Quality Control Experiences and Effectiveness in a Large-Scale Film Digitization Project

Peter Schallauer

JOANNEUM RESEARCH DIGITAL

Darrell Myers

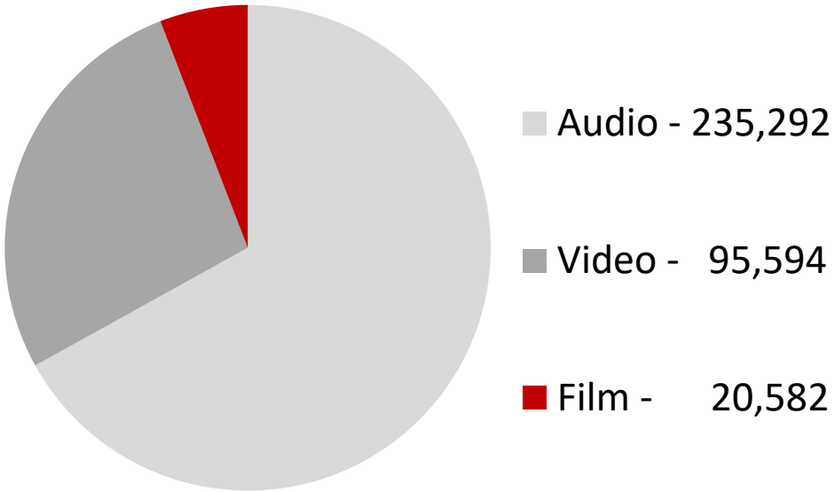
AMIA Conference
Nov. 19th 2020, Online

Media Digitization & Preservation Initiative

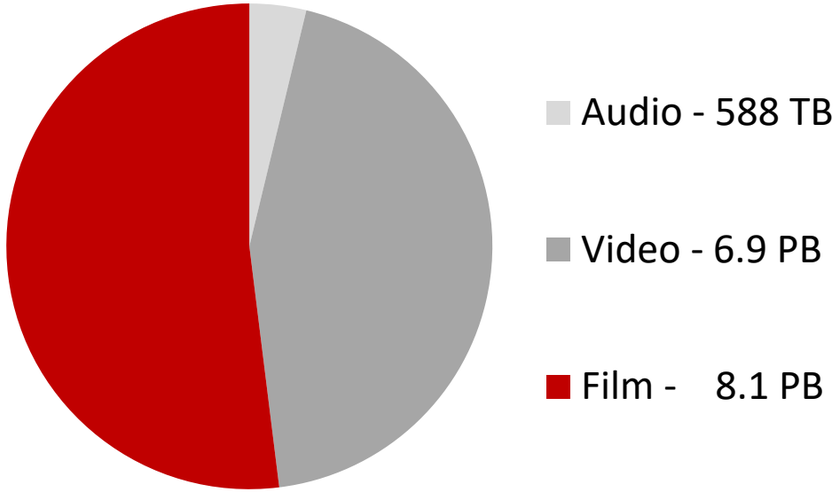
- Digitally preserve all significant audio, video, and film (~24 Petabytes)
  - Audio/Video: ~325,000 recordings within 5 years
  - **Film: ~25,000 reels within 3 years**
- Completed: IU Bicentennial year (2020)

Preserved Titles
11/8/20

Number of Items Digitized



Storage Usage (15.7 PB)



