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

Peter Schallauer

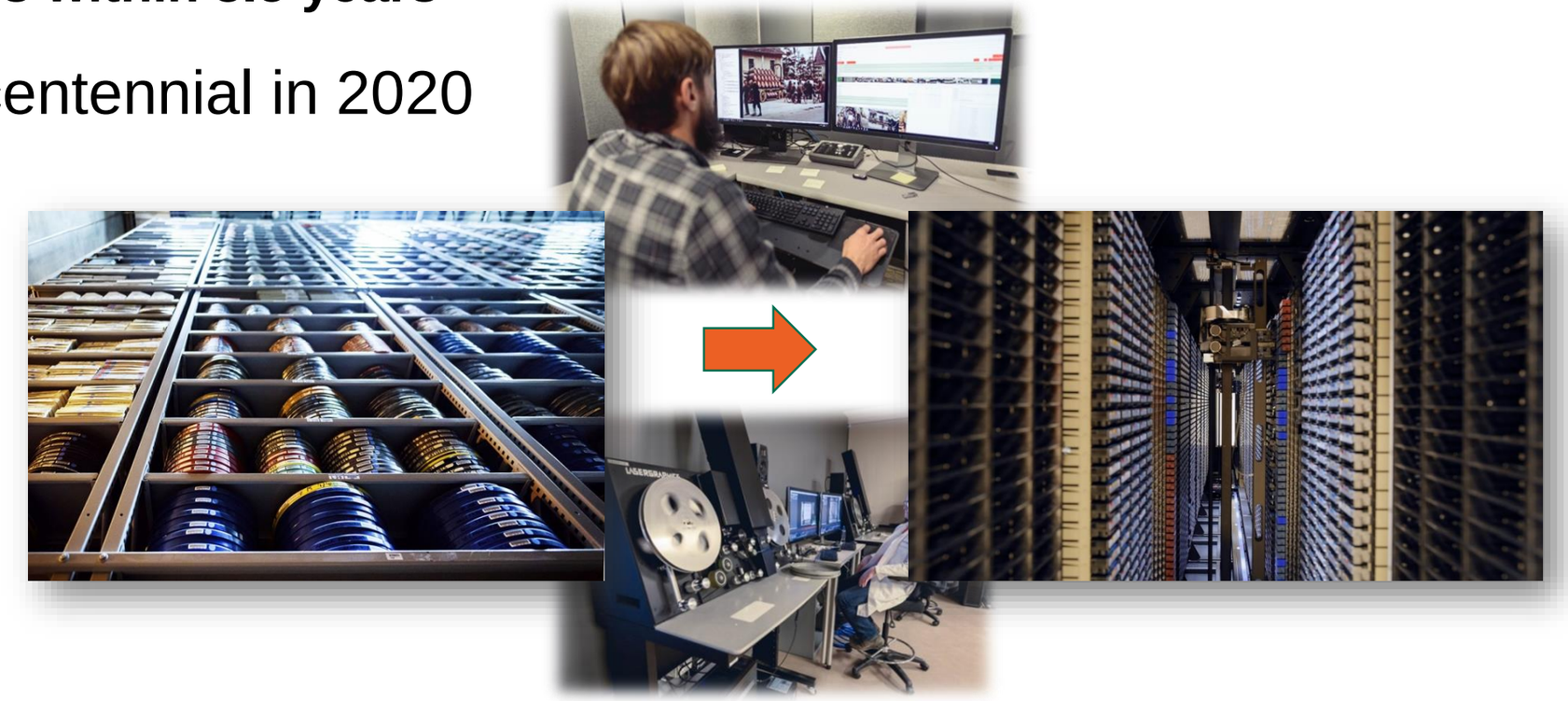
JOANNEUM RESEARCH DIGITAL

Darrell Myers

Joint Technical Symposium
Oct. 4th 2019, Hilversum

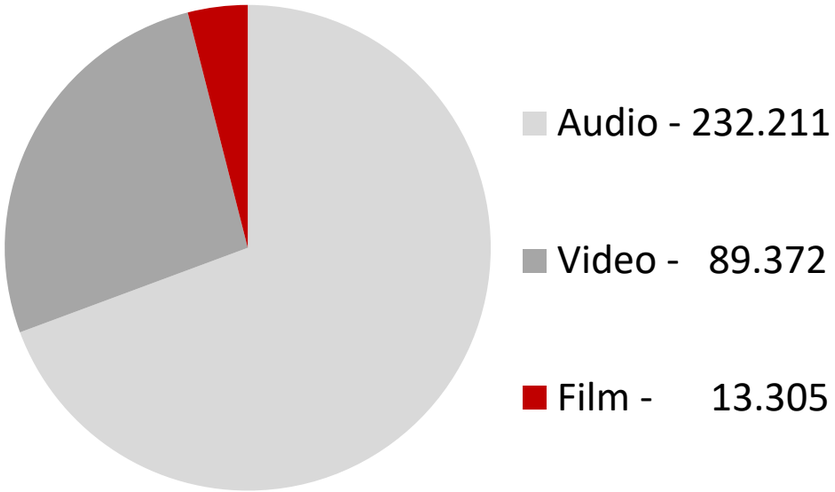
Media Digitization & Preservation Initiative

