FRAME tECH 2021

An overview of an AV archive, from
physical assets management to digitization
and digital objects preservation.

17 → 19 & 22 → 24 / 11

FRAME

Co-funded by the
Creative Europe Programme
of the European Union

EBU ACADEMY

FIAT IFTA

ina

Open Source Quality Control and Quality Management

MediaArea

Open source team company focused on digital media analysis.

Why quality control is important?

Why quality control is important?

If you don’t handle that now, content is definitely partially lost.





Why quality control is important?

Your digital copy provider is available now
In the future, your content provider may have disappeared, trashed content…

The original tape is still there, but for how long?
The original tape is still there, but for how long?
Or your tape becomes no more readable.

Why quality control is important?

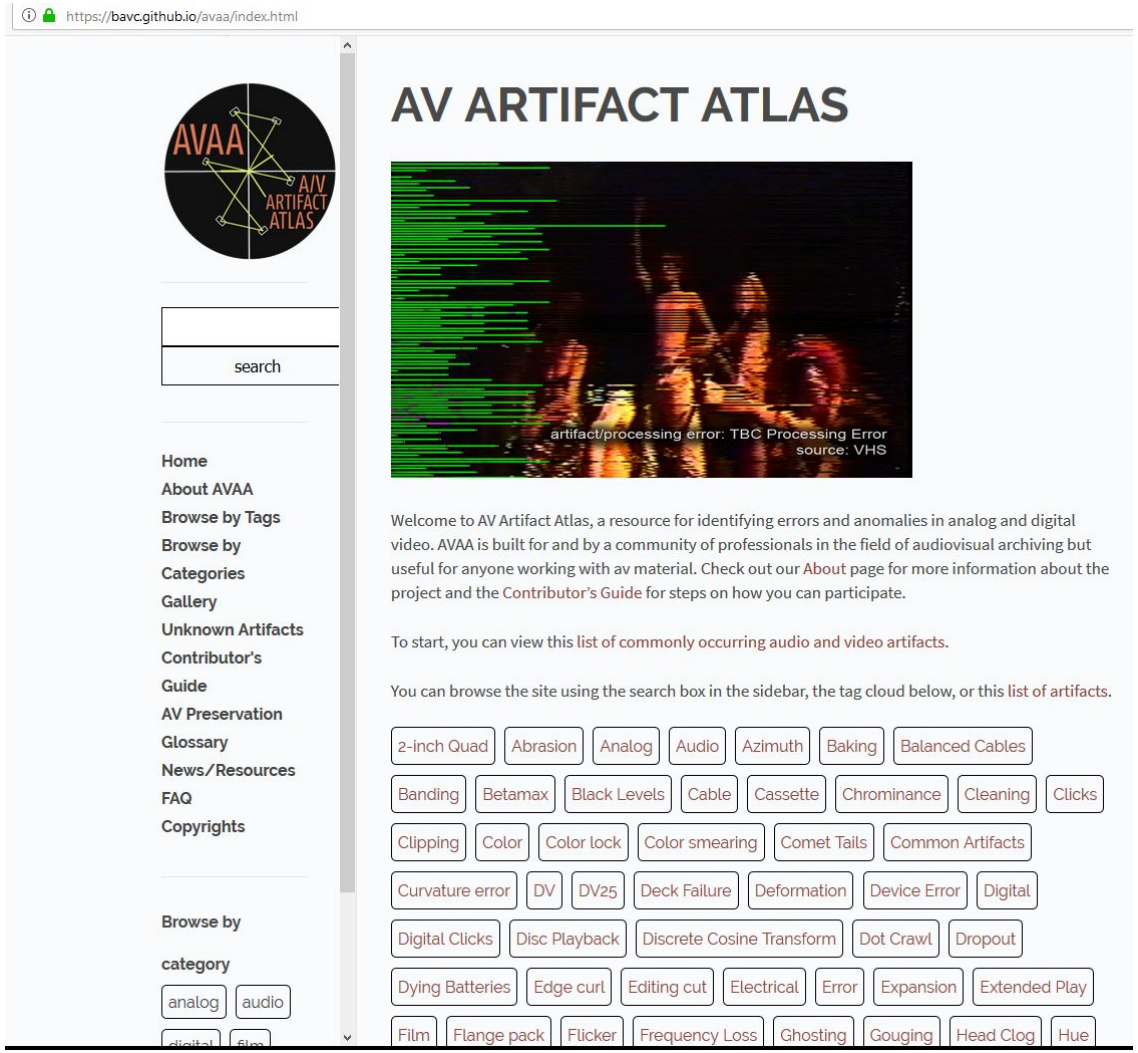
Easier to catch quality issues and fix them now than later
When someone else discovers quality issues, it may be too late.

Why quality control is important?

You are not alone

Example of online support:
AV Artifact Atlas, a resource for identifying errors and anomalies in analog and digital video. AVAA is built for and by a community of professionals in the field of audiovisual archiving but useful for anyone working with av material.

https://www.avartifactatlas.com/



Why quality control is important?

You are not alone

Example of symposiums:
“No Time To Wait” symposiums are free events focused on open media, open standards, and digital audiovisual preservation.
Symposium program is made by archivists, who present how they work and the associated (open source) tools.

Organization of symposiums has a cost.
- Hosting granted by a main sponsor
- Other archives or vendors may provide additional financial support

“No Time To Wait” 2021 will be at Sound & Vision in Amsterdam on December 8-10.
Other sponsors are AV Preservation by reto.ch, Centre national de l’audiovisuel Luxembourg (CNA), The National Library of Norway and the International Federation of Film Archives (FIAT/IFTA)

FRAME

Overview of open source

Overview of open source

The archival domain:

- Relatively small market (niche)
- Off-the-shelf products often focus on different use cases
- Highly specialized demands
- High potential for vendor dependence...

Overview of open source

Proprietary tools and formats are linked to one, and only one, vendor

- If the vendor disappears, it is impossible to get bug fixes
- Without source code, you can not adapt yourself to new platforms: are you sure you'll be able to run the software in 100 years?
- Vendor decides on the new features
- Are you big enough for making the vendor develop your prefered feature?

Overview of open source

The four freedoms of open source (a.k.a. free software)

The freedom to run the program as you wish, for any purpose
The freedom to study how the program works, and change it so it does your computing as you wish
The freedom to redistribute copies so you can help your neighbor
The freedom to distribute copies of your modified versions to others

Overview of open source

Reusability of existing solutions

- Free choice of (local) support/suppliers
- No black-box
- Common tools/codebase = larger userbase
- Less "forced" upgrades

Overview of open source

Benefits of paying for open source software:

- Better support/updates
- Pooling resources
- Designed for your use-cases
- Overall better cost-effectiveness
- Public money = public solution?

Overview of open source

Open source does not mean lack of professional support:

- Support contracts
- Paid installation/integration
- Hire developers

Non-financial contributions:

- Documentation
- Testing / bug-reports
- Helping others

Overview of open source

A real-world example: The lossless video codec FFV1 (RFC 9043)

- Austrian National Archive (Mediathek) wanted to do lossless digital video archiving
- Not satisfied with existing products (Interoperability issues)
- Found FFV1 in FFmpeg: Excellent codec, but we wanted/needed more...
- Contacted and hired a FFV1 developer
- Other parties involved for advices (pooled resources)
- Budget calculated in reference to costs of proprietary alternatives
- Now FFV1 version 3: faster and integrity-aware
- Important: we published our experiences with FFV1
- So: other archives using FFV1 now profit from improvements, too!

Overview of open source

Quality of tools

Chicken and egg: if everyone waits, nothing happens
We start with needs easy to handle and we create dedicated open source tools
Tools become bigger, step by step, when more people join

YOU can decide about participating in lowering the overall cost for archives.

Overview of open source

Quality of tools

Chicken and egg: if everyone waits, nothing happens
We start with needs easy to handle and we create dedicated open source tools
Tools become bigger, step by step, when more people join

YOU can decide about participating in lowering the overall cost for archives.

Overview of open source

Quality of tools

Chicken and egg: if everyone waits, nothing happens
We start with needs easy to handle and we create dedicated open source tools
Tools become bigger, step by step, when more people join

YOU can decide about participating in lowering the overall cost for archives.

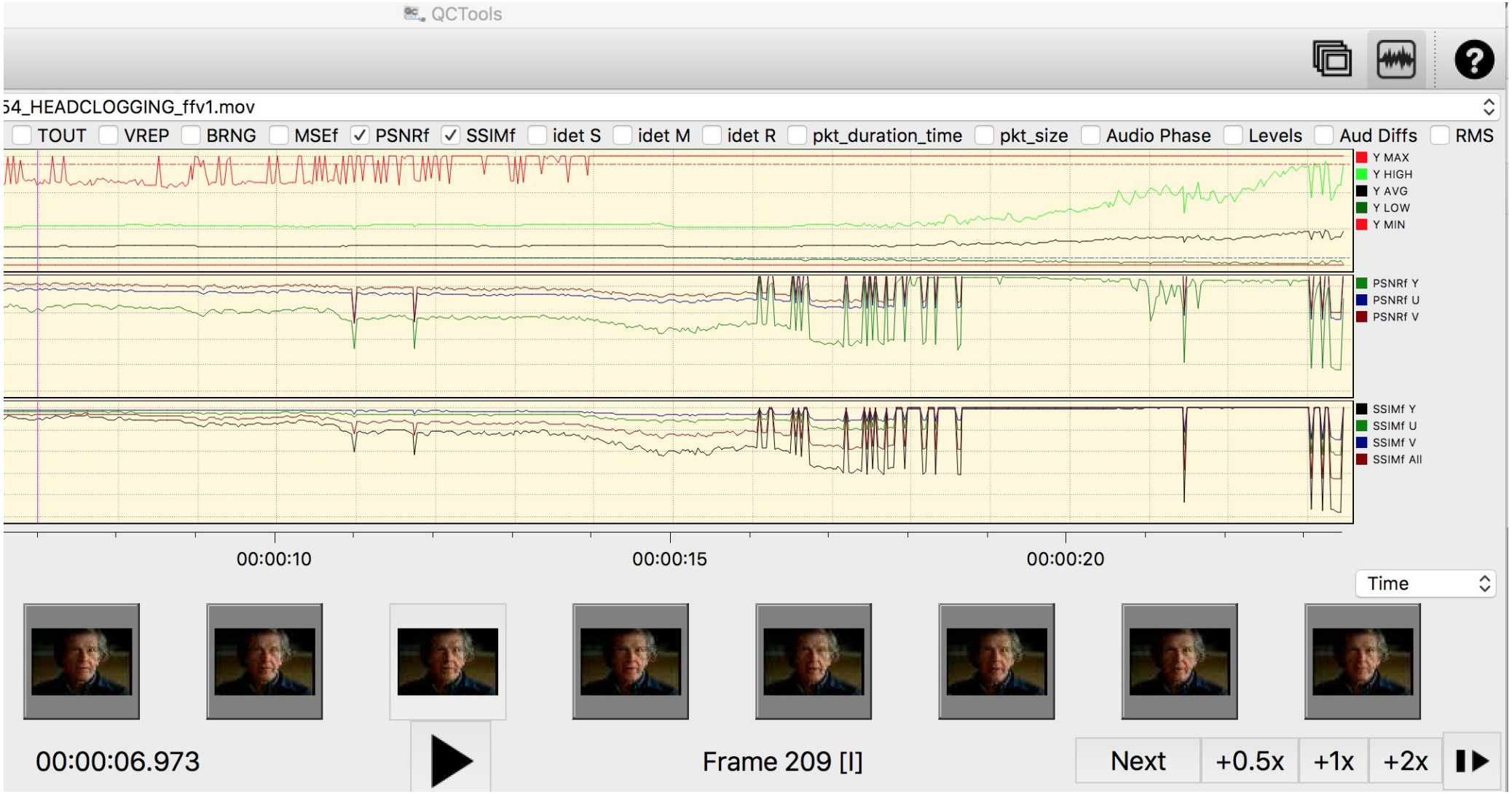
QCTools

QCTools

Tool that helps users analyze and understand their digitized video files through use of audiovisual analytics and filtering.