Hours

Audio – 17,8255.03

Video – 11,4884.03

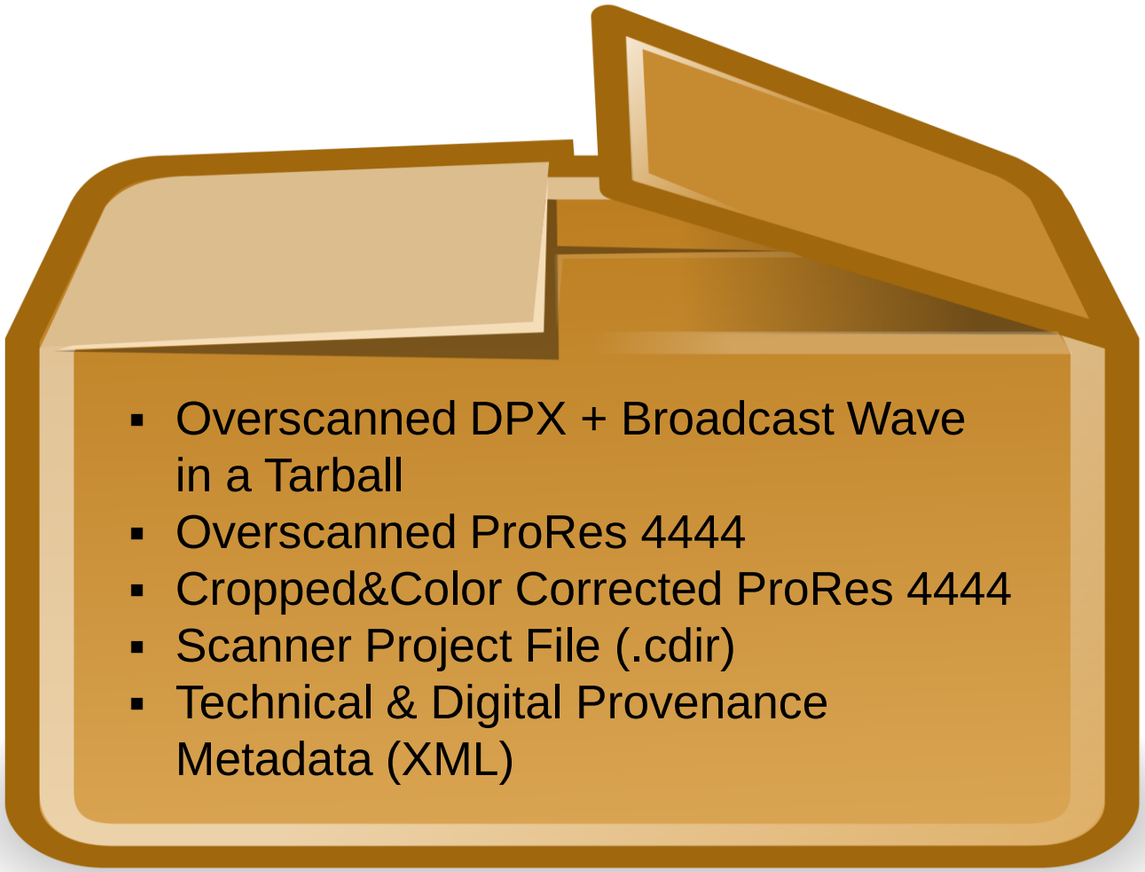
Film – 5,959.69

Preserved Titles Detail
11/8/20

Type	Count	SDA Usage(G)	Avg Size(G)	Duration(h)	Avg Duration(m)
1-Inch Open Reel Video Tape	4081	147499.70	36.14	2522.97	37.09
1/2-Inch Open Reel Video Tape	516	12933.91	25.07	284.37	33.07
2-Inch Open Reel Video Tape	11	1251.77	113.80	8.23	44.90
45	4032	2084.07	0.52	513.77	7.65
78	34545	16590.46	0.48	3693.68	6.42
8mm Video	1014	79439.60	78.34	1364.10	80.72
Aluminum Disc	1299	846.87	0.65	187.68	8.67
Audiocassette	54494	262607.50	4.82	64514.28	71.03
Betacam	19206	838767.80	43.67	13806.07	43.13
Betamax	1228	106029.25	86.34	1963.95	95.96
CD-R	12156	14549.34	1.20	10472.77	51.69
Cylinder	6643	2031.79	0.31	322.43	2.91
DAT	8996	24929.38	2.77	17671.37	117.86
DV	2681	115314.12	43.01	2960.24	66.25
DVD	4779	32784.69	6.86	4987.06	62.61
Film	20582	8198652.21	398.34	5959.69	17.37
Lacquer Disc	6109	6794.47	1.11	1586.73	15.58
LP	38657	114740.65	2.97	28130.52	43.66
Open Reel Audio Tape	68094	142700.17	2.10	51055.63	44.99
Other Analog Sound Disc	267	464.64	1.74	106.17	23.86
U-matic	15279	541644.15	35.45	9321.25	36.60
VHS	46799	4999635.08	106.83	77665.79	99.57
All Types	351468	15662291.64	44.56	299098.76	51.06

Quality Control Needs - Film

- Submission Package QC
  - Handle Up to 27 TB per Day
- Image and Sound QC
  - 100% Mezzanines (Overscanned+Cropped)
  - 10% Preservation Masters
  - < 16 hrs content / day
- Fast Identification, Communication & Reporting for Potential Failures



Submission Package QC - IU Build

- Fully automated QC confirming ….

