Assessing the Preservation of DPX Image Sequences within MoMA’s Digital Preservation Ecosystem
Six-channel video (color, sound)
MoMA

Digital Repository for Museum Collections (DRMC)
MoMA

Physical artwork folder

Digital asset management system (DAM)

Physical art storage

Collection management system (TMS)

Electronic artwork folder

Digital Repository (DRMC)
01111010
10010001
11111011
10110111
01001010

Dear future,
This string of bytes represent a video file. It is 51 minutes long, and is in the MOV file format. It’s dimensions are 720x1280 pixels, and it has a color space of RGB.
Sincerely,
Archivematica
Lime Kiln Club Field Day
1913/2014
MoMA

File: Film - Confidential
3799.DG

Lime Kiln Club Field Day
Bert Williams Picnic
Studio: Biograph Company
1913
Digital

<table>
<thead>
<tr>
<th>Component Number</th>
<th>Component Name</th>
<th>Count</th>
<th>Active in TMS</th>
<th>Component Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>3799.A/1</td>
<td>.mov ProRes 422</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R1.x1</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R1.x2</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R2.x1</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R2.x2</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R3.x1</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R3.x2</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R4.x1</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R4.x2</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R5.x1</td>
<td>2K DPX color corrected</td>
<td></td>
<td></td>
<td>Part of artwork</td>
</tr>
</tbody>
</table>

Description:
Made in 2014 by Deluxe as 2K raw scans from FGM #8660.

Component Barcode:

Storage Comments:

Preparation Comments:

Installation Comments:

Flex Fields:

Add:

Delete:
# MoMA

**Film - Confidential**

**3799.DG**

**Lime Kiln Club Field Day**
*Bert Williams Picnic*
Studio: Biograph Company

1913
**Digital**

<table>
<thead>
<tr>
<th>Component Number</th>
<th>Component Name</th>
<th>Count</th>
<th>Active in TMS</th>
<th>Component Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>3799.DG</td>
<td>.mov ProRes 422 (digital file)</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R1.x1</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R1.x2</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R2.x1</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R2.x2</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R3.x1</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R3.x2</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R4.x1</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R4.x2</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
<tr>
<td>8660.R5.x1</td>
<td>2K DPX color corrected</td>
<td>✔️</td>
<td>✔️</td>
<td>Part of artwork</td>
</tr>
</tbody>
</table>

**Description**

**Physical Description**
Made in 2014 by Deluxe as 2K raw scans from FGM #8660.

**Storage Comments**

**Preparation Comments**

**Installation Comments**

**Component Barcode**

**Dimensions**
MoMA

In TBs

Total Film Dept Material

FY18  FY19  FY20  FY21  FY22
I'll end it around here or another slide that basically describes the problem we face and you can pick it up with the analysis of DPX and how we're trying to solve the problem. Sound good?

Peter Oleksik; 23/10/2018
MoMA

The challenges of DPX

- Massive quantity of DPX files coming into MoMA’s collection
- Complexity of DPX in terms of ingest
- Ingest workflow was designed around single file ingest, not complex multi-file sequences
- How to capture the preservation decision making process
- Adapting our ever-evolving condition assessment and documentation workflows to multi-file sequences
- Bridging knowledge gaps between vendors and conservation department
MoMA

The challenges of DPX

- Lack of consensus amongst the field in terms of the handling and/or preservation of DPX materials
- File format sustainability unknowns
- Curation of material (what to keep and why?)
Research into the DPX format

- Researching and examining the DPX file format.
  - What it is?
  - How is it supposed to be rendered or played back?
- Characterization methods, metadata extraction, embedding metadata, tools
- How others are working or managing large quantities of DPX files
Establishment and Refinement of Internal Procedures and Workflows (within and across depts) to capture and assess material in an efficient and responsible manner

Trying to understand the significant properties of DPX

Strategies to document process history and decision making in the DPX file format itself
MoMA
Metadata

● Reviewing FADGI guidelines
● Analysis of Tools
  ○ MediaInfo
  ○ Exiftool
  ○ Imagemagick
  ○ DPX Analytics
  ○ GraphicsMagick
● Analysis of Terminology
  ○ how fields are translated
    From one tool to the next
MoMA

Playback, transcoding, and packaging tools

- Playback
  - ffplay
  - DaVinci Resolve
  - Adobe Premiere
- Looking into lossless transcoding options for efficient storage and preservation
  - MXF/OP1a
  - MXF/JPEG2000
  - FFV1/MKV
- Lossless compression/packaging tools
  - ZIP
  - TAR
Preliminary results

1. 2K resolution DPX directory composed of 14,619 files amounted to 103.03 GB transformed into a 70.51 GB MKV using RAWcooked
2. 2K resolution DPX directory composed of 49,003 files, amounting to 616.99 GB transformed into a 210.14 GB MKV using RAWcooked
3. 2K resolution DPX directory composed of 17,525 files amounting to 145.57 GB transformed into a 50.52 GB zip directory
1. ~5 hours to process a valid bag of FFV1/MKV in Archivematica
2. Research into 'store AIP' failures
3. Working with Artefactual to create a separate archivematica (v 1.8) pipeline exclusively for processing DPX
MoMA

Next steps

- More testing including escrow retrieval
- Deep analysis of our research results
- Develop a methodology for the condition assessment of DPX files
- Attending #NTTW3
- Starting robust dialogue with colleagues and peers including a convening for a peer forum in November to see and discuss what others in the field are doing
Questions?