- Digitally preserve all significant audio, video, and film (~24 Petabytes)
  - Audio/Video: ~325,000 recordings within 5 years
  - **Film: ~30,000 reels within 3.5 years**
- Complete by IU Bicentennial in 2020



Preserved Titles
30 September 2019

Number of Items Digitized

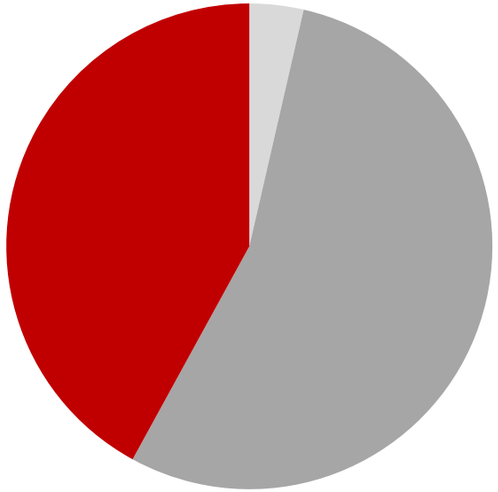


Audio - 232.211

Video - 89.372

Film - 13.305

Storage Usage (12.3 PB)



Audio - 436 TB

Video - 6,6 PB

Film - 5,1 PB

Hours

Audio – 175.460

Video – 108.153

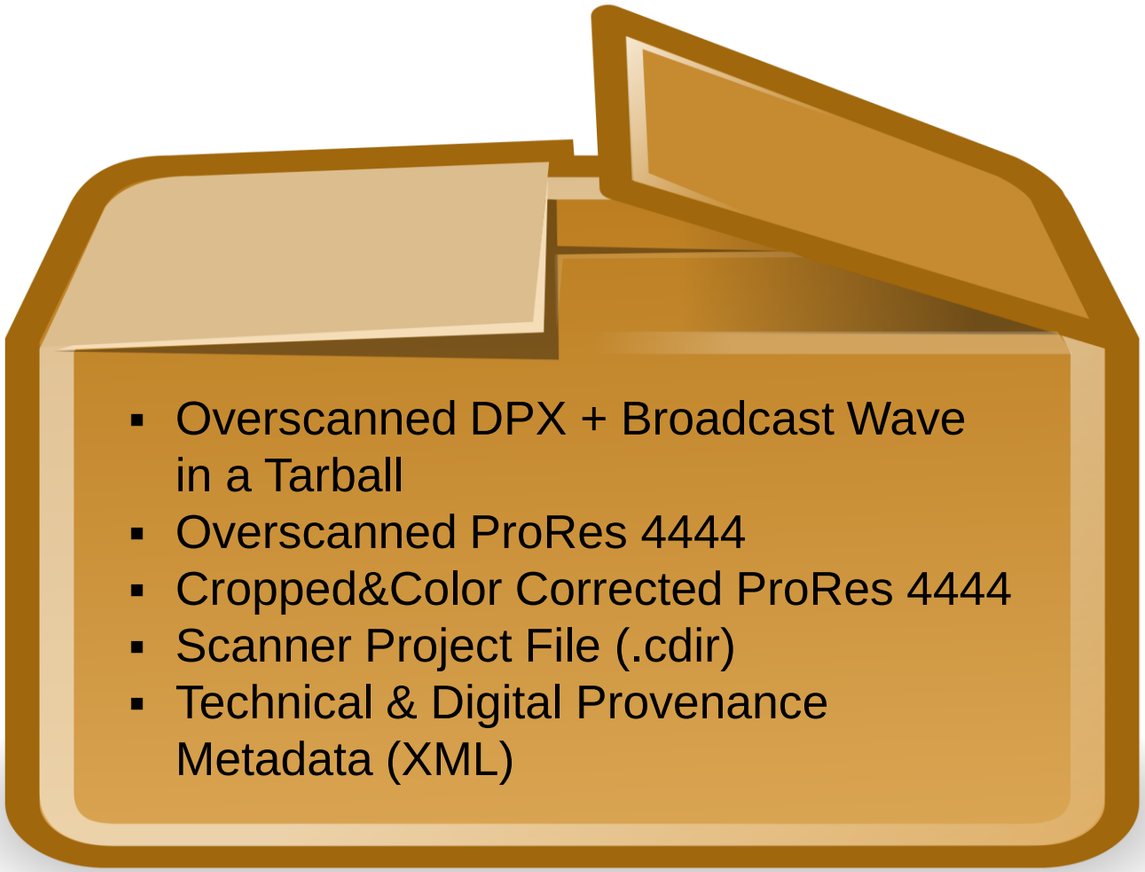
Film – 3.747

Preserved Titles Detail
30 September 2019

Type	Count	SDA Usage(G)	Avg Size(G)	Duration(h)	Avg Duration(m)
1-Inch Open Reel Video Tape	4054	146734.71	36.20	2509.03	37.13
1/2-Inch Open Reel Video Tape	371	9942.52	26.80	223.86	36.20
45	4032	2084.07	0.52	513.77	7.65
78	34449	16553.92	0.48	3684.83	6.42
8mm Video	992	78538.10	79.17	1347.94	81.53
Aluminum Disc	1250	822.15	0.66	182.36	8.75
Audiocassette	52808	254589.29	4.82	62641.12	71.17
Betacam	19086	832001.93	43.59	13691.40	43.04
Betamax	1224	105468.27	86.17	1954.16	95.79
CD-R	11928	14149.66	1.19	10278.29	51.70
Cylinder	6643	2031.77	0.31	322.43	2.91
DAT	8996	24929.38	2.77	17671.37	117.86
DVD	2919	20248.62	6.94	3061.27	62.92
Film	13305	5066499.65	380.80	3746.61	16.90
Lacquer Disc	5776	6397.61	1.11	1496.22	15.54
LP	38657	114740.65	2.97	28130.52	43.66
Open Reel Audio Tape	67654	141319.73	2.09	50534.83	44.82
Other Analog Sound Disc	18	16.71	0.93	4.12	13.74
U-matic	15223	540272.96	35.49	9295.91	36.64
VHS	45503	4909617.15	107.90	76069.12	100.30
All Types	334888	12286958.84	36.69	287359.16	51.48