QCTools


QCTools
54_HEADCLOGGING_ffv1.mov
TOUT
VREP
BRNG
MSEf
PSNRf
SSIMf
idet S
idet M
idet R
pkt_duration_time
pkt_size
Audio Phase
Levels
Aud Diffs
RMS
Y MAX
Y HIGH
Y AVG
Y LOW
Y MIN
PSNRf Y
PSNRf U
PSNRf V
SSIMf Y
SSIMf U
SSIMf V
SSIMf All
00:00:10
00:00:15
00:00:20
Time
00:00:06.973
Frame 209 [I]
Next
+0.5x
+1x
+2x

QCTools


CSV
/Users/jerome/Programmation/QCTools/Project/QtCreator/293_0049_jitter_ffv1_h264.mp4
Y values
U values
V values
YDiff
YDiffX
UDiff
VDiff
Diffs
TOUT
VREP
BRNG
HEAD
QC Tools - /Users/jerome/Programmation/QCTools/Project/QtCreator/293_0049_jitter_ffv1_h264.mp4
Previous
00:00:00.067
(Frame 2)
Next
Filters >
Previous
00:00:00.000 (Frame 0)
Next
Normal
Field split
Field diff
Field histogram
Chroma split
Interlace
Waveform
Waveform (Field 1 Only)
Waveform (Field 2 Only)
Waveform (field-split)
Broadcast Range Pixels
Broadcast Range (field-split)
Vectorscope
Vectorscope (Field 1 Only)
Vectorscope (Field 2 Only)
Vectorscope (field-split)
Temporal Outlier Pixels
Temporal Outlier Pixels (field-split)

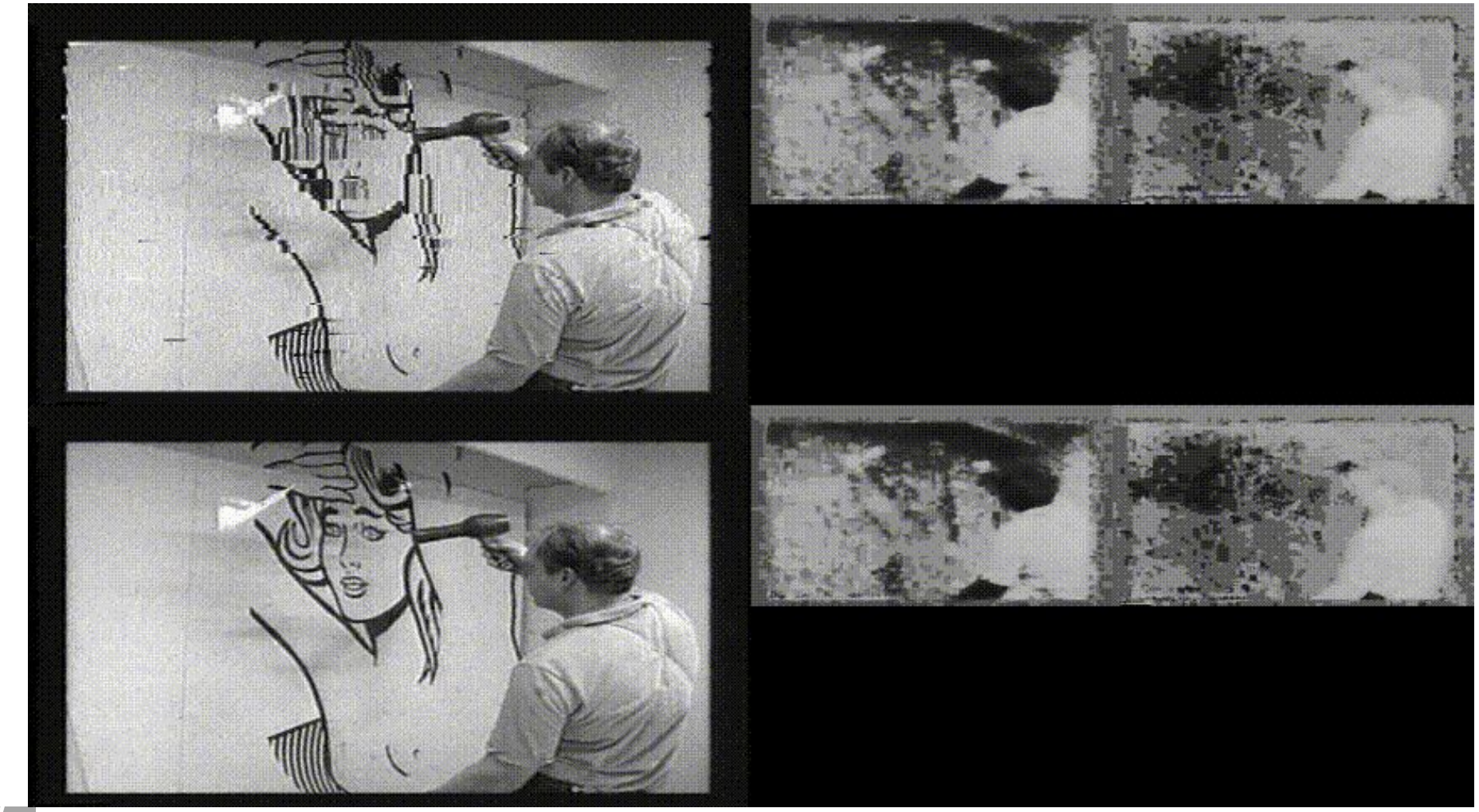
QCTools

Help in detecting issues. Example with field 2 missing color.



QCTools

Only 1 field has issues.

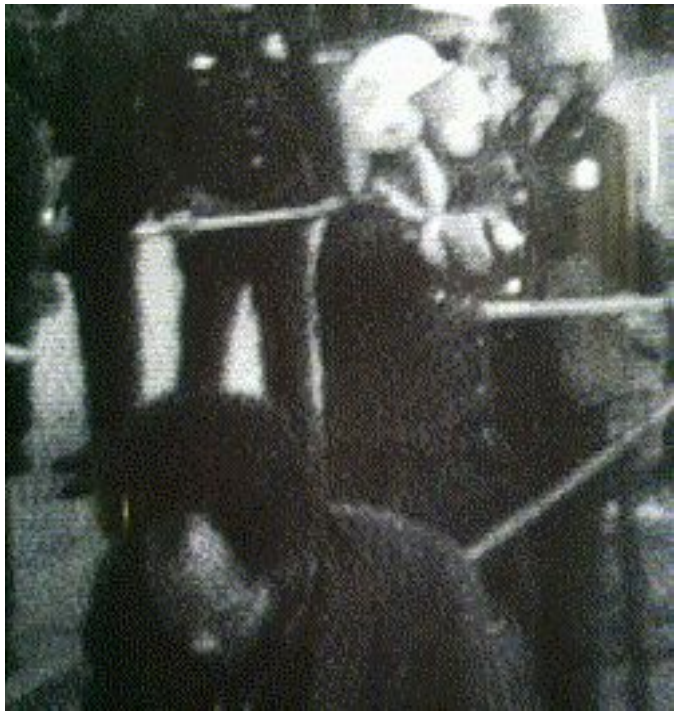


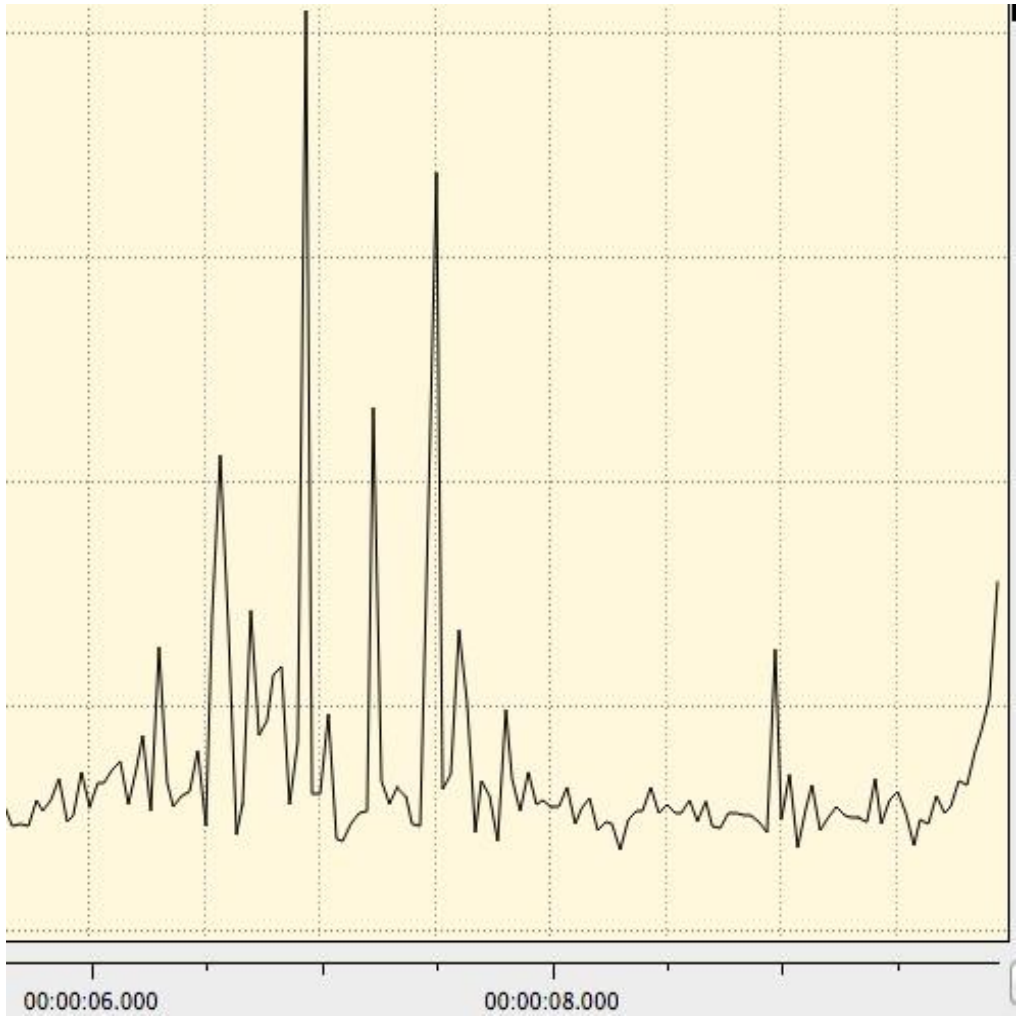
QCTools

Example of drop out.