✓ Package Structure
✓ Checksums
✓ File Naming Conventions
✓ Valid Bag
✓ Correct # of DPX Frames

✓ Compare Deliverables to XML:
- Durations
- Bit Rate
- Color Space
- # Audio and Video Channels
- Codec
- Frame Rate
- Height x Width
- Pixel Format

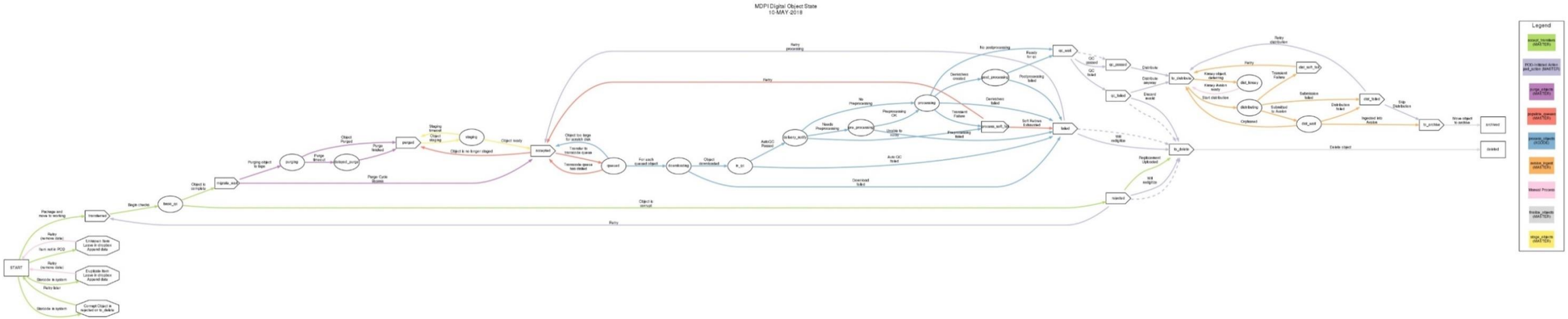


Image and Sound QC - VidiCert

- Confirm – Preservation Overscan
  - Completeness
  - No Image / Audio Loss
  - No Image / Audio Corruption
  - Playback Speed
  - Film Prepared well for Scanning
  - Faithful Representation of Original
- Confirm – Access Copy (Cropped)
  - Minimal Frame Lines
  - Acceptable Color Correction
- Create QC Report