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

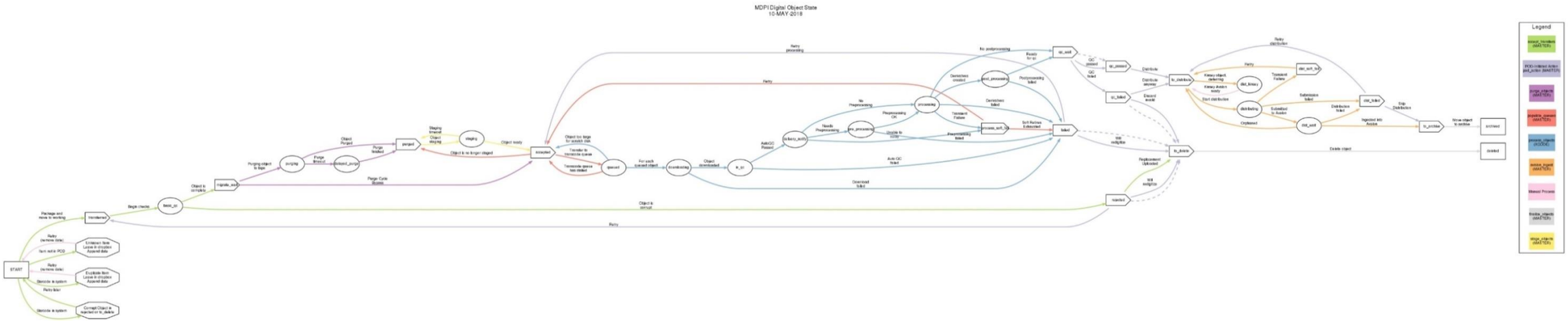


Image and Sound QC - VidiCert

- 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

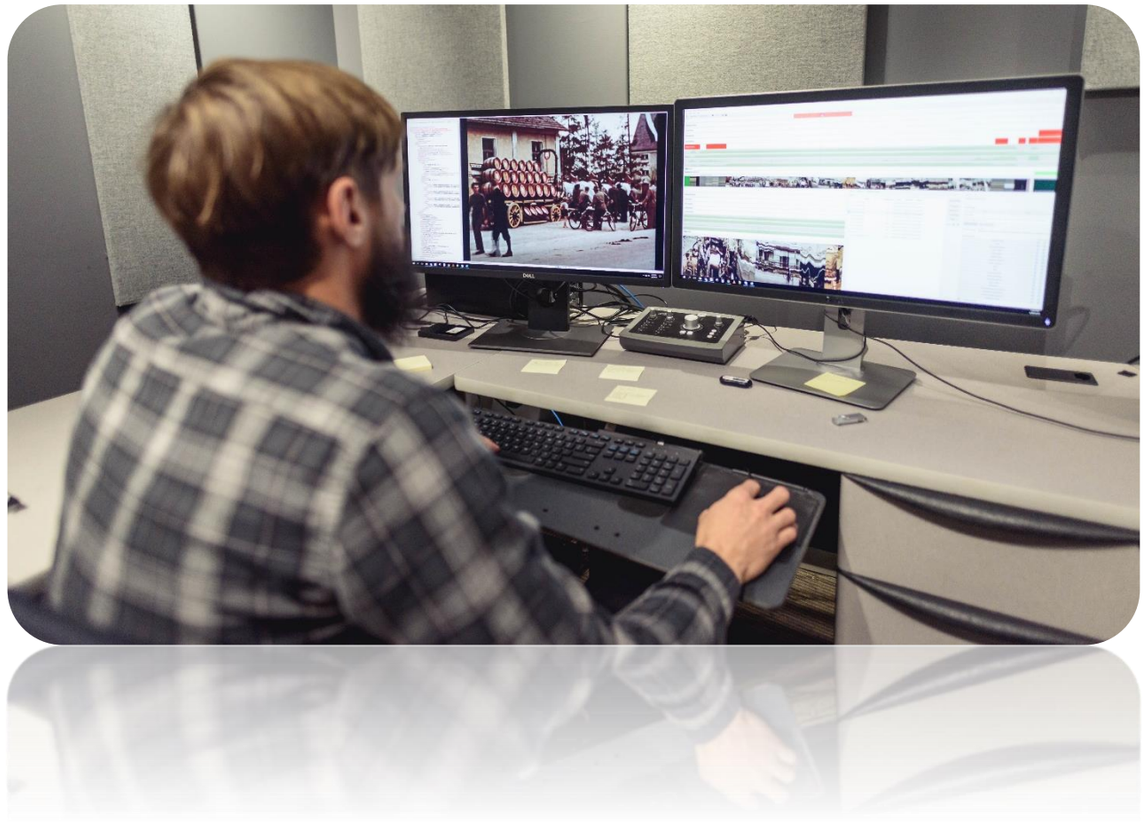


Image and Sound QC Identified Defects

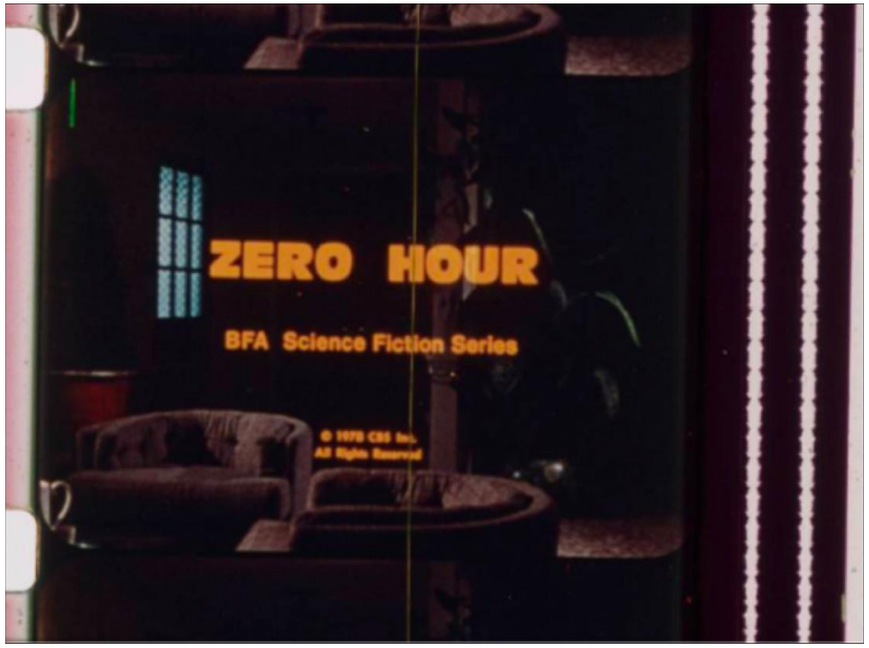
Dirt / Dust

Gamut/Clipping

