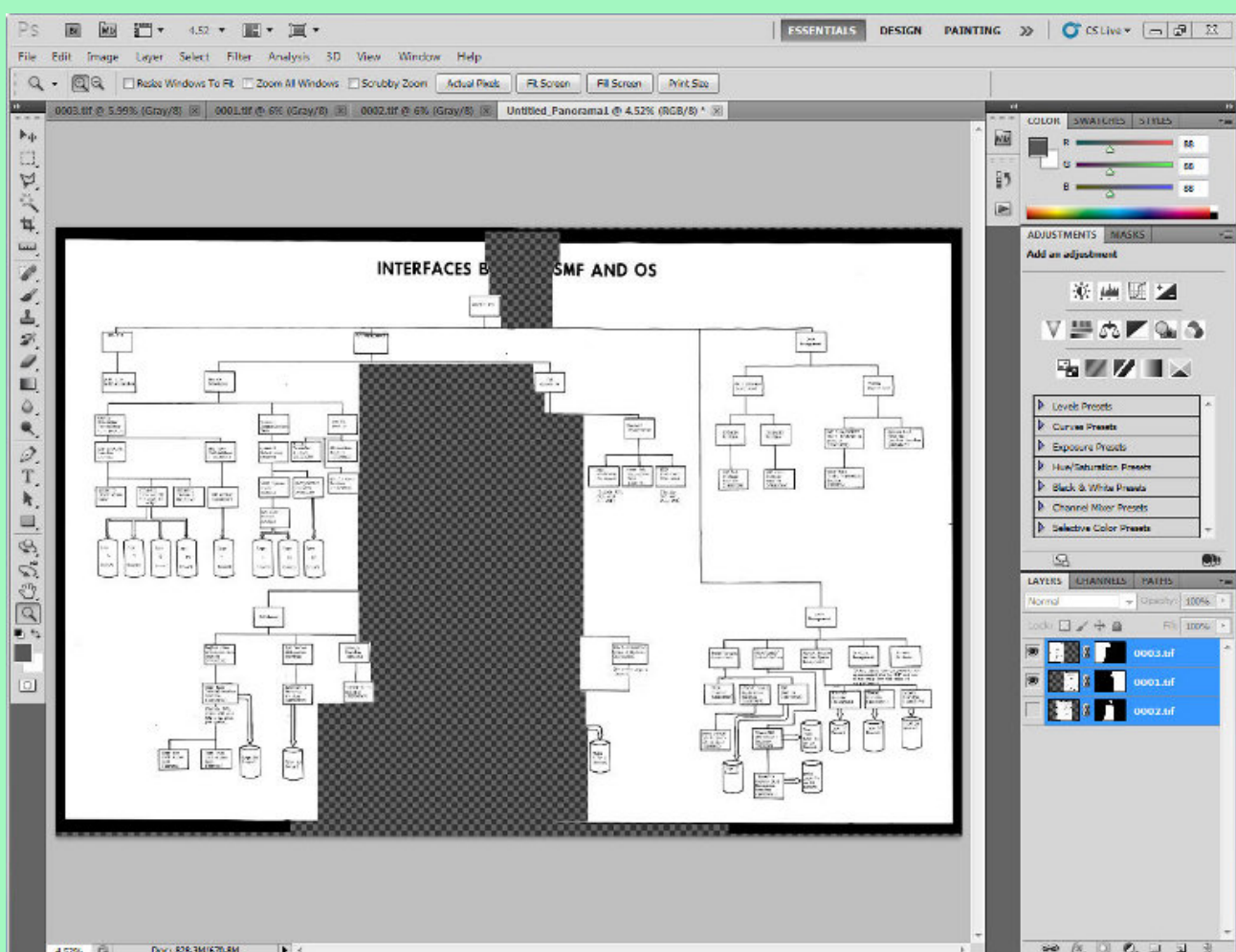
[5] Large pages are stitched together in Photoshop from scanned pieces

