

# Digital Preservation: Tales from the Precipice between theory and practice

## Metadata that supports preservation management

- **Some fields are automatically captured** through DSpace system. Others required deeper analysis.
- **Data was transferred from deprecated fields** for consistent usage and format.
- **IR curation tools allow for batch updates** to only metadata without having to re-ingest the digital files.

Data	PREMIS	Qualified Dublin Core
Size	in size under objectCharacteristics	dc.format.extent
CHECKSUM and CHECKSUMTYPE	in fixity under objectCharacteristics	dc.description.provenance
MIMETYPE	in format under objectCharacteristics	dc.format.mimetype
Unique designation	In IdentifierValue objectCharacteristics	dc.identifier.uri
Format Designation	in format under objectCharacteristics	dc.type.dcmi
Creation of object info	objectDateCreated ByApplication	dc.date.digital
Creation of object info	objectCreating	dc digitization .specifications

Metadata Crosswalk of PREMIS to Qualified Dublin Core (QDC) Elements



Monica Rivero  
Amanda Focke  
Ying Jin  
Sid Byrd

<https://scholarship.rice.edu/handle/1911/75103>  
DLF 2013 Austin, Texas

## The Challenges

Inconsistent usage of metadata elements and data values across time. For example:

<b>dc.type</b> dc.format.medium dc.format.original dc.format.extent dc.format.mimetype (bitstream format)	<b>dc.rights</b> <b>dc.rights.uri</b> bundle: CC-LICENSE text bundle: LICENSE text  <b>dc.identifier.digital</b> dc.identifier dc.identifier.other
--	---

Most visits originate from outside the IR

Master files stored at secured but difficult to access location without any structured system of organization

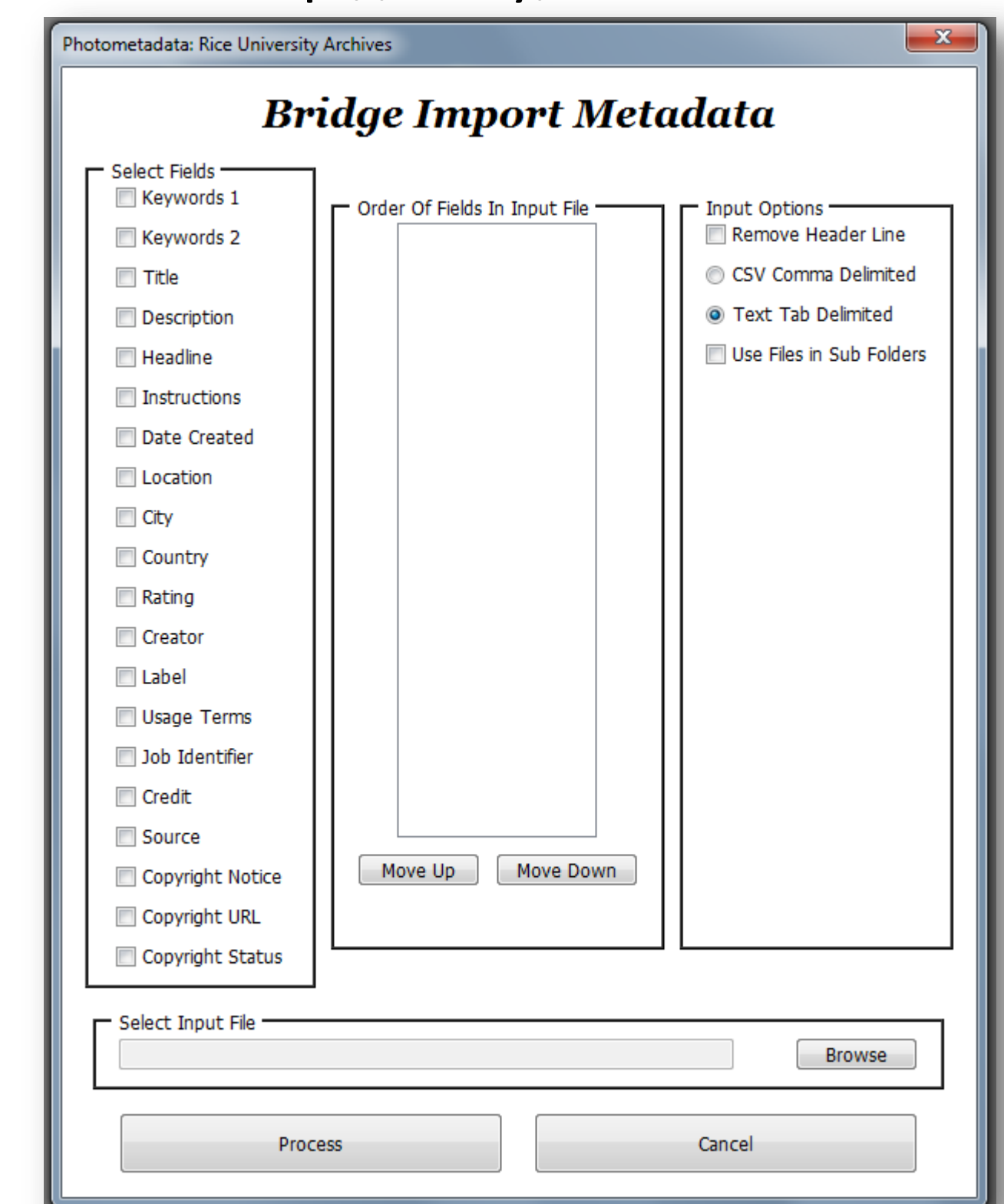
## Ingest master files to the IR

- **Better management of high quality files** by ingesting masters to the IR in “dark” or hidden bundles.
- **The relationship between the master file and the metadata record is maintained** in this environment and master files are automatically included in Duracloud service for preservation.
- **Master files can be easily accessed by archivists** for ad hoc patron requests.
- **A custom ingest perl script allows file-only ingests**, so the process can be applied to pre-existing collections.

## Embed metadata to Files

To ensure basic description travels with images and as a fail safe for masters.

Photometadata import script  
<http://bit.ly/1aTKoAx>



Then overnight script converts TIFFs to JPEG2000, retaining embedded metadata

## Tools

- **Droid** <http://sourceforge.net/projects/droid/>
- **Jhove** <http://sourceforge.net/projects/jhove/>
- **ExiftTool** <http://www.sno.phy.queensu.ca/~phil/exiftool/>
- **DSpace:**
  - Metadata Batch import/export
  - Curation tasks: Profile Bitstreams Format, Count Masters