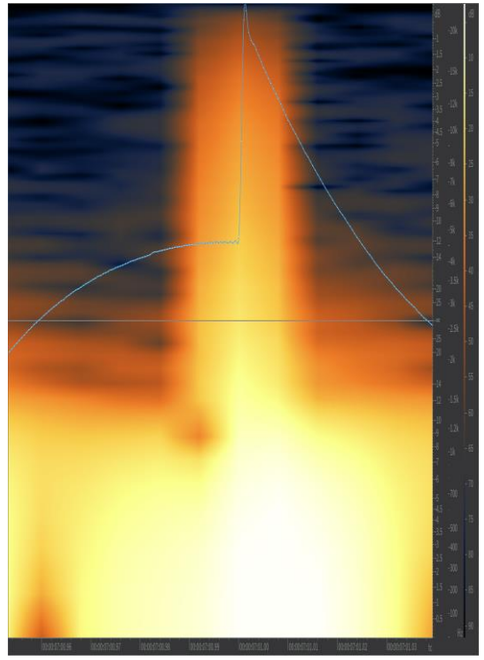
Foreign Object in Frame

Color Correction Error

Image and Sound QC Identified Defects



Crookedness



Interstitial Errors

Added Tones

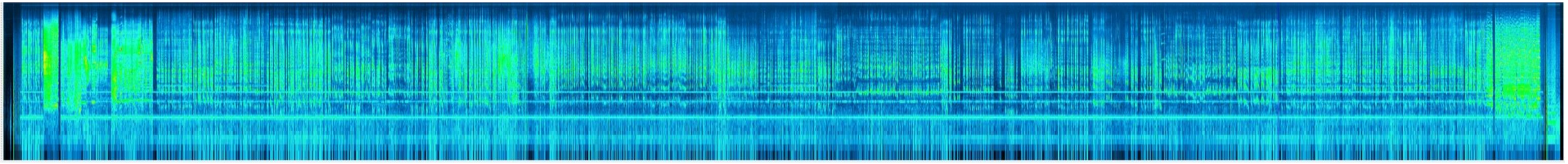
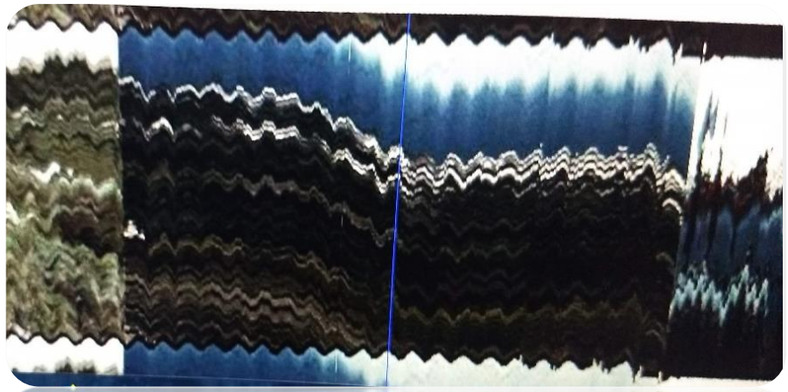


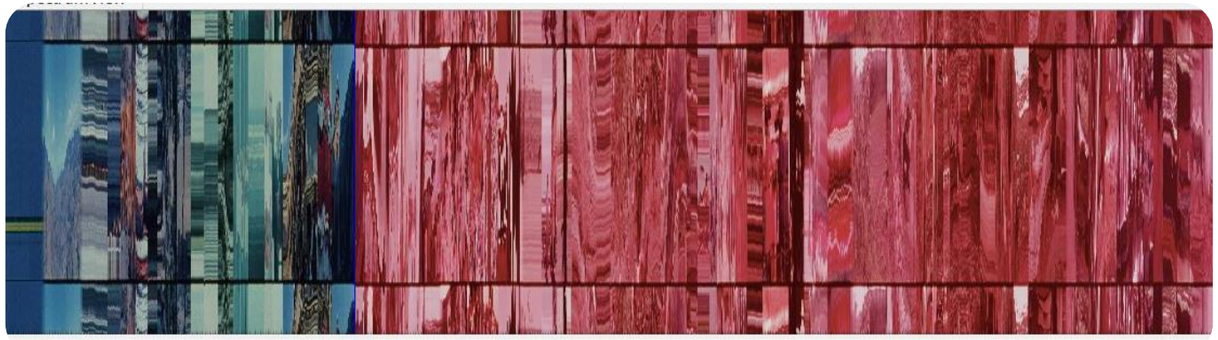
Image and Sound QC
Identified Defects & Reporting