Stage Three: Processing

1. Preliminary QC step to identify scanning errors

2. Oversized pages are processed manually in Photoshop

3. All other pages are processed with Scan Tailor

4. Oversized page files are re-integrated with the page files

5. As necessary, pages are rotated so that all text is horizontal

6. Final QC check

7. A PDF file is created for each report and files are sorted into a standard internal folder structure
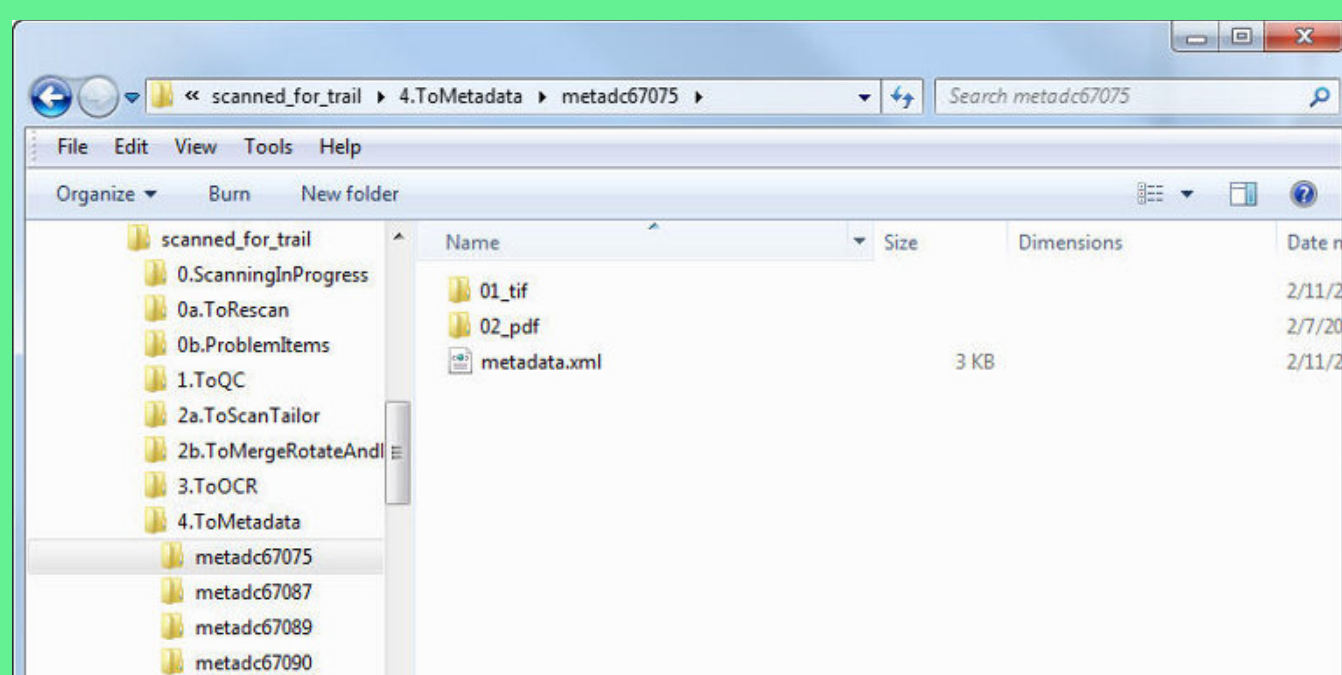