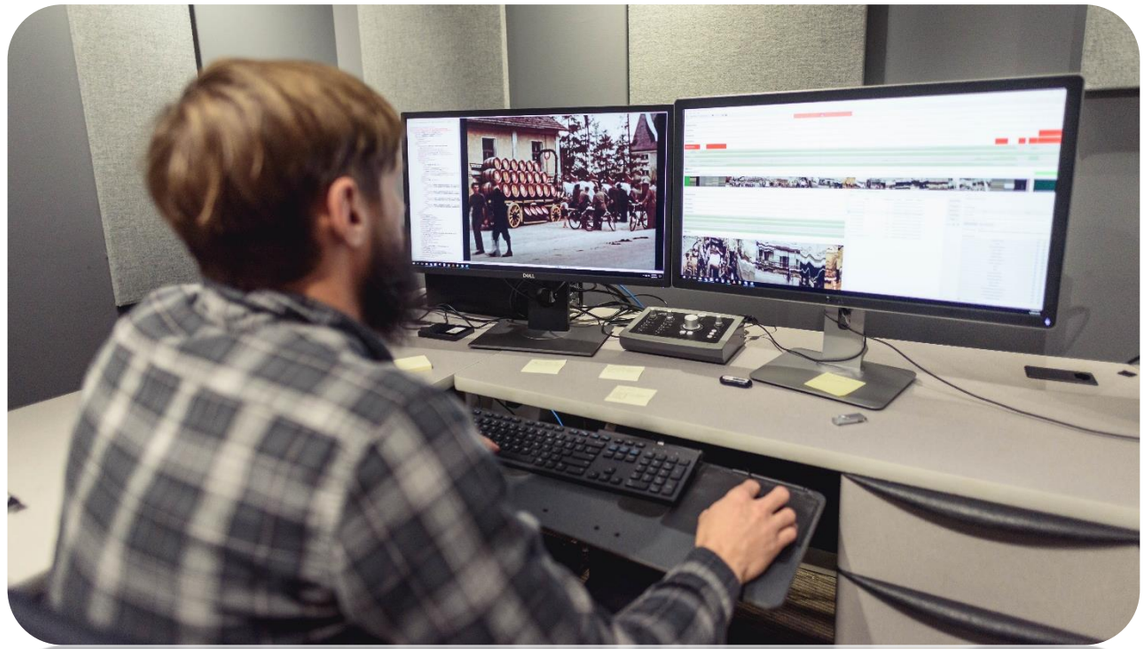


Image and Sound QC Identified Defects

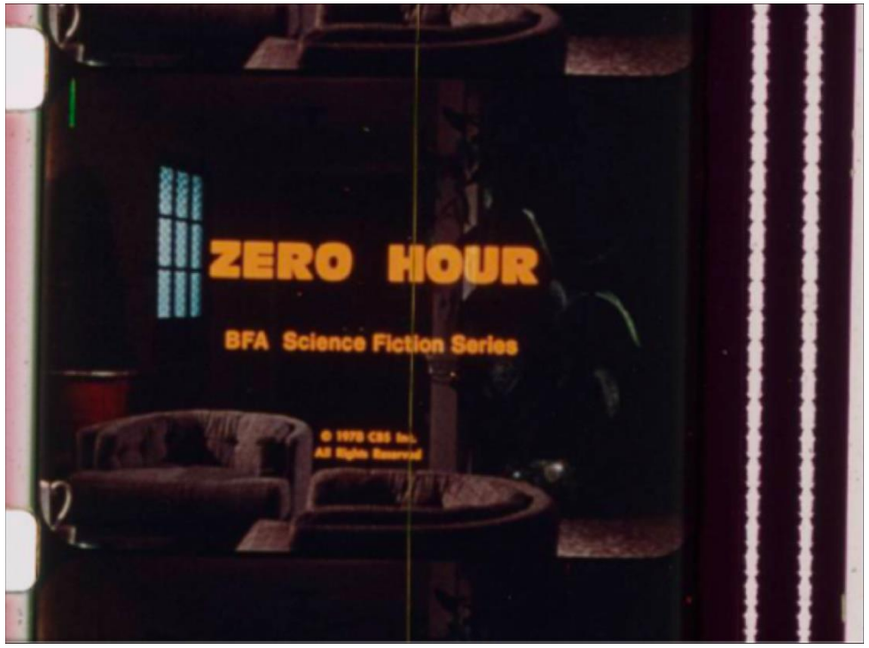
Dirt / Dust

Gamut/Clipping

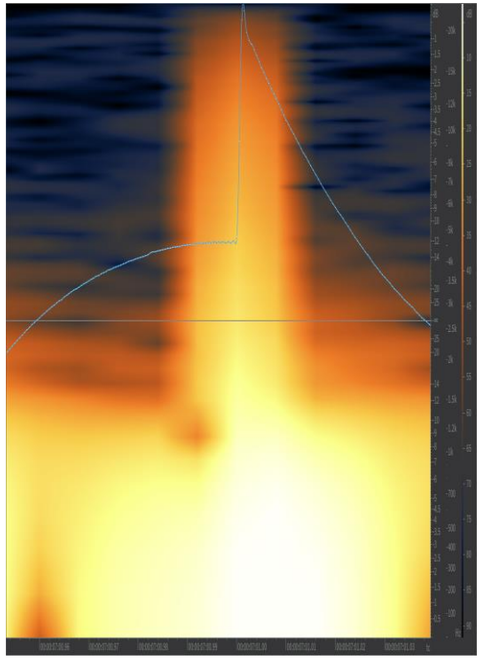
Foreign Object in Frame

Color Correction Error

Image and Sound QC Identified Defects



Crookedness



Interstitial Errors

Added Tones

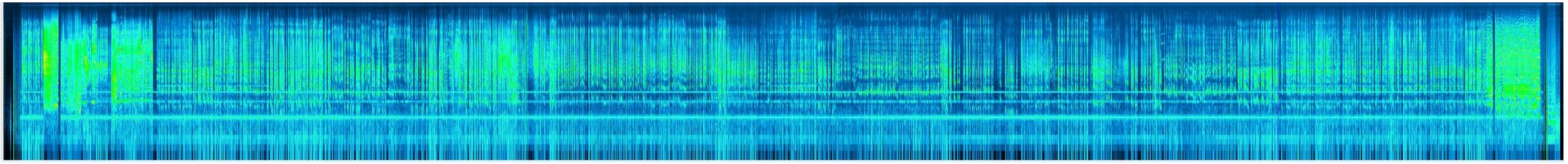
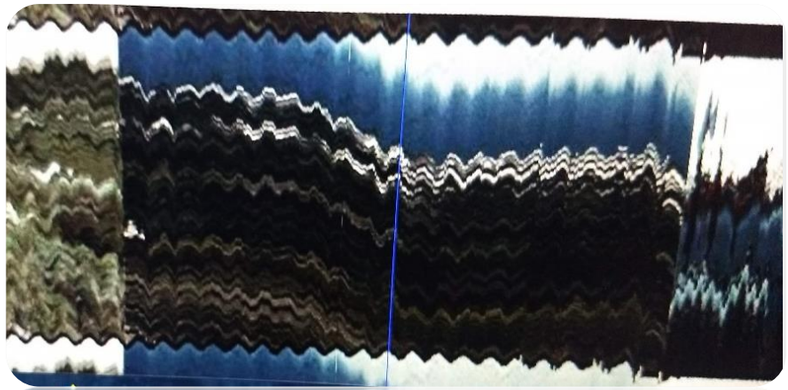


