

# MediaConch

Implementation and policy checking  
on FFV1, Matroska, LPCM, and more



Jérôme Martinez  
MediaArea.net SARL

Open Source Preservation Workshop - April 2016
















































# What is MediaConch?

MediaConch is a conformance checker

- Implementation checker
- Policy checker
- Reporter
- Fixer

# What is MediaConch?

## Implementation and Policy reporter

| MediaConch              |   | Checker   | Result  | Policies  | Display  | Help |                     |
|-------------------------|---|---|---|---|--|------|---------------------|
|                         |   |   |   |   |  |      | × Close all results |
| Files                   | Implementation  | Policy  | MediaInfo   | MediaTrace  | Status   |      |                     |
| Coconut.mp4             | ✓ Valid       | ✗ Is Matroska       |       |       | Analyzed    |      |                     |
| Exampleä.mp4            | ✓ Valid       | ✗ Is Matroska       |       |       | Analyzed    |      |                     |
| ffv1_3 - Copie (2).m... | ✓ Valid       | ✓ Is Matroska       |       |       | Analyzed    |      |                     |
| ffv1_3 - Copie.mkv      | ✓ Valid   | ✓ Is Matroska   |   |   | Analyzed  |      |                     |
| ffv1_3.mkv              | ✓ Valid   | ✓ Is Matroska   |   |   | Analyzed  |      |                     |

# What is MediaConch?

## Example of report

| <b>MKV/FFV1 (reformatted QT/v210)</b>   |                |         |          |         |              |
|---|----------------|---------|----------|---------|--------------|
| Policy check for Matroska-wrapped FFV1 and LPCM file, transcoded from QuickTime-wrapped v210 and LPCM file. |                |         |          |         |              |
| <b>General Format equals Matroska</b>   |                |         |          |         |              |
| <b>Context (field): Format</b>  |                |         |          |         |              |
| <b>Context (value): Matroska</b>  |                |         |          |         |              |
| tracktype   | tracktypeorder | trackid | actual   | outcome | reason       |
| General   |                |         | Matroska | ✓ pass  |              |
| <b>Video Width is equal to 720 (pixels)</b>   |                |         |          |         |              |
| <b>Context (field): Width</b>   |                |         |          |         |              |
| <b>Context (value): 720</b>   |                |         |          |         |              |
| tracktype   | tracktypeorder | trackid | actual   | outcome | reason       |
| Video   |                | 1       | 640      | ✗ fail  | is not equal |

# What is 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                                    |

# What is MediaConch?

Inspect your files

| Offset     | Key                                    | Value     |
|------------|--|-----------|
| 0x00000000 | EBML (47 bytes)                        |           |
| 0x0000002f | Segment (126120 bytes)                 |           |
| 0x0000002f | Header (12 bytes)                      |           |
| 0x0000003b | SeekHead (66 bytes)                    |           |
| 0x0000007d | Void (157 bytes)                       |           |
| 0x0000011a | Info (81 bytes)                        |           |
| 0x0000016b | Tracks (167 bytes)                     |           |
| 0x0000016b | Header (12 bytes)                      |           |
| 0x00000177 | TrackEntry (155 bytes)                 |           |
| 0x00000177 | Header (9 bytes)                       |           |
| 0x00000180 | TrackNumber - 1 (3 bytes)              |           |
| 0x00000183 | TrackUID - 1 (4 bytes)                 |           |
| 0x00000187 | FlagLacing - 0 (3 bytes)               |           |
| 0x0000018a | Language - eng (7 bytes)               |           |
| 0x00000191 | TrackType - 1 (3 bytes)                |           |
| 0x00000194 | DefaultDuration - 40000000 (8 bytes)   |           |
| 0x0000019c | CodecID - V_MS/VFW/FOURCC (17 bytes)   |           |
| 0x000001ad | Video (16 bytes)                       |           |
| 0x000001bd | CodecPrivate - Copy of vids (85 bytes) |           |
| 0x000001bd | Header (3 bytes)                       |           |
| 0x000001c0 | Size                                   | 81 (0x51) |

# What is MediaConch?

## Policy editor

To create policies, users may select fields from MediaInfo's stream types (General, Video, Audio, etc.) and apply validators (Equal, Not Equal, Greater Than, etc.) to the associated field values. For example, a policy item may state that the "Width" field from the file's Video stream equal 640 pixels. When adding the field's value, however, be sure to leave off any associated strings, such as "pixels" "Kbps" "fps", etc.

**Rule name \***

Quicktime wrapper

**Editor type :**  Editor  Free text

**Track type \***

General

**Field \***

Format\_Profile

**Occurrence \***

\*

**Validator \***

Is equal (==)

**Value**

QuickTime

### CAVPP Preservation Video Files rules

Quicktime wrapper



10 bit uncompressed 4:2:2 video



Total Bitrate



Frame Rate 29.970



Aspect Ratio 4/3



Is Interlaced



Audio is 48kHz



Audio is stereo



Audio is 24 bit



Add a new rule

# MediaConch interfaces

- Graphical interface
- Web interface
- Command line
- Server (REST API)
- (Work in progress) a library (.dll/.so/.dylib)



# MediaConch output formats

- XML (native format)
- Text
- HTML
- (Work in progress) PDF
- Tweakable! (with XSL)