8. Reports are flagged for OCR in the workflow
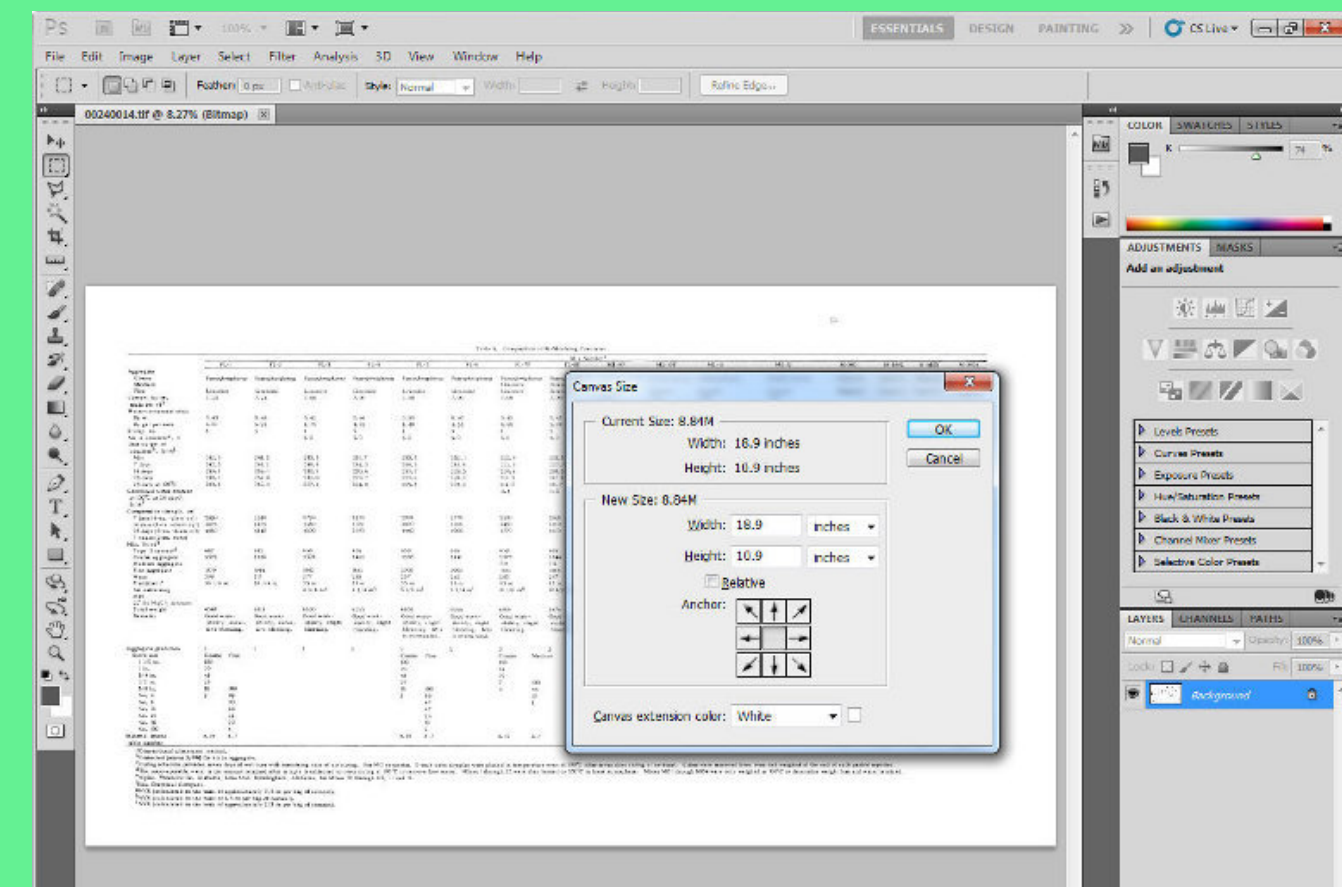
[2] Oversized pages are deskewed and the size is changed to match the physical page
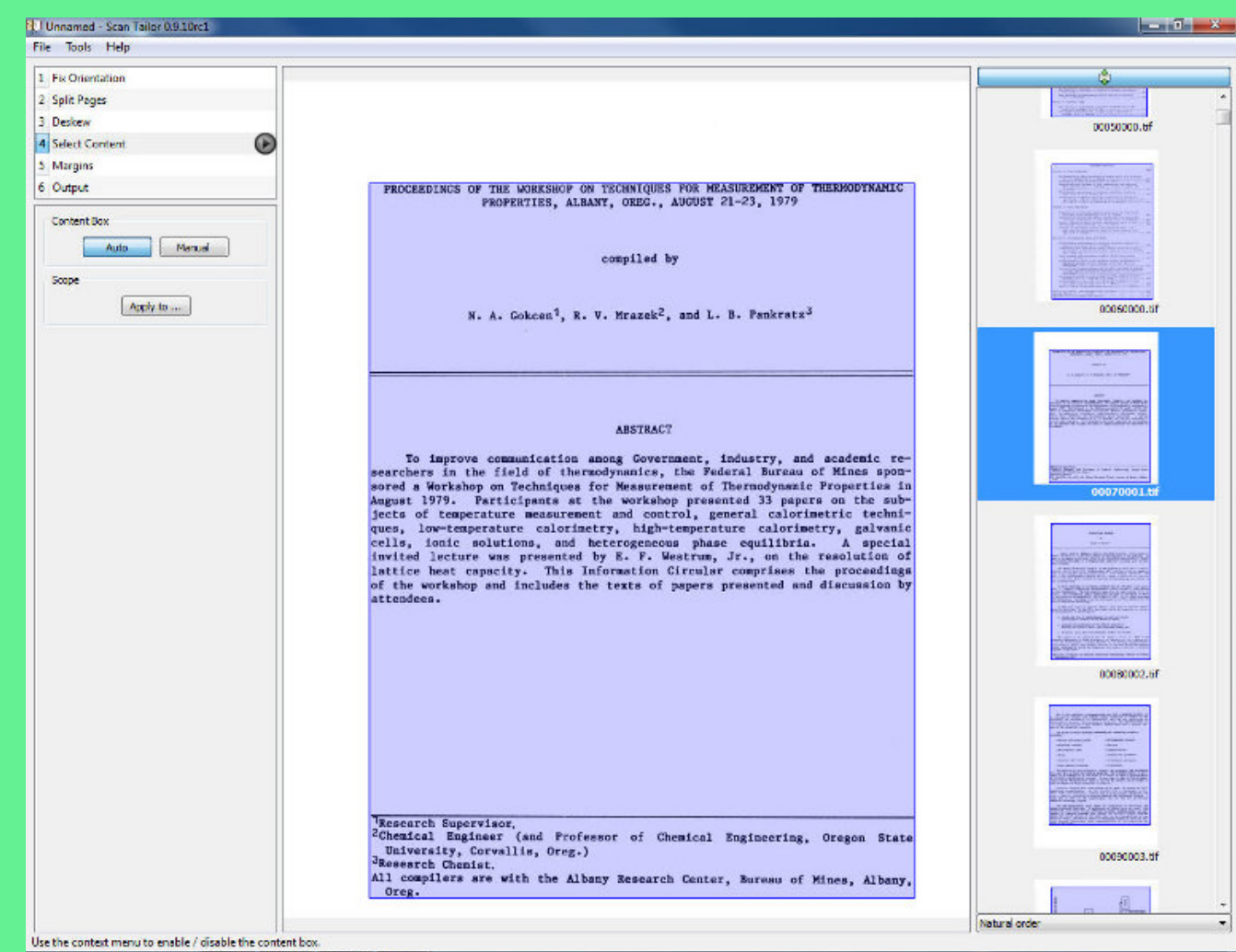
[3] Scan Tailor deskews the text, cleans the pages, resizes