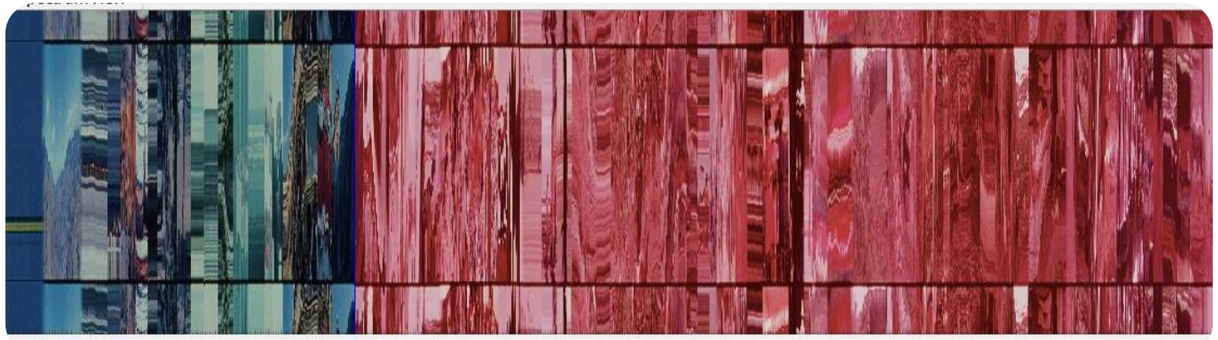
Image and Sound QC
Identified Defects & Reporting

Framing Errors & Image Loss



Unsteadiness

PDF
Reports



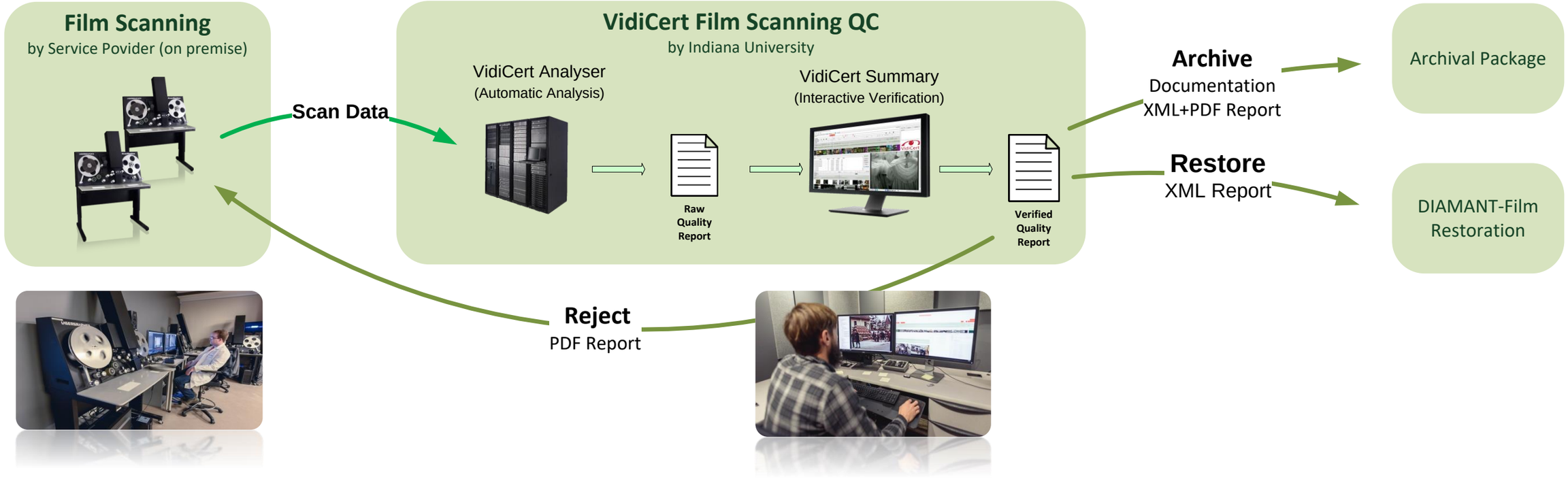
Faded Stock – Color Correction Needed

Image and Sound Quality Control

How does it work?

How effective is it in operations?