Framing Errors & Image Loss



Unsteadiness

PDF
Reports



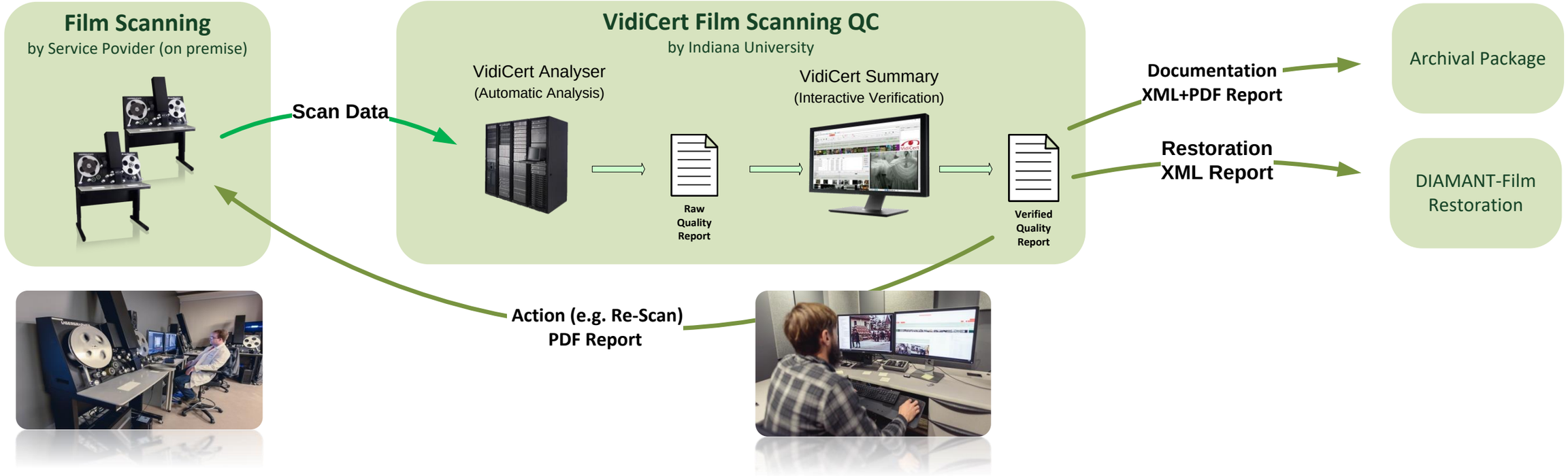
Faded Stock – Color Correction Needed

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

Image and Sound QC
Automated Defect 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