Dropout is e.g. a brief loss of RF caused by a media defect such as a scratch on the media surface.

Can be detected through “TOUT” graph.



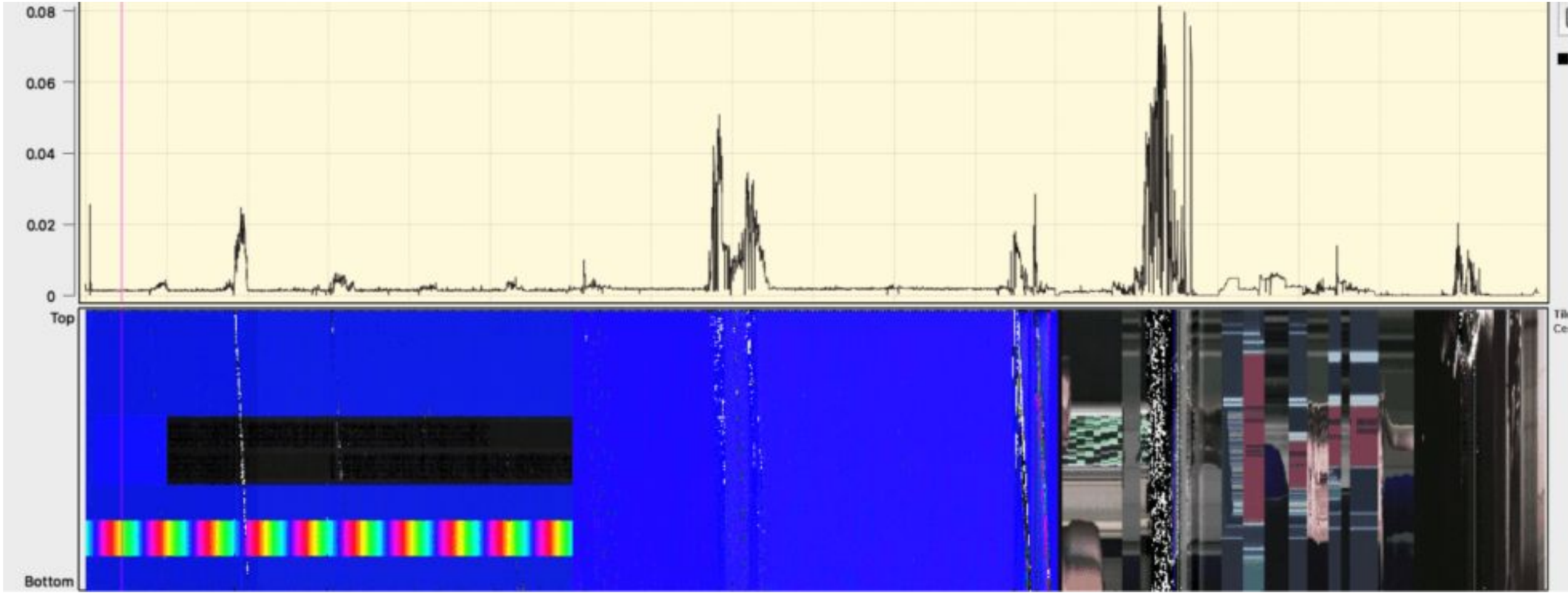


QCTools

Graph Domain	Graph(s) Name	8-bit values	10-bit values	Ideal Range (8/10)
Y Channel	Y MIN, Y LoW, Y AVG, Y HIGH, Y MAX	0-255	0-1023	16/64 (Y Low) 128/512 (Y AVG) 235/940 (Y HIGH)
UV Channels	MIN, LoW, AVG, HIGH, MAX	0-255	0-1023	16/64 (Low) 85/341-170/682 (AVG) 240/960 (HIGH)
Saturation	SAT MIN, SAT LoW, SAT AVG, SAT HIGH, SAT MAX	0-88.7 (75 %) 88.7-118.2 (100%) 118.2-181.02 (Illegal YUV)	0-354.8 (75%) 354.8-472.8 (100%) 472.8-724.08 (Illegal YUV)	75% values
Hue	Hue	0-360°	0-360°	120-147°
Temporal outliers	ToUT	0-1		0-0.009
Vertical Line Repetitions	VREP	0-1		0
Mean Square Error per Field	MSEf	0-1		As close to 0 as possible
Peak Signal to Noise Ratio	PNSR	0-60 dB		30-50 dB

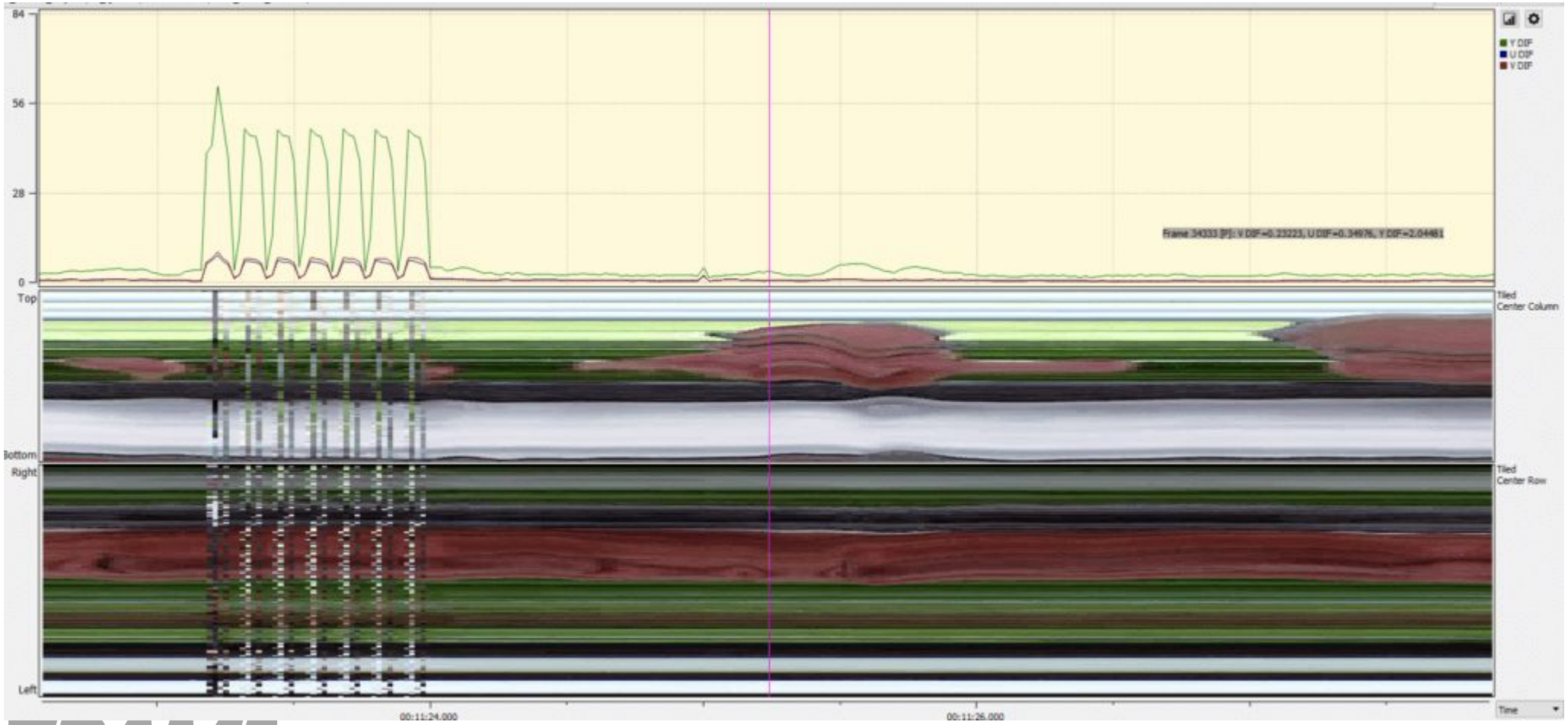
QCTools

In new version 1.2, sponsored by Technische Informationsbibliothek (TIB): panels
This allows a way to provide a visual impression of the content against the traditional QCTools graphs



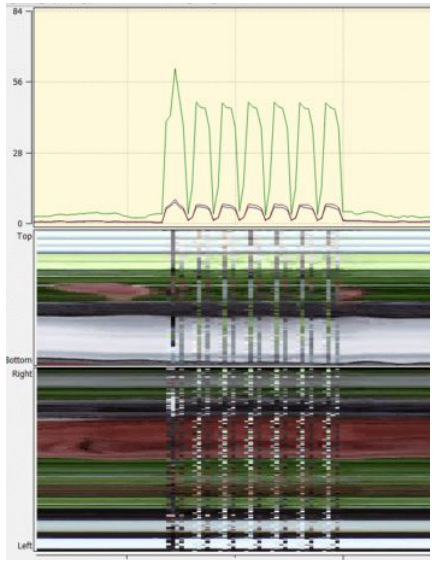
QCTools

How to catch e.g. vertical lines in a digitized digital betacam:
What we have with QCTools