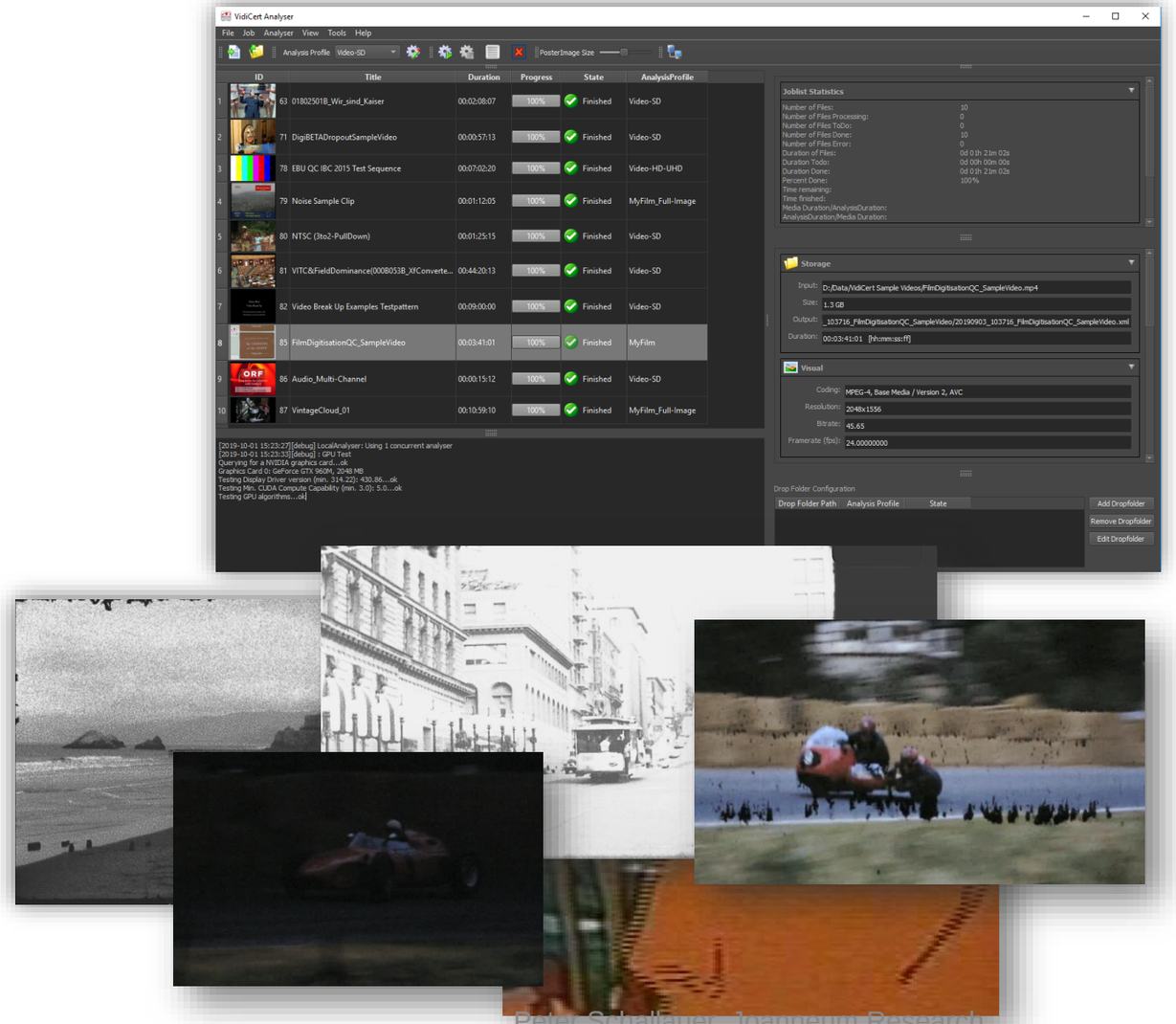
Image and Sound QC
How does it work?



- Scanning Operations
  - 2 scanners,
    two 6 hrs shifts per scanner per day
  - Creating up to 4 hrs of content per shift
    = up to 16hrs of content per day
- QC Operations
  - QC of overscan and crop version = **up to 32hrs of QC content per day**
  - 3 VidiCert Analyser‘s, 3 VidiCert Summary stations
  - Current **QC throughput**: **~50 files = 15 hrs per person** in a 8 hours shift