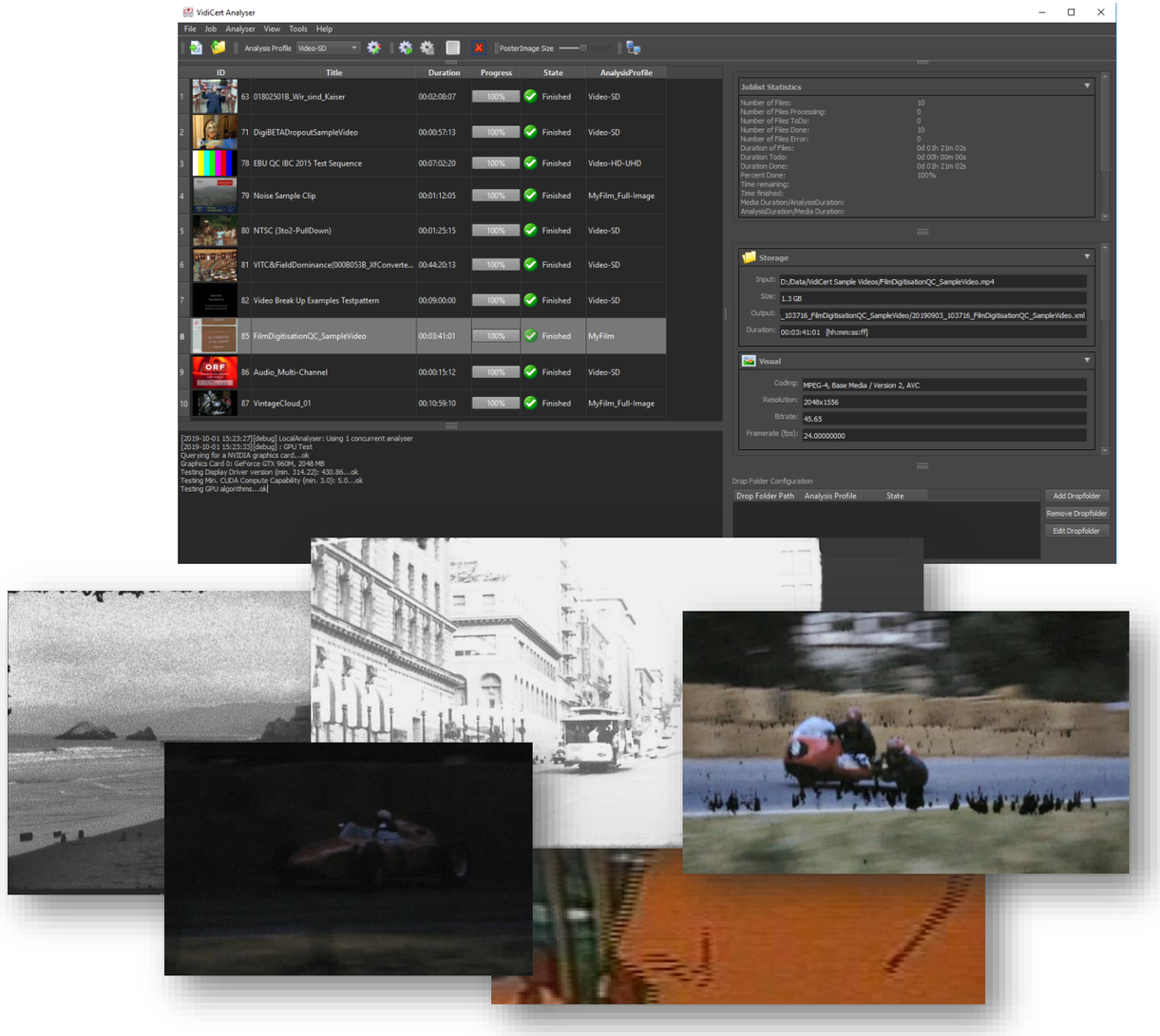


Image and Sound QC
Interactive Defect Verification

- Time-efficient human verification of automated detections
  - Advanced summarization and navigation by timeline based views for each defect and quality measure
  - Full featured desktop player: zoom, defect and graphics overlays, time accurate playback, second monitor full screen, frame rate check, selectable audio channels, SDI out
  - Integration of external time based metadata, e.g. film scanner sensor metadata
  - Fully customizable user interface (presets for different QC tasks)
  - Efficient time-based human annotation
  - Job-time optimistation capability – trade-off human effort against verification accuracy
- QC Reports
  - Machine readable XML
  - Human readable pdf
- DIAMANT Restoration Report

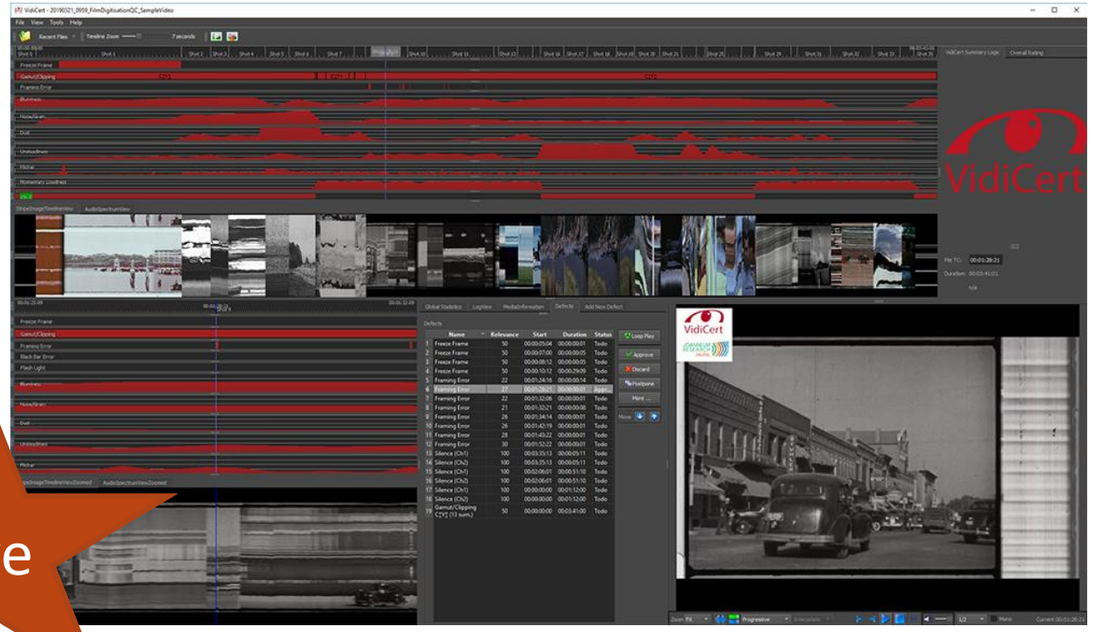
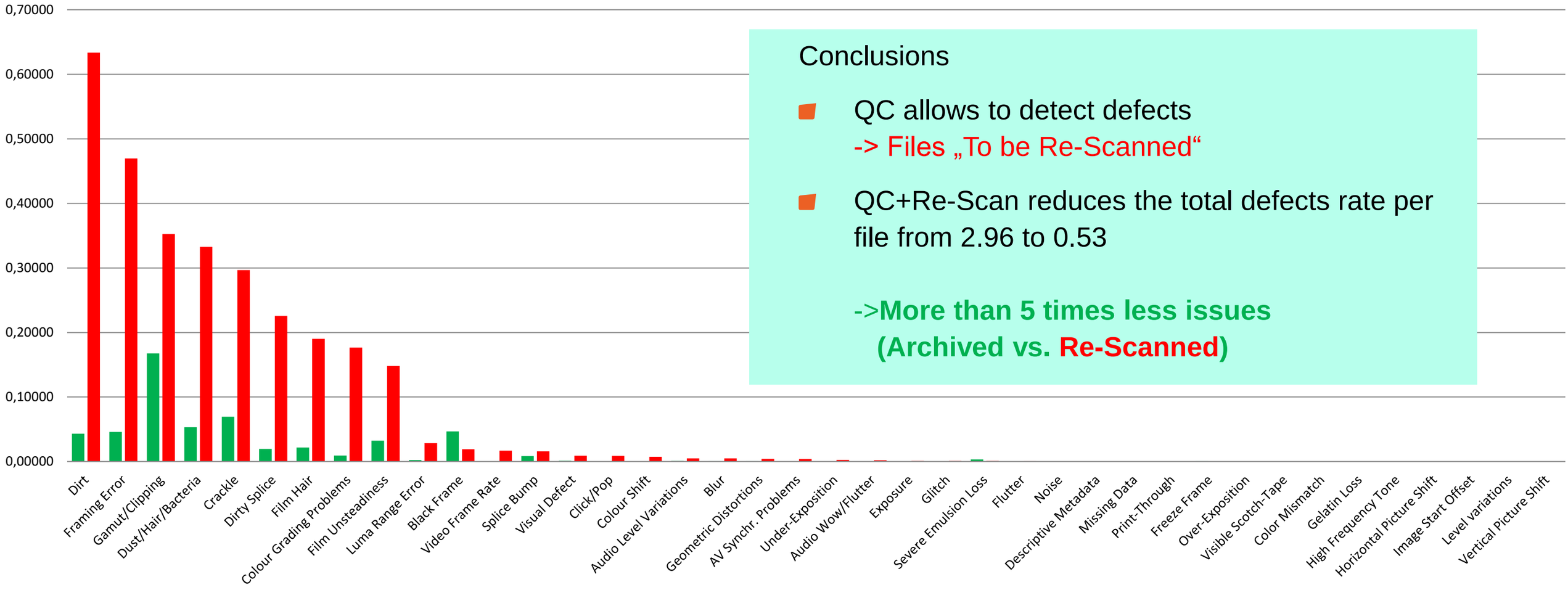