each image to a uniform page size, and compresses the files
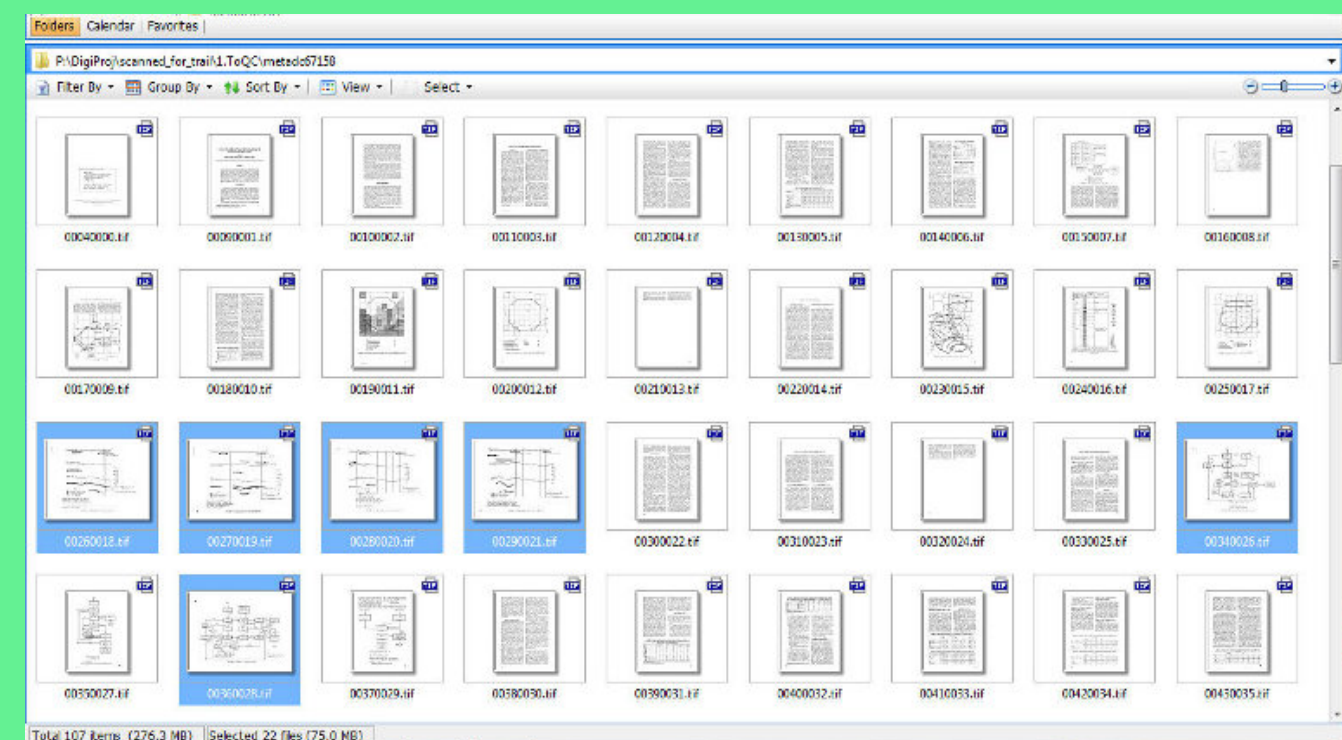
[5] Pages with text printed vertically are rotated so that content is readable
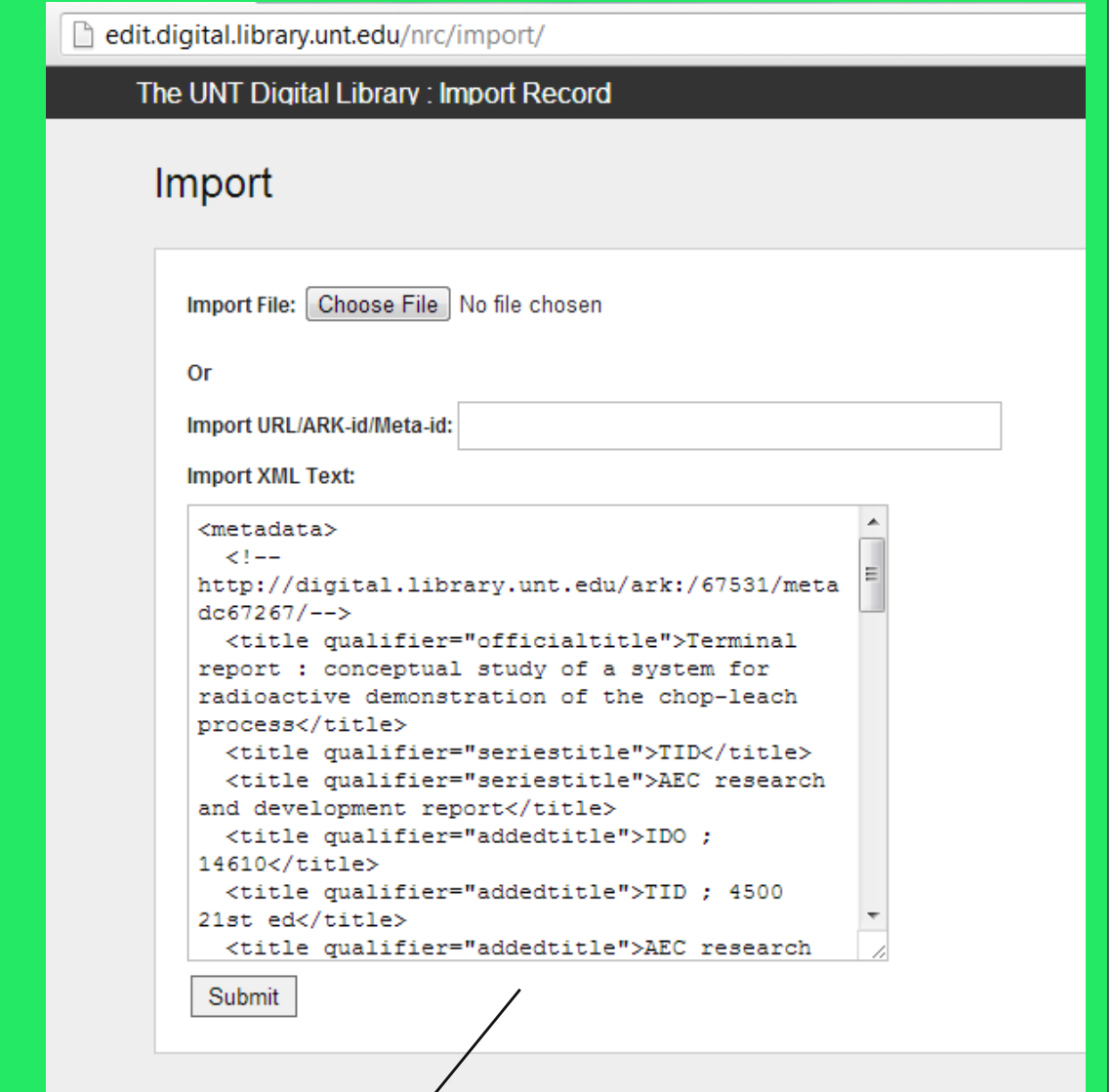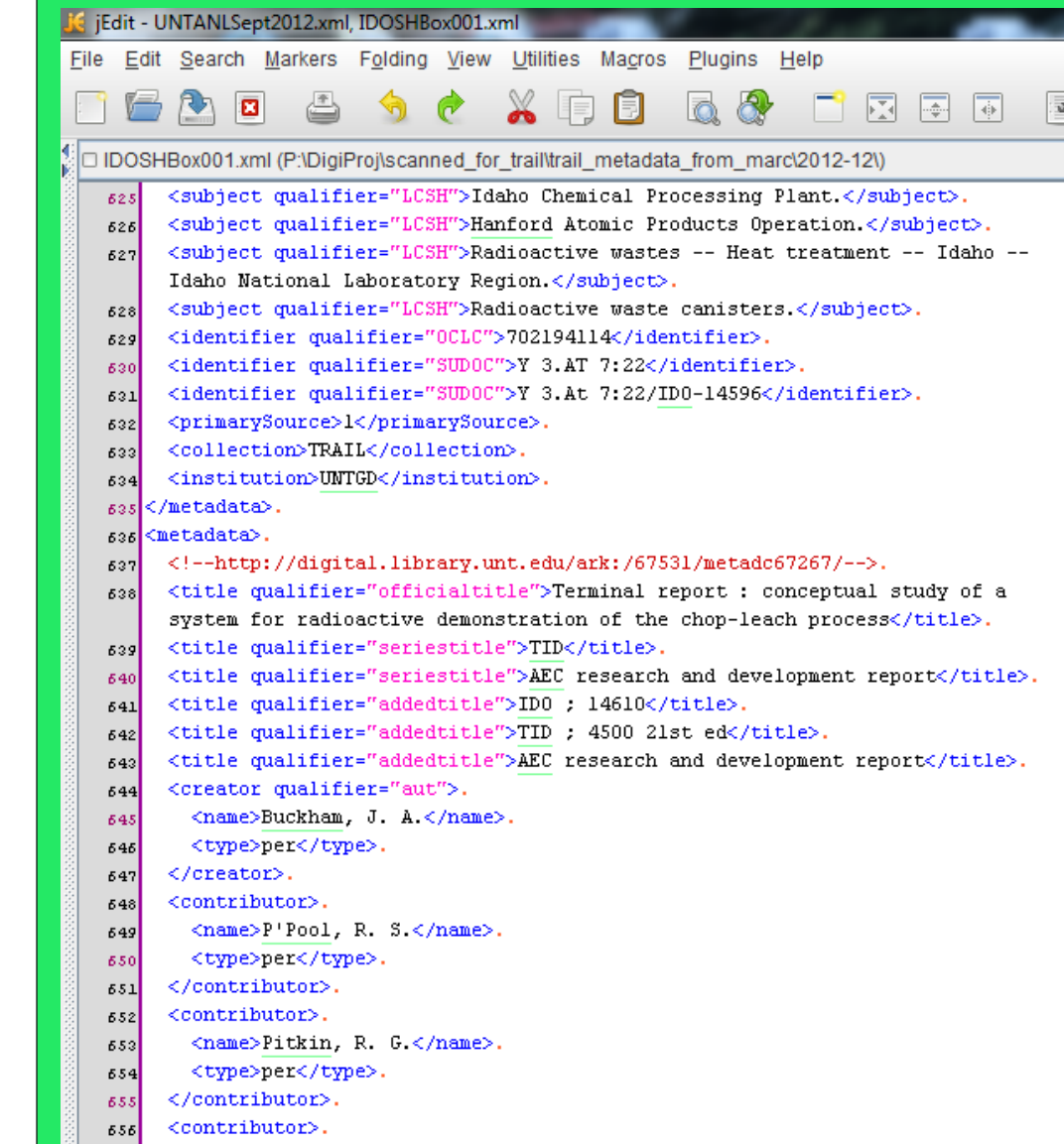