QCTools

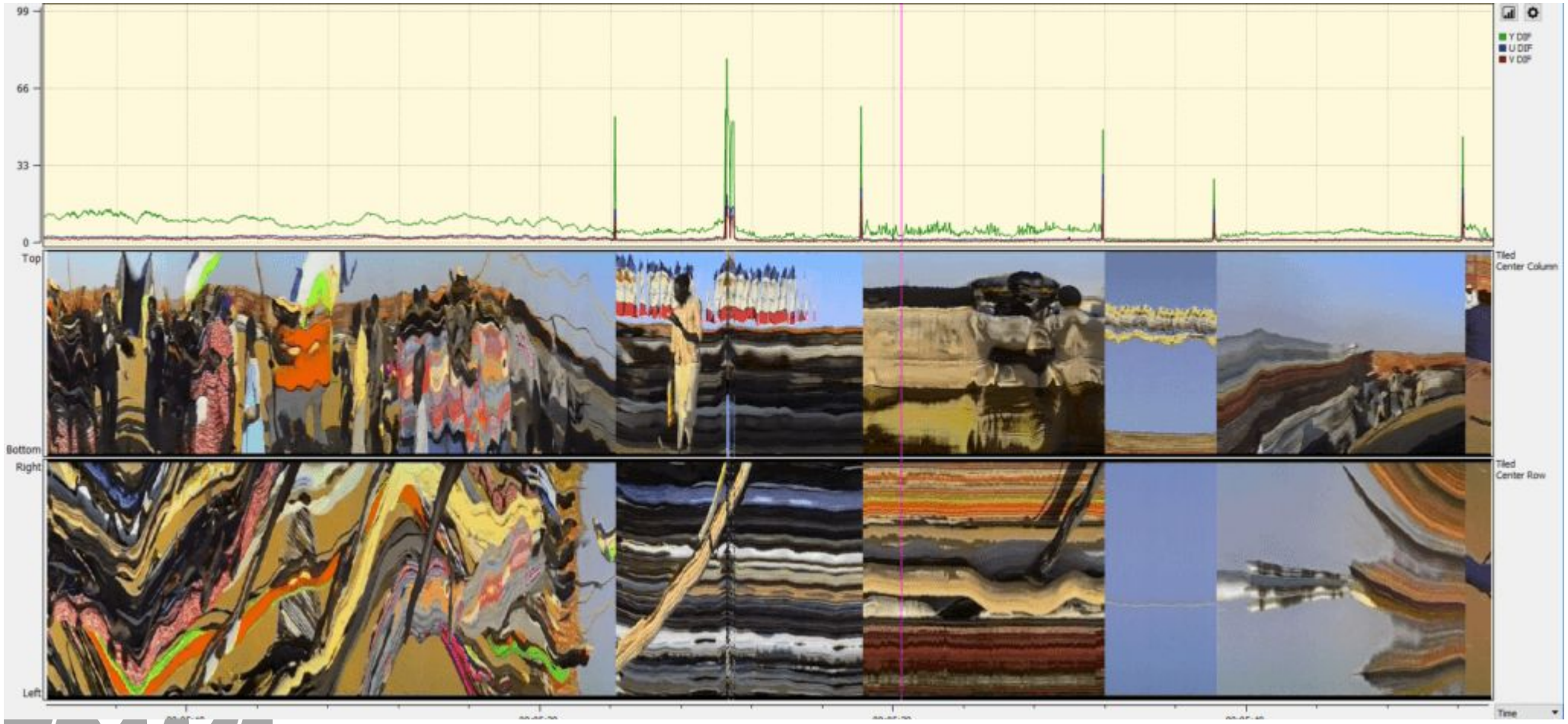
How to catch e.g. vertical lines in a digitized digital betacam:
The scanned image!





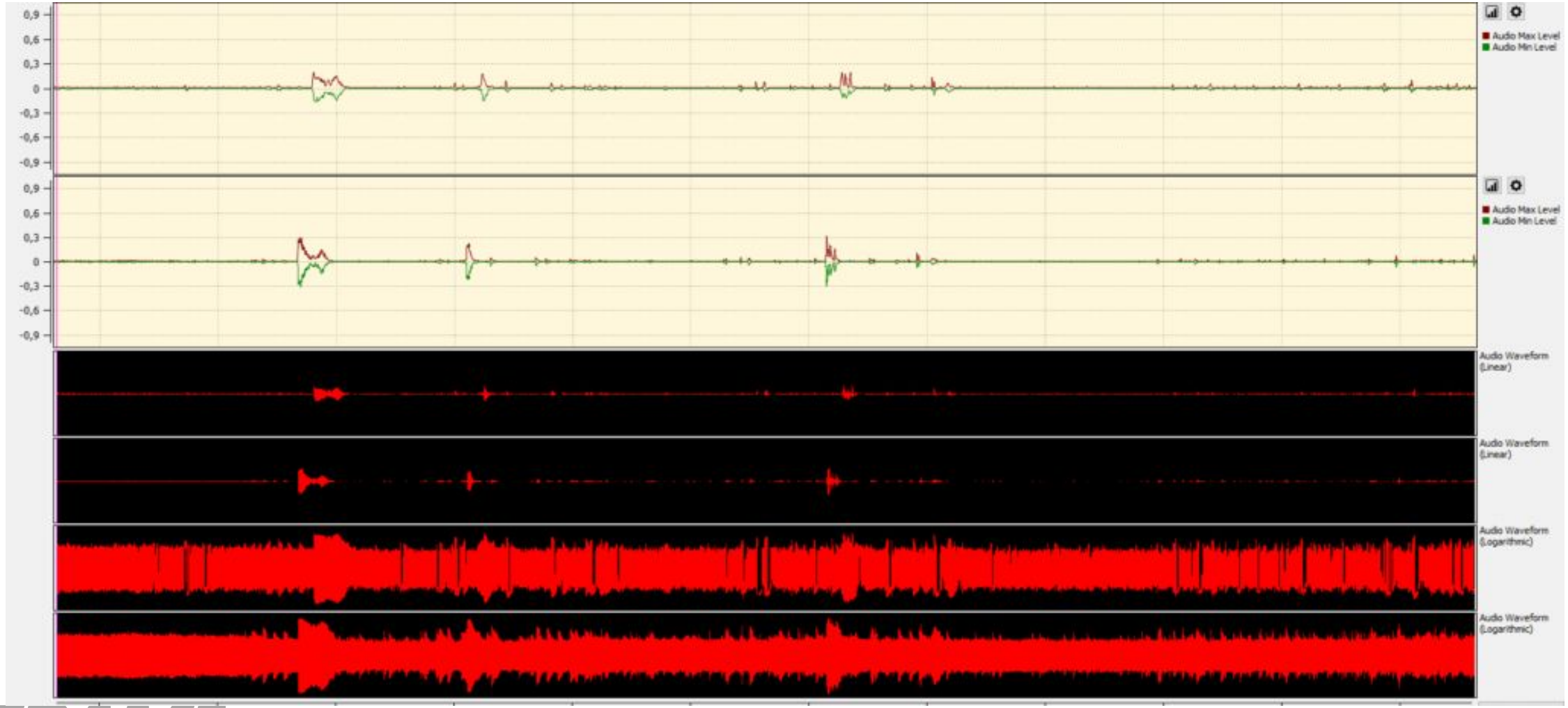
QCTools

Big differences in luminance/colors may be normal (scene change), but sometimes it is not, panels helps to quickly see if there is a scene change



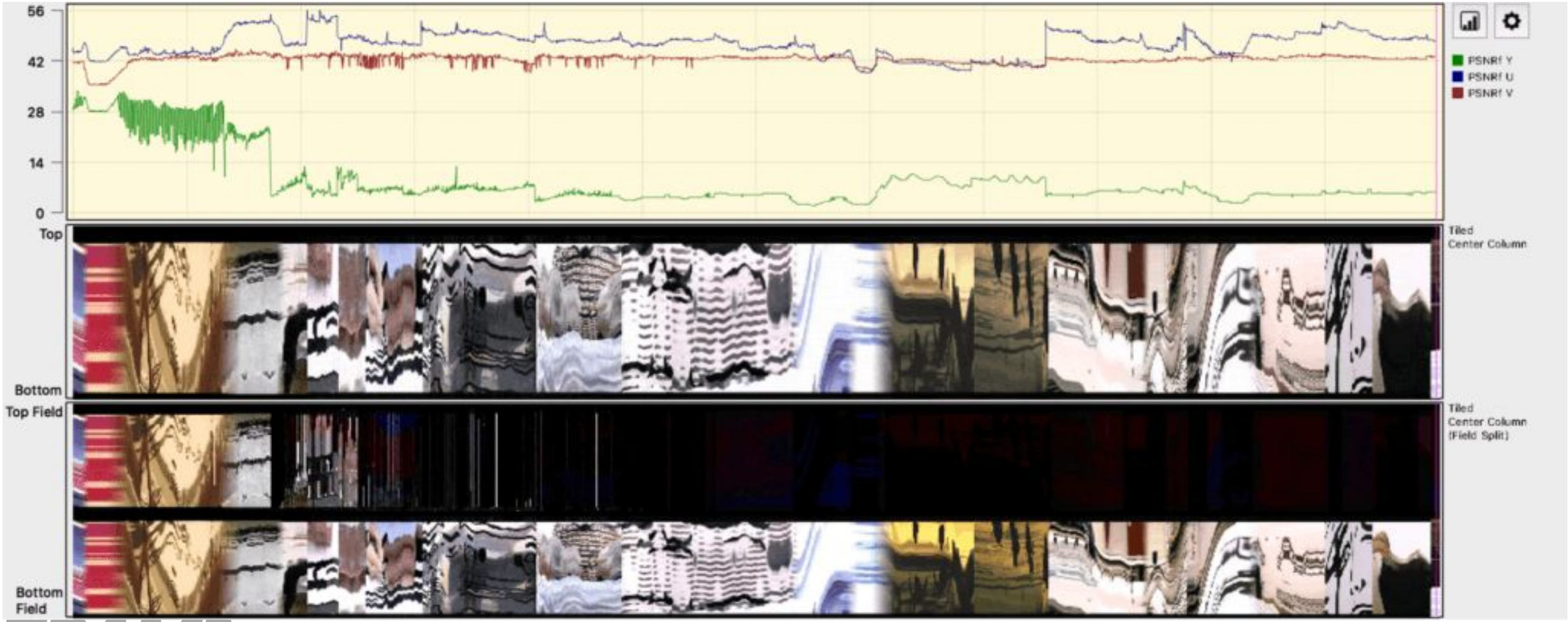
QCTools

Catching audio alignment when the soundtrack is in multiple forms e.g. optic and magnetic.



QCTools

Field split issue detection was manual in the previous versions of QCTools. Now more automated.