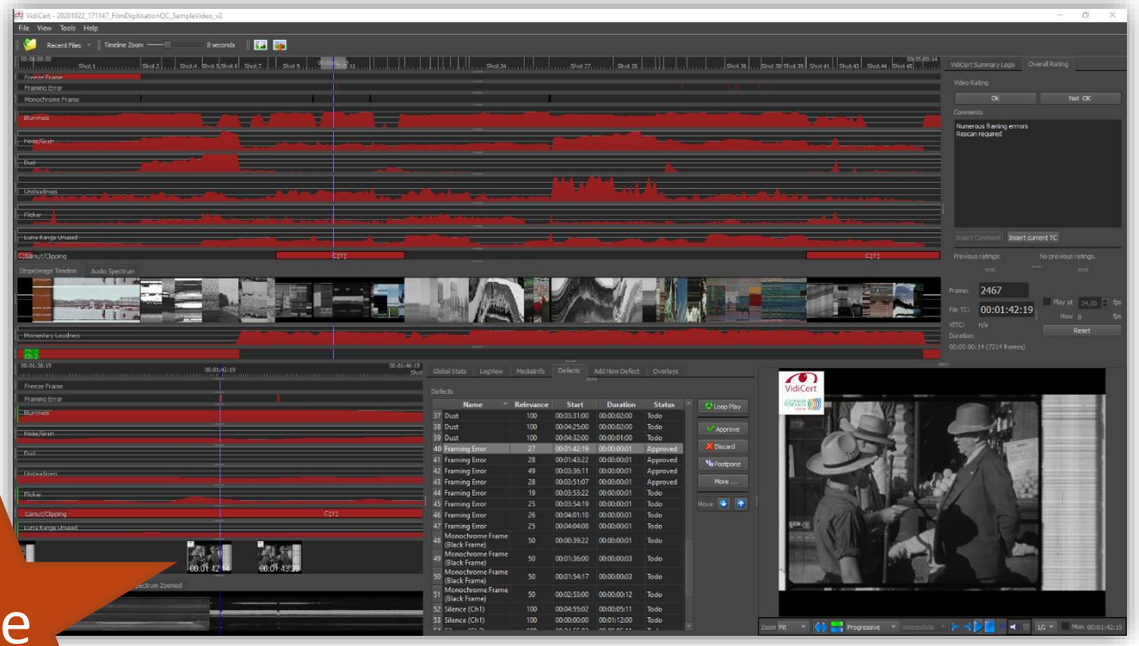
The Image and Sound QC Solution @ IU
Automated Scanning Issue Detection Functions

- Gamut/Clipping (Under&Over Exposure)
- Freeze Frame
- Framing Error
- Dust/Dirt/Hair Level
- Unsteadiness Level
- Film Grain Noise Level
- Out of Focus / Blurriness Level
- Flicker Level, Flash Light
- Contrast/Luminance Range
- Black & Single Coloured Frames
- Black Bar / Aspect Ratio
- Macroblocking
- Audio Silence, Loudness, Superimposed Sound
- Integration of scanner sensor data (optional)
  - Perforation/Shrinkage
  - Splices



The Image and Sound QC Solution @ IU
Interactive Defect Verification

- Time-efficient user interface for
  - Verifying AV digitisation issues detected automatically
  - Annotating AV issues manually
  - Checking content completeness and correctness (image, sound, timecode)
  - Documenting the QC decision
- Creating QC reports
  - Machine readable XML
  - Human readable pdf
- DIAMANT Restoration Report