[7] All files for an individual report are in a folder labeled with the unique identifier (metadc##); each file format is in a separate, nested folder and the metadata file lives in the parent folder (once completed)
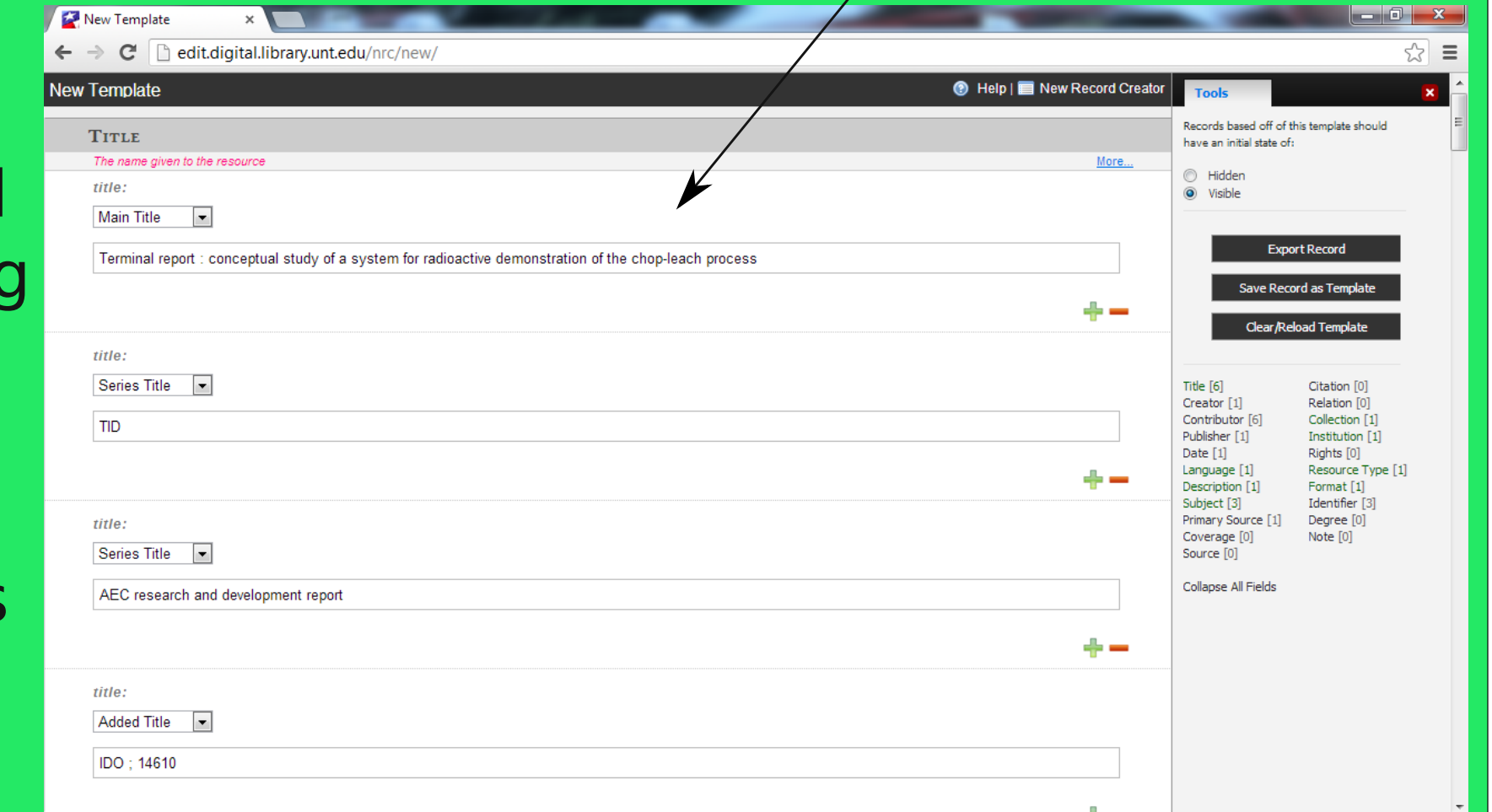
Stage Four: Metadata

MARC records are converted to UNTL xml files

Individual records are pasted into the web-based New Record Creator

The record is displayed in an editing form - information is changed or added as necessary

Stage Five: Online

UNT DIGITAL LIBRARY

The Development and Evaluation of Effective Symbol Signs

UNIVERSITY OF NORTH TEXAS

Hannah Tarver   hannah.tarver@unt.edu

Contributors: Shannon Willis
                Meredith Lahaie

UNT Libraries