QCTools

And more automation with a command line tool (qcli) for preprocessing files.

MediaConch

MediaConch

MediaConch is a conformance checker

- Implementation checker
- Policy checker
- Reporter
- Fixer



MediaConch

How it was developed:

Initial development sponsored by European Union (PREFORMA project)

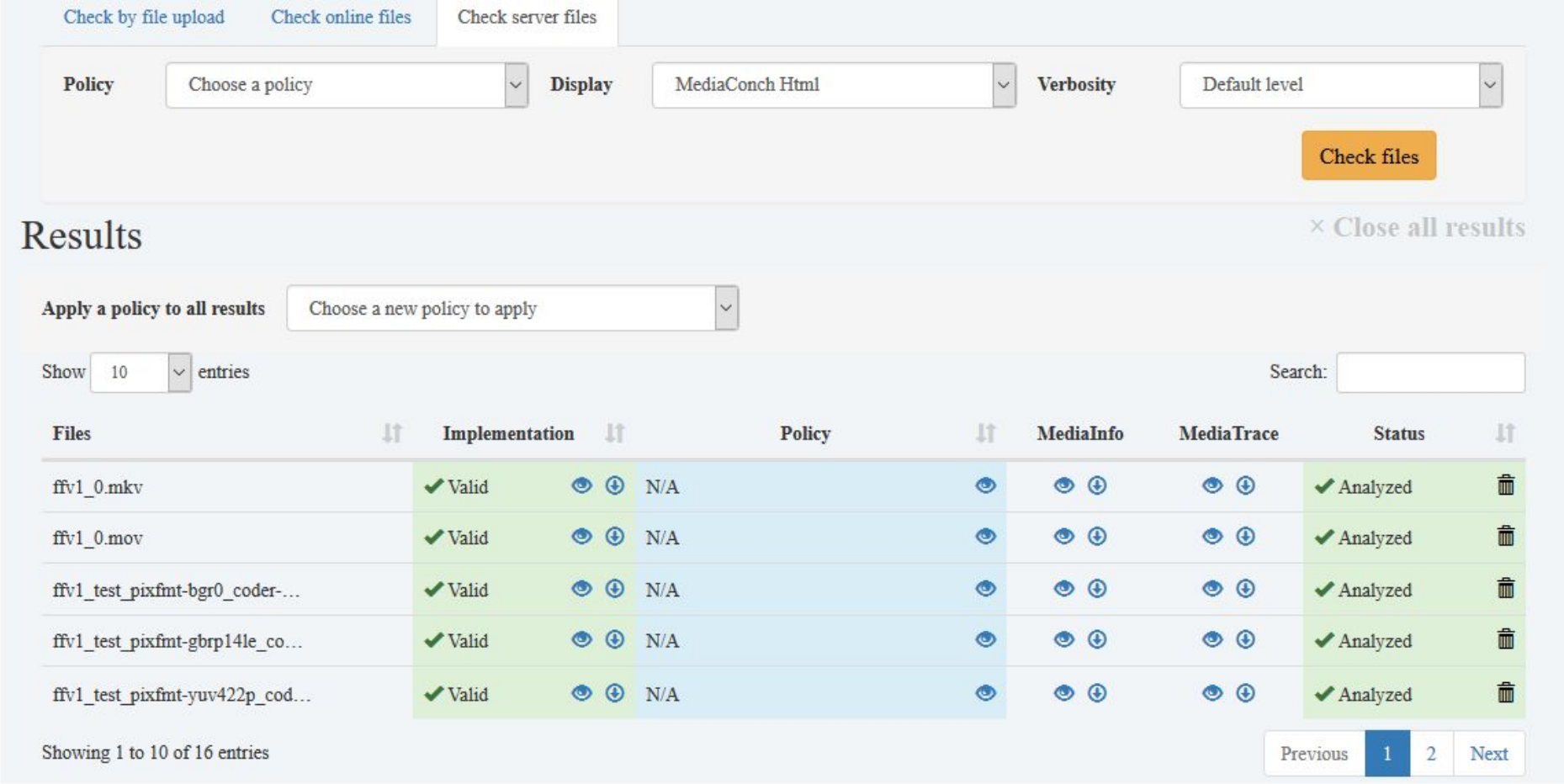
Focused on FFV1 and Matroska.

We added all formats supported by MediaInfo to the policy checker.



MediaConch

Implementation and Policy reporter



MediaConch

Compared to a specification
or
Compared to your policy

Implementation report:

MediaConch Report

File: C:/temp/FFV1+PCM_WithChecksum_Untouched.mkv

MediaConch EBML Implementation Checker

Toggle all verbosity:

- **EBML-ELEM-START** Tests run: 1 | Results: ✓
- **EBML-VER-COH** Tests run: 1 | Results: ✓
- **EBML-DOCVER-COH** Tests run: 1 | Results: ✓
- **EBML-ELEMENT-VALID-PARENT** Tests run: 87 | Results: ✓
- **EBML-ELEMENT-NONMULTIPLES** Tests run: 70 | Results: ✓
- **EBML-ELEMENT-CONTAINS-MANDATES** Tests run: 43 | Results: ✓
- **EBML-ELEMENT-IN-SIZE-RANGE** Tests run: 43 | Results: ✓
- **EBML-VALID-MAXID** Tests run: 1 | Results: ✓
- **EBML-VALID-MAXSIZE** Tests run: 1 | Results: ✓
- **HEADER-ELEMENTS-WITHIN-IDLENGTH-LIMIT** Tests run: 1 | Results: ✓
- **ELEMENTS-WITHIN-MAXIDLENGTH** Tests run: 1 | Results: ✓
- **HEADER-ELEMENTS-WITHIN-MAXSIZELENGTH** Tests run: 1 | Results: ✓
- **ELEMENTS-WITHIN-MAXSIZELENGTH** Tests run: 1 | Results: ✓
- **MKV-SEEK-RESOLVE** Tests run: 4 | Results: ✓
- **EBML-CRC-FIRST** Tests run: 6 | Results: ✓
- **EBML-CRC-VALID** Tests run: 6 | Results: ✓
- **MKV-VALID-TRACKTYPE-VALUE** Tests run: 2 | Results: ✓
- **MKV-VALID-BOOLEANS** Tests run: 3 | Results: ✓

MediaConch FFV1 Implementation Checker

- **FFV1-SLICE-CRC-VALID** Tests run: 4 | Results: ✓

MediaConch PCM Implementation Checker

Policy report:

MediaConch Report

File: C:/temp/FFV1+PCM_WithChecksum_Untouched.mkv

- Example MKV FFV1 digitization policy ✗ fail

  Example of a digitization specification of analog SD video to FFV1 and Matroska.

  Type: and | **Rules run:** 17 | **Fail count:** 5 | **Pass count:** 12

  - **Is it Matroska?** ✓ pass
  - **Matroska version 4 or greater?** ✓ pass
  - **Unique ID is present?** ✓ pass
  - **Is the video FFV1?** ✓ pass
  - **FFV1 is version 3.4 or later?** ✓ pass
  - **FFV1 is encoded in GOP size of 1?** ✗ fail
  - **FFV1 uses slice crcs?** ✓ pass
  - **Display Aspect Ratio is 4/3?** ✗ fail (Actual: 1.222)
  - **Frame Rate is Constant?** ✓ pass
  - **ColorSpace is YUV?** ✗ fail (Actual: RGB)
  - **Chroma Subsampling is 4:2:2?** ✗ fail
  - **Audio is PCM?** ✓ pass
  - **Audio is 48000 Hz?** ✓ pass
  - Is this NTSC or PAL SD? ✗ fail
  - Bit Depth is 8 or 10? ✓ pass
  - Audio is Stereo or Mono? ✓ pass
  - Bit Depth is 16 or 24? ✓ pass

FRAME

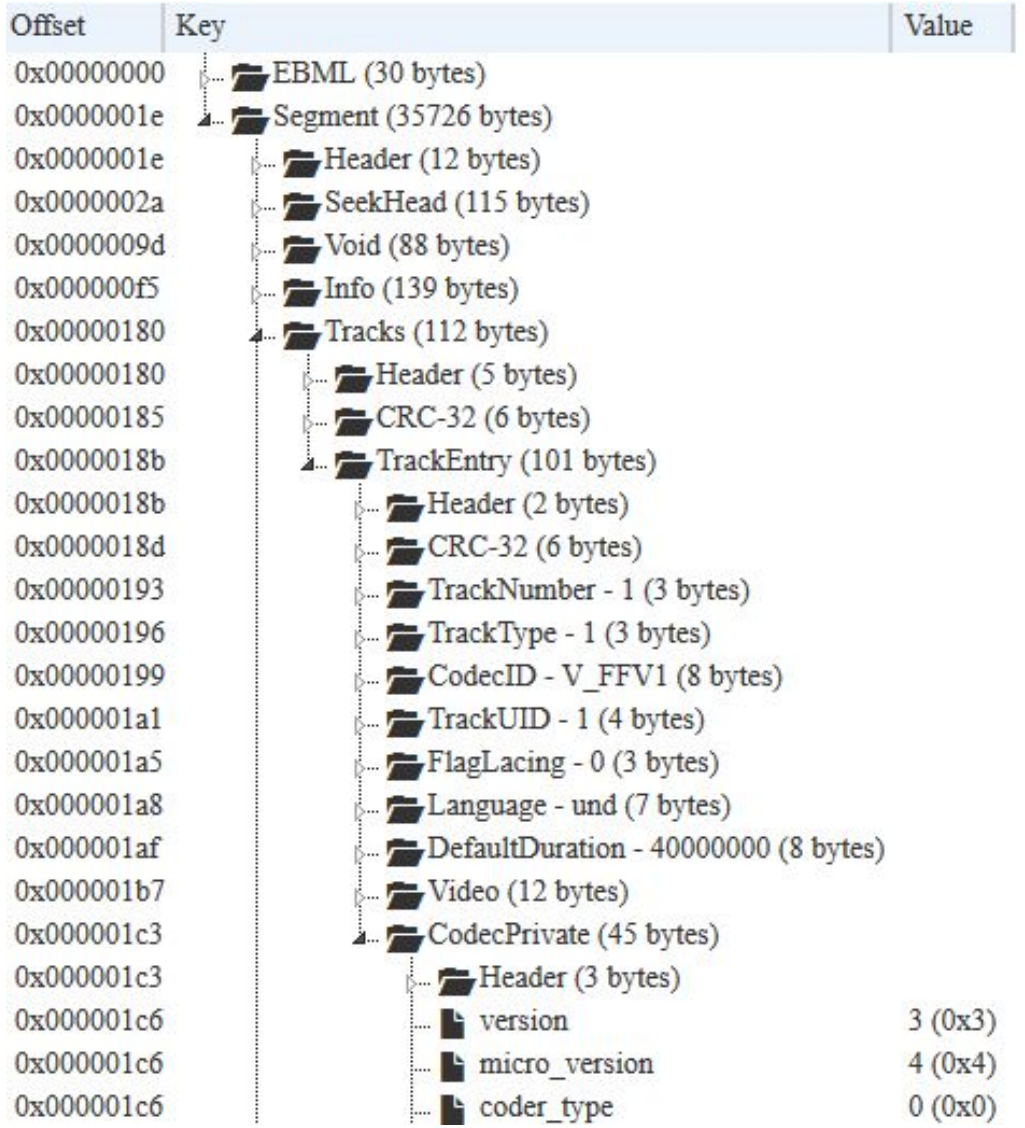
MediaConch

General information about your files

Key	Value
C:/Programmation/PreFormaMediaInfo/SampleTestFiles/FFV1/ffv1_3.mkv	
General	
UniqueID	88323790047680325859674626238128084708
Format	Matroska
Format_Version	4
FileSize	126167
Duration	1.000
OverallBitRate	1009336
FrameRate	25.000
FrameCount	25
StreamSize	2511
Video	
StreamOrder	0
ID	1
UniqueID	1
Format	FFV1
Format_Version	3.4
CodecID	V_MS/VFW/FOURCC / FFV1
Duration	1.000
BitRate	989250
Width	320