Image and Sound QC
How effective is it in operations?

- Defect Statistics
  - Dec. 2017 – Aug. 2019
  - 12073 titles/reels
  - 24146 files / 6526 hours QC‘d (overscanned and cropped)
  - Avg. File/reel duration: 16 min. 13 sec.
  - Comparison between
    - Files to be Archived
    - Files/reels to be Re-scanned

Image and Sound QC
Most Critical Issues & Effectiveness

Defects per File in Files to be Re-Scanned vs. Files to be Archived

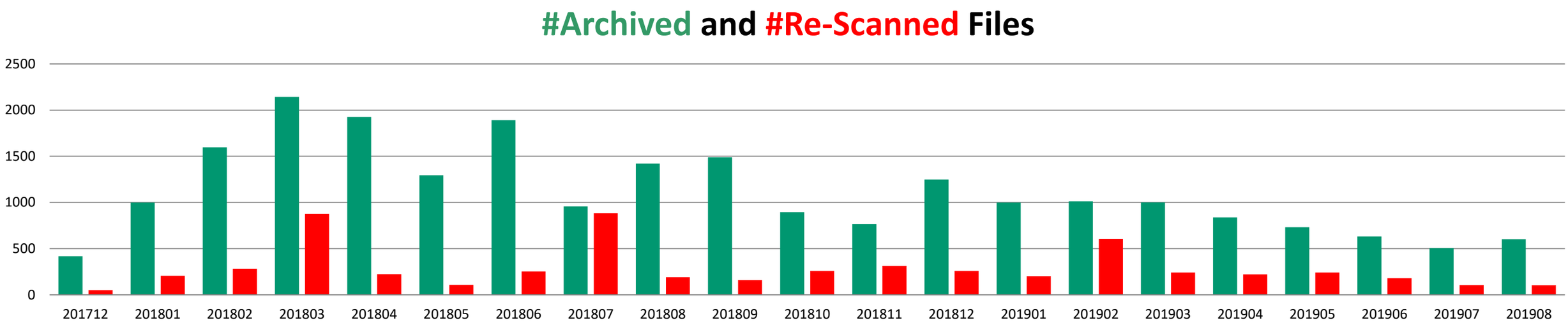


Conclusions

- QC allows to detect defects
  -> Files „To be Re-Scanned“
- QC+Re-Scan reduces the total defects rate per file from 2.96 to 0.53

->**More than 5 times less issues (Archived vs. Re-Scanned)**

MDPI Film - Operations Overview
Re-Scans per Month



- Goal of QC is to detect defects in the short term and to drive down rescans in the mid term
- IU requests some rescans for reasons other than scanning defects
- 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

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 5
- 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

www.vidicert.com