Live

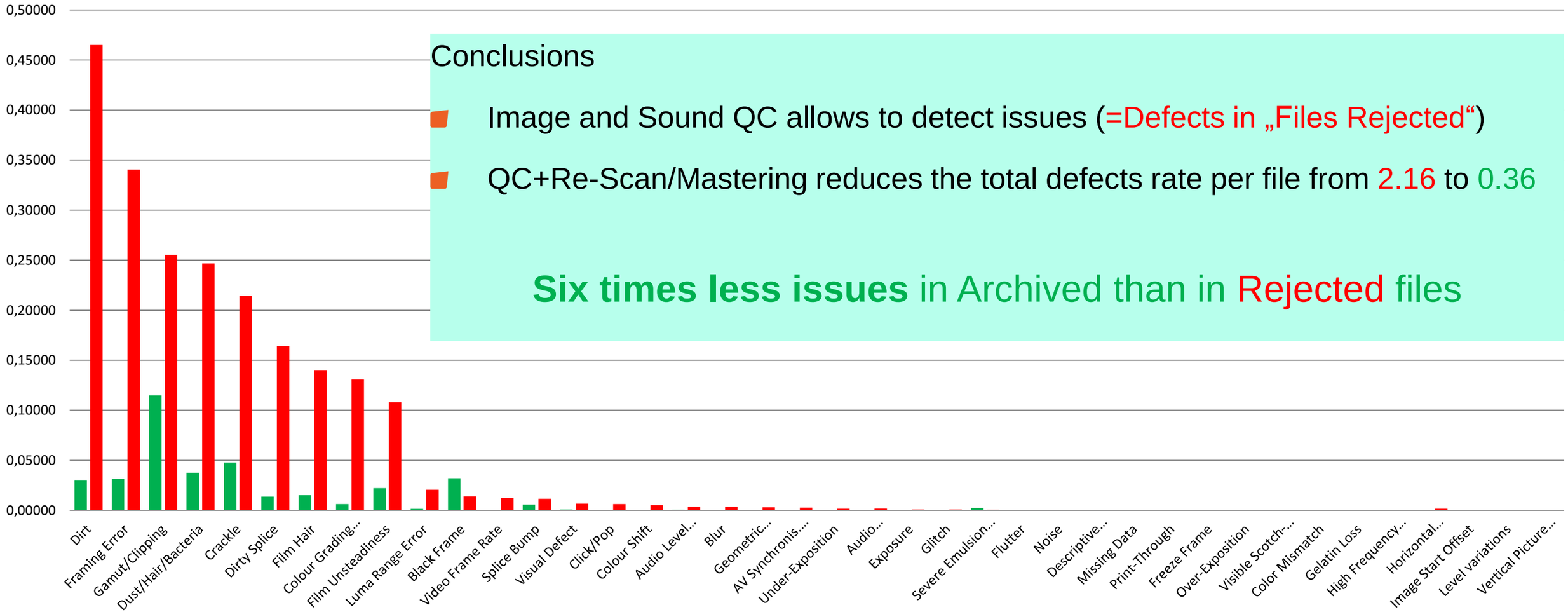
Peter Schallauer, Joanneum Research

Image and Sound QC
How effective is it in operations?

- Defect Statistics
  - Dec. 2017 – Sept. 2020
  - 17595 titles/reels
  - 35194 files / 9968 hours QC‘d (overscanned and cropped)
  - Avg. File/reel duration: 16 min. 59 sec.
  - Comparison between
    - Files to be Archived
    - Files/reels Rejected

Image and Sound QC
Most Critical Issues & Effectiveness

Avg. Defects Rate per File in Files Rejected vs. Files Archived



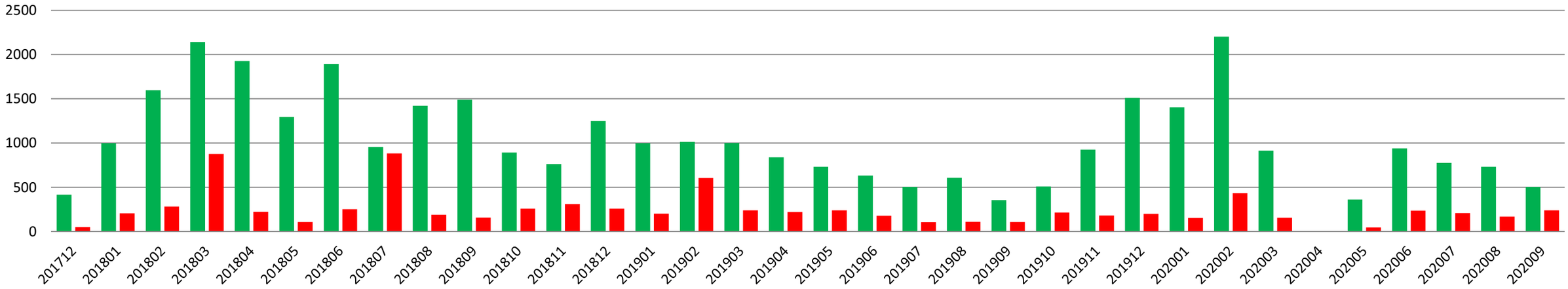
Conclusions

- Image and Sound QC allows to detect issues (=Defects in „Files Rejected“)
- QC+Re-Scan/Mastering reduces the total defects rate per file from 2.16 to 0.36

Six times less issues in Archived than in Rejected files

MDPI Film - Operations Overview

#Archived and #Rejected Files per Month



- Goal of QC is to detect systematic scanning equipment issues and individual film scanning/mastering issues
- IU requests some rescans for reasons other than scanning/mastering issues
- Requirements and complexity change for some collections and film stocks
- Scanner and film cleaner equipment issues led to spikes in March and July 2018 and in February 2019 -> were detected and solved
- Consequences of Covid19 in April and May 2020

Conclusions

- Approach to QC is a project strength
  - Allows IU to better understand its diverse collections and adapt workflows
  - Enables high quality archive package
- Submission Package QC
  - Ensures packages meets archiving standards
- VidiCert Image and Sound QC
  - Integrates very well with IU workflow and Submission Package QC
  - Image and audio issues can be detected quickly
  - Detailed automatic and interactive detection functions helps in finding the origin of an issue
  - QC + Re-Scan reduces the total defects rate by more than a factor of 6
- QC strengthens IU‘s relationship with the service provider
- Very beneficial cooperation between IU and JR

Thank you!

Darrell Myers
dsmyers@iu.edu

Peter Schallauer
peter.schallauer@joanneum.at

mdpi.iu.edu

JOANNEUM RESEARCH DIGITAL

VidiCert

www.vidicert.com