MediaConch

Inspect your files

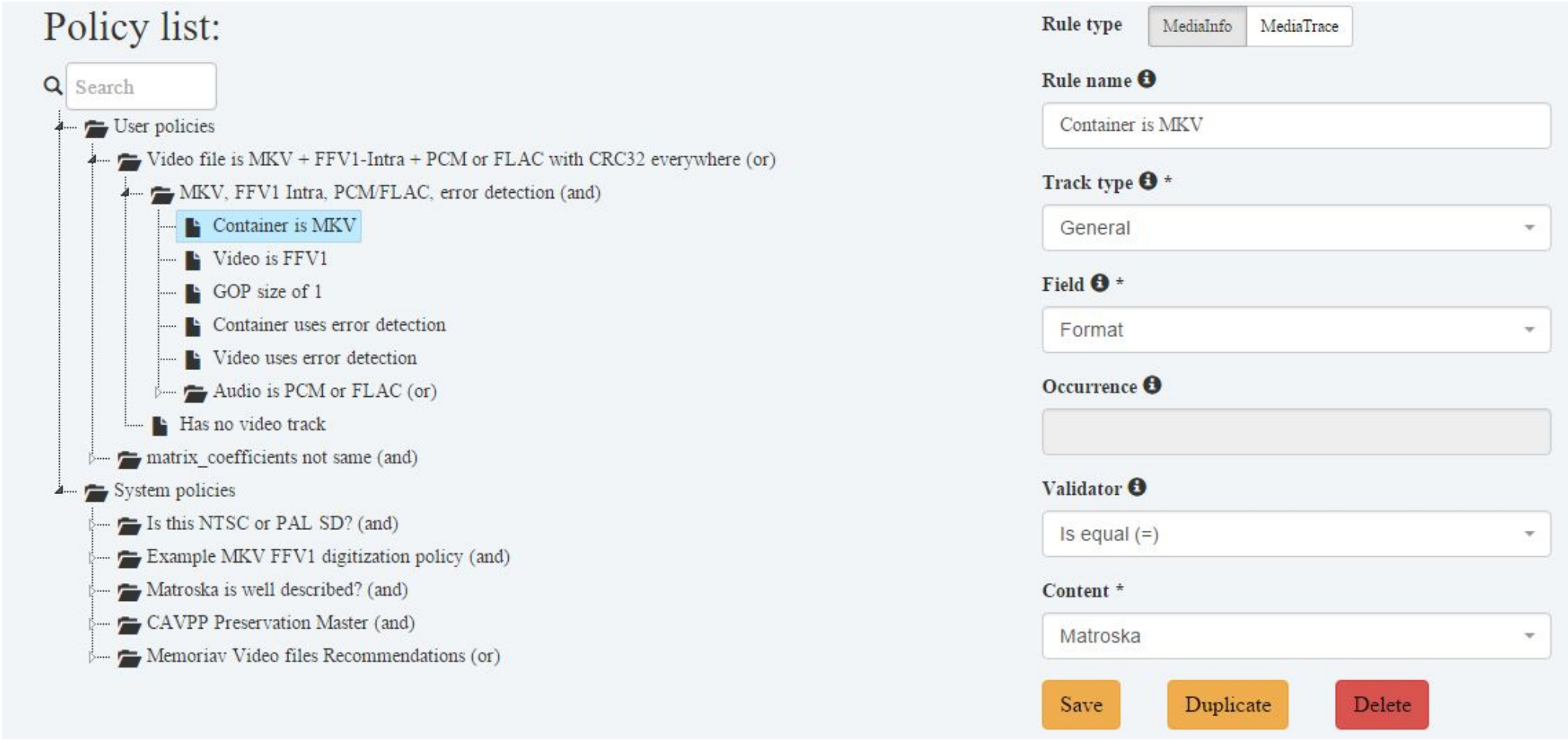


Offset	Key	Value
0x00000000	EBML (30 bytes)	
0x0000001e	Segment (35726 bytes)	
0x0000001e	Header (12 bytes)	
0x0000002a	SeekHead (115 bytes)	
0x0000009d	Void (88 bytes)	
0x000000f5	Info (139 bytes)	
0x00000180	Tracks (112 bytes)	
0x00000180	Header (5 bytes)	
0x00000185	CRC-32 (6 bytes)	
0x0000018b	TrackEntry (101 bytes)	
0x0000018b	Header (2 bytes)	
0x0000018d	CRC-32 (6 bytes)	
0x00000193	TrackNumber - 1 (3 bytes)	
0x00000196	TrackType - 1 (3 bytes)	
0x00000199	CodecID - V_FFV1 (8 bytes)	
0x000001a1	TrackUID - 1 (4 bytes)	
0x000001a5	FlagLacing - 0 (3 bytes)	
0x000001a8	Language - und (7 bytes)	
0x000001af	DefaultDuration - 40000000 (8 bytes)	
0x000001b7	Video (12 bytes)	
0x000001c3	CodecPrivate (45 bytes)	
0x000001c3	Header (3 bytes)	
0x000001c6	version	3 (0x3)
0x000001c6	micro_version	4 (0x4)
0x000001c6	coder_type	0 (0x0)

FRAME

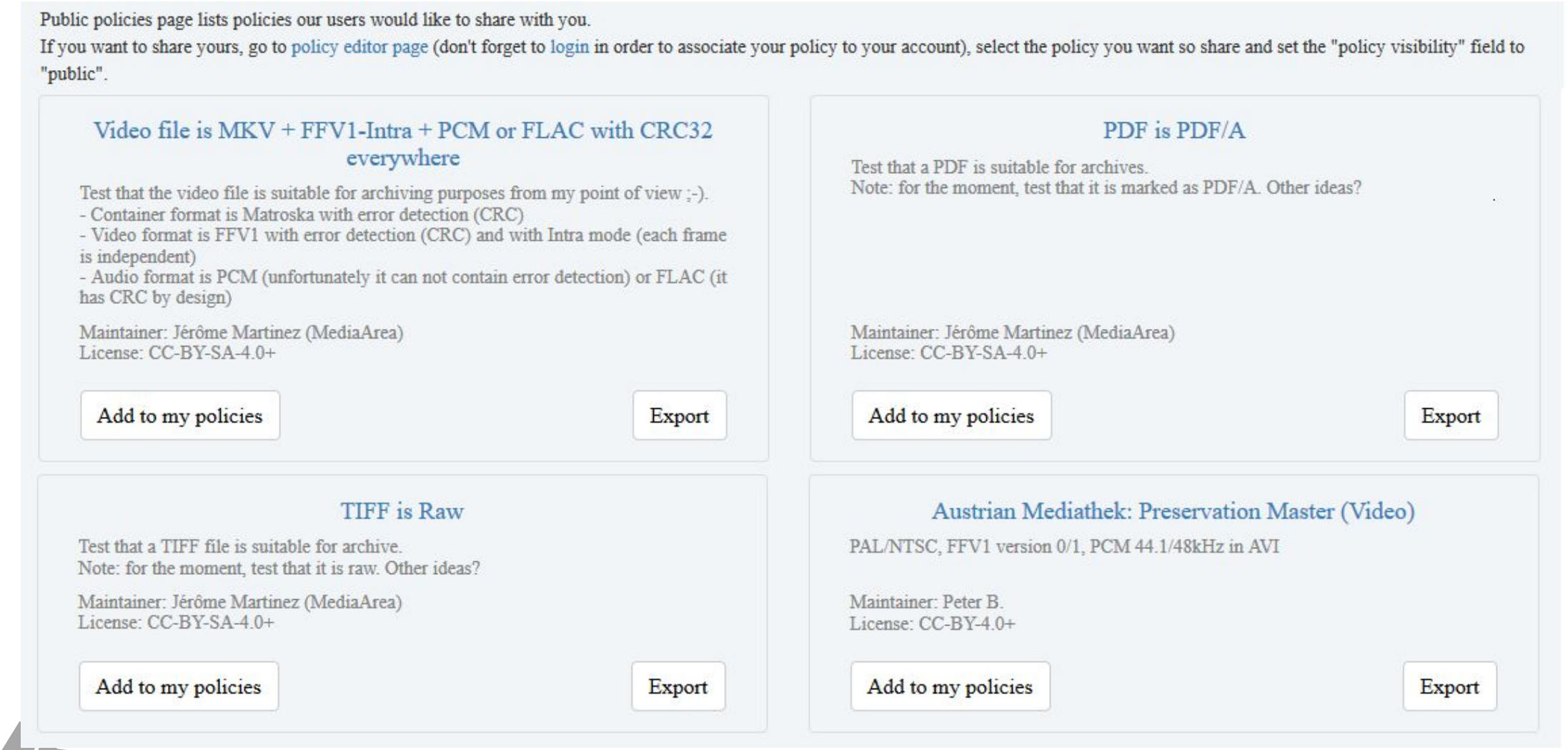
MediaConch

Policy editor



MediaConch

Public policies



DVRescue



DVRescue

DV (Digital Video) is analog.

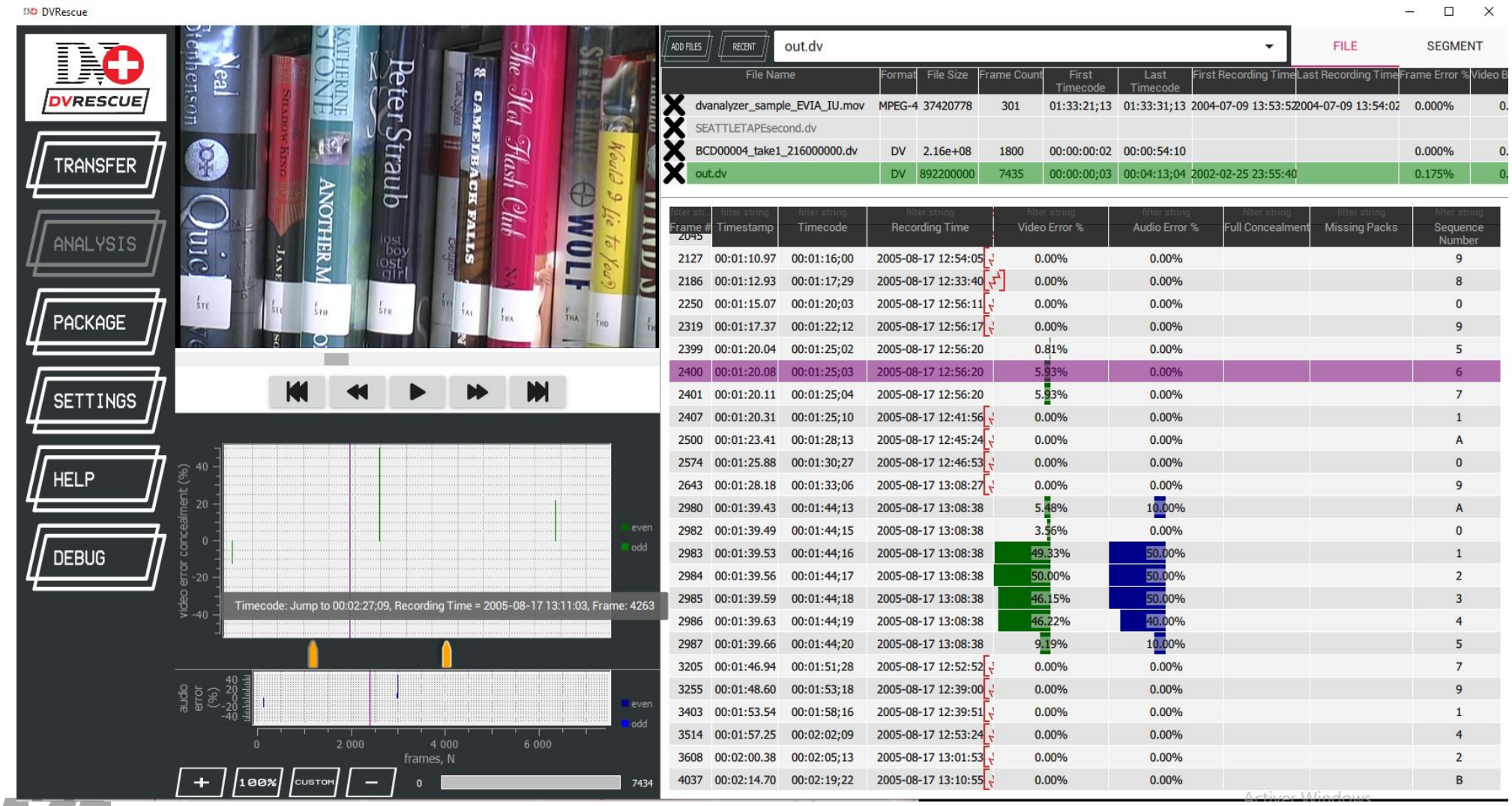
Tape will not last forever.
Players are disappearing...



DVRescue

DVRescue supports improving DV files that have already been created by allowing such files to be assessed, so that the software may selectively retry portions of the tape and incorporate any improvements into the existing file.

DVRescue


DVRescue
DVRESCUE
TRANSFER
ANALYSIS
PACKAGE
SETTINGS
HELP
DEBUG
ADD FILES
RECENT
out.dv
FILE
SEGMENT
File Name
Format
File Size
Frame Count
First Timecode
Last Timecode
First Recording Time
Last Recording Time
Frame Error %
dvanalyzer_sample_EVIA_IU.mov
MPEG-4
37420778
301
01:33:21;13
01:33:31;13
2004-07-09 13:53:52
2004-07-09 13:54:02
0.000%
SEATTLETAPEsecond.dv
BCD00004_take1_216000000.dv
DV
2.16e+08
1800
00:00:00:02
00:00:54:10
0.000%
out.dv
DV
892200000
7435
00:00:00;03
00:04:13;04
2002-02-25 23:55:40
0.175%
Frame #
Timestamp
Timecode
Recording Time
Video Error %
Audio Error %
Full Concealment
Missing Packs
Sequence Number
2127 00:01:10.97 00:01:16;00 2005-08-17 12:54:05 0.00% 0.00% 9
2186 00:01:12.93 00:01:17;29 2005-08-17 12:33:40 0.00% 0.00% 8
2250 00:01:15.07 00:01:20;03 2005-08-17 12:56:11 0.00% 0.00% 0
2319 00:01:17.37 00:01:22;12 2005-08-17 12:56:17 0.00% 0.00% 9
2399 00:01:20.04 00:01:25;02 2005-08-17 12:56:20 0.81% 0.00% 5
2400 00:01:20.08 00:01:25;03 2005-08-17 12:56:20 5.93% 0.00% 6
2401 00:01:20.11 00:01:25;04 2005-08-17 12:56:20 5.93% 0.00% 7
2407 00:01:20.31 00:01:25;10 2005-08-17 12:41:56 0.00% 0.00% 1
2500 00:01:23.41 00:01:28;13 2005-08-17 12:45:24 0.00% 0.00% A
2574 00:01:25.88 00:01:30;27 2005-08-17 12:46:53 0.00% 0.00% 0
2643 00:01:28.18 00:01:33;06 2005-08-17 13:08:27 0.00% 0.00% 9
2980 00:01:39.43 00:01:44;13 2005-08-17 13:08:38 5.48% 10.00% A
2982 00:01:39.49 00:01:44;15 2005-08-17 13:08:38 3.56% 0.00% 0
2983 00:01:39.53 00:01:44;16 2005-08-17 13:08:38 49.33% 50.00% 1
2984 00:01:39.56 00:01:44;17 2005-08-17 13:08:38 50.00% 50.00% 2
2985 00:01:39.59 00:01:44;18 2005-08-17 13:08:38 46.15% 50.00% 3
2986 00:01:39.63 00:01:44;19 2005-08-17 13:08:38 46.22% 40.00% 4
2987 00:01:39.66 00:01:44;20 2005-08-17 13:08:38 9.19% 10.00% 5
3205 00:01:46.94 00:01:51;28 2005-08-17 12:52:52 0.00% 0.00% 7
3255 00:01:48.60 00:01:53;18 2005-08-17 12:39:00 0.00% 0.00% 9
3403 00:01:53.54 00:01:58;16 2005-08-17 12:39:51 0.00% 0.00% 1
3514 00:01:57.25 00:02:02;09 2005-08-17 12:53:24 0.00% 0.00% 4
3608 00:02:00.38 00:02:05;13 2005-08-17 13:01:53 0.00% 0.00% 2
4037 00:02:14.70 00:02:19;22 2005-08-17 13:10:55 0.00% 0.00% B
video error concealment (%)
even
odd
Timecode: Jump to 00:02:27;09, Recording Time = 2005-08-17 13:11:03, Frame: 4263
audio error (%)
frames, N
100%
CUSTOM
7434

DVRescue

Help in DV tape dump automation

- Automate tape to file
- Automatically cut the file depending on aspect ratio change, non consecutive timecode…
- Show a table with a list of detected errors per frame
- Merge 2 inputs in a better content, peeking good blocks from each file

DVRescue

Still work in progress

- Automation of the rewind and dump again when a problem is detected during dump of the tape

RAWcooked

RAWcooked

The issue with raw A/V files

Huge size (4K+ is there, can be 100 MB/frame, several TB per hour)

1 file per video frame (thousands of files in a directory)

Not playable as is by several players (VLC...)

So many DPX or TIFF format flavors (interoperability issues)

RAWcooked

But… FFV1

Lossless video compression format

Open source, patent free

Adopted by several archives

Standardized (IETF, RFC 9043)

Frames are divided by slices, with checksums

RAWcooked

Compression!

Example with 1 second at 24 fps 10-bit HD film on a 6-core (12-thread) Skylake-X CPU:

- 24 DPX files (or in ZIP/TAR uncompressed): 189 MB
- 1 compressed ZIP file: 175 MB in 10 seconds
- 1 compressed LZMA2 file: 154 MB in 30 seconds
- 1 FFV1/MKV Intra 16-slice file: 105 MB in 1.5 seconds

RAWcooked

Advantages

- Open source and free (FFmpeg)
- Package is 1.5-3x smaller than DPX/TIFF
- Checksum by "Cluster" (usually 1 second) at container level
- Checksum by "Slice" (you choose how many per frame) at video level
- Files are natively playable by several players (VLC, FFplay...)

RAWcooked

Disadvantages of FFV1 alone

- You lose some metadata
  (DPX/TIFF header: scan software, some colorimetry info, film type, DPX time code, shutter angle, gamma...)
- Complicated command line

But...

RAWcooked

A new tool

Easy: just a short command line: "rawcooked YourDirectoryName"
Store DPX/TIFF headers/footers in a specific Matroska attachment
Store other sidecar files as Matroksa attachments
Output is a single Matroska/FFV1/FLAC file
Encoding is reversible (bit-by-bit to original files): "rawcooked YourMatroskaFileName.mkv"
Integrated quality management "rawcooked --all YourDirectoryName" adds hash of input files, check the conformance of DPX, check that the files can correctly reverted...

RAWcooked

Use case

Archive asks a digitalization to their supplier: classic workflow with the scanner + "rawcooked --all YourDirectoryName"
Transport... (2x less file sizes, less costly)
Archive receives content & checks the integrity: "rawcooked --check YourMatroskaFileName.mkv"
Archive can visually check the content with e.g. VLC Media Player
Storage (cost divided by 2 due to compression)
Revert to exact original DPX if someone needs it: "rawcooked YourMatroskaFileName.mkv"

RAWcooked

Work in progress, always

- Lack of GUI
- New features planned: error correction codes, speedup improvements...

RAWcooked

Users:

- AV Preservation by reto.ch (main sponsor)
- National Audiovisual Centre Luxembourg (CNA)
- National Library of Norway
- Irish Film Institute (IFI)
- Northwest University Library
- National Library of Wales
- Walter J. Brown Media Archives
- The MediaPreserve
- British Film Institute

RAWcooked

A different way to do open source

Open source code provided without lock to sponsors
Deliveries on our website are with a lock
DPX 8/10 bit RGB & WAV 2ch 48kHz flavors are usable by default
We provide a key for other format flavors and features (temporary key possible)
1000 € for first flavor/feature
+ 500 € per additional flavor/feature
To be compared with storage cost saving (storage cost divided by 2)

Some other tools

MediaInfo

Convenient unified display of the most relevant technical and tag data for video and audio files.

Output: text, XML, HTML, PBCore, EBUCore, NISO Z39.87...

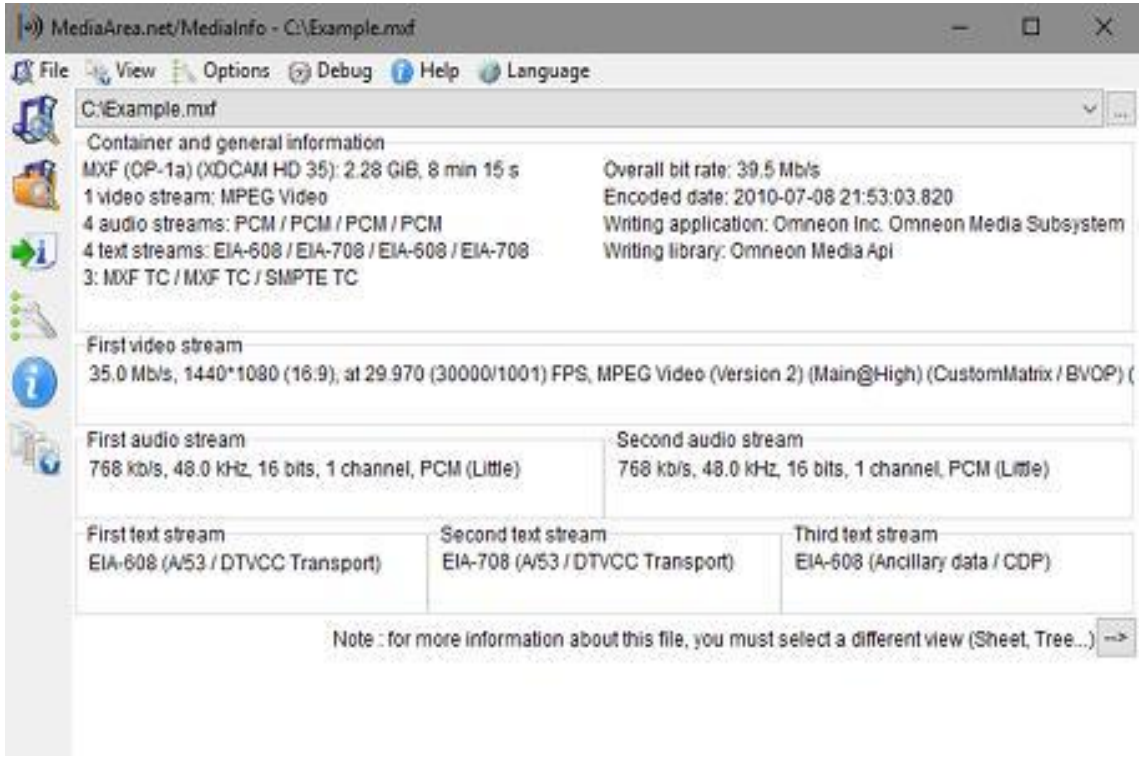
Container: format, profile, commercial name of the format, duration, overall bit rate, writing application and library, title, author, director, album, track number, date, duration...

Video: format, codec id, aspect, frame rate, bitrate, color space, chroma subsampling, bit depth, scan type, scan order...

Audio: format, codec id, sample rate, channels, bit depth, language, bit rate...

Text: format, codec id, language of subtitle...

Chapters: count of chapters, list of chapters...

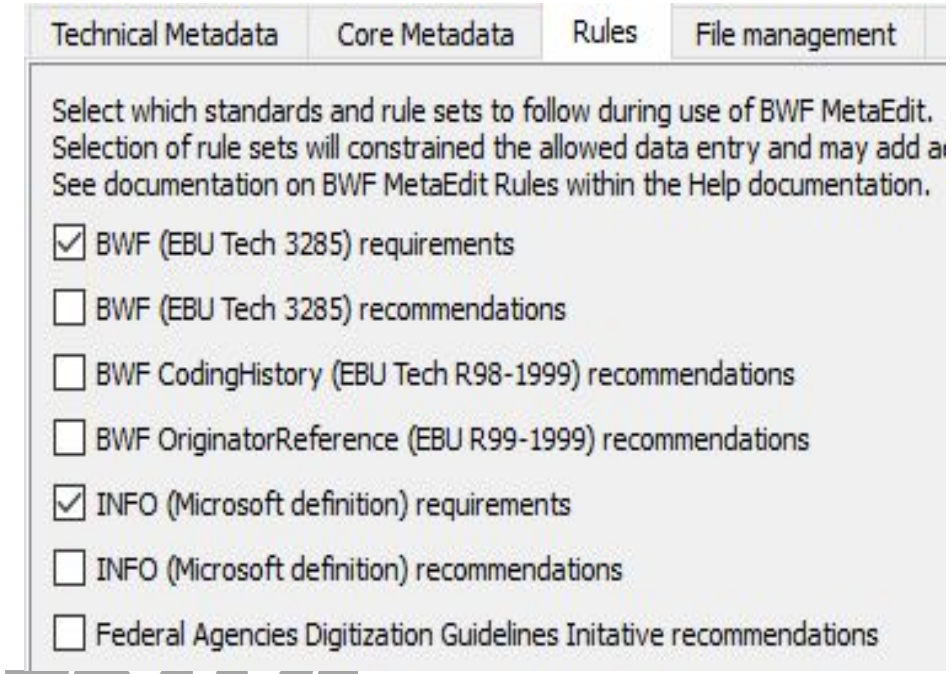


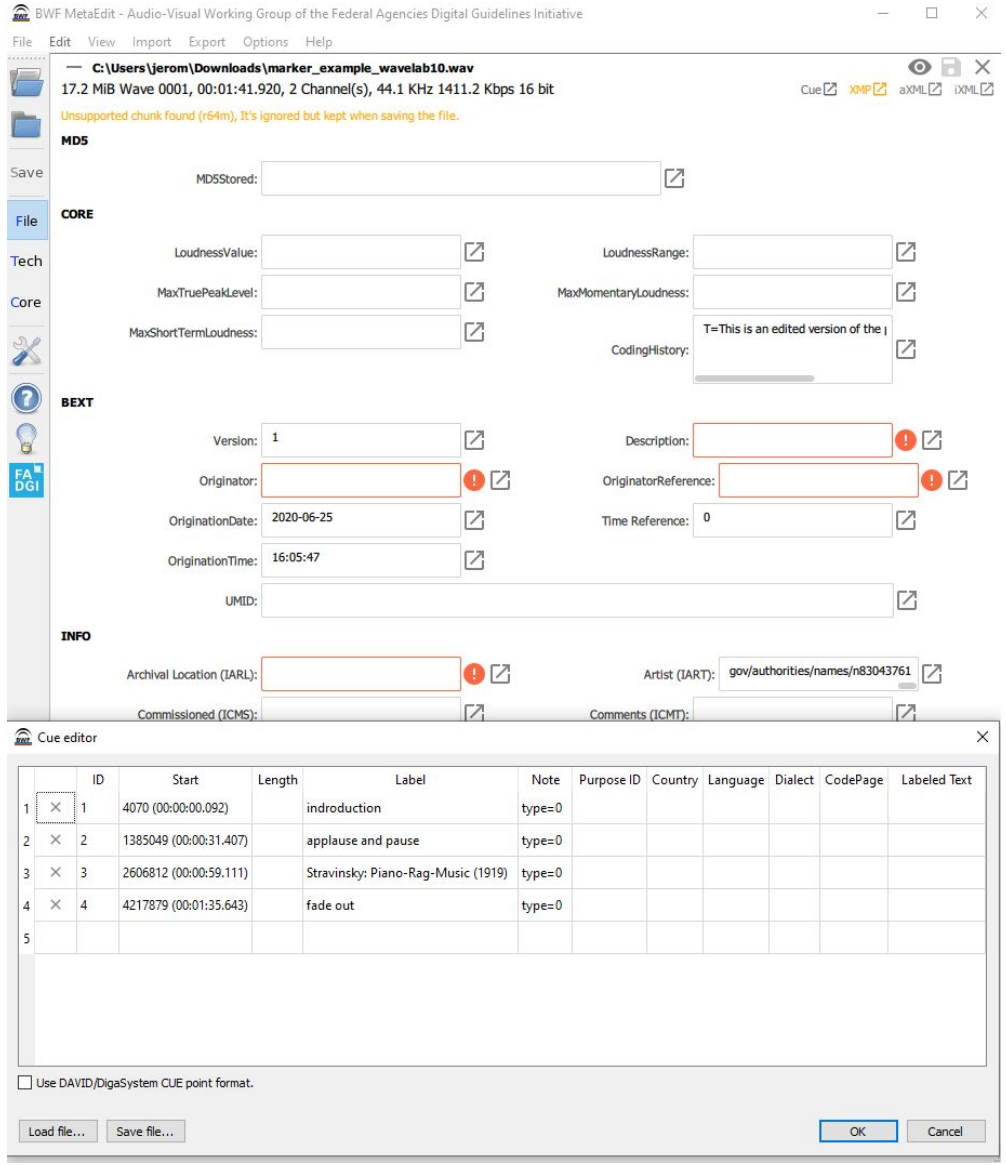
BWF MetaEdit

Supports embedding, validating, and exporting of metadata in Broadcast WAVE Format (BWF) files.

It supports the FADGI Broadcast WAVE Metadata Embedding Guidelines.

Last version includes a Cue Editor





Platform support

Platform support

We try to support as many platforms as possible:

Windows 7+
macOS 10.10+
Debian 8+
Ubuntu 16.04+
CentOS/RHEL 7+
openSUSE/SLE 12+

CPU: x86, x86_64/x64, ARM (including Apple M1),...

How was it created?

How was it created?

By users for users!

Each tool started from the need of a user.

MediaArea is here to technically help, but users are at the driver seat.

No restriction on the usage

Stay in touch

http://mediaarea.net jerome@mediaarea.net

License (except images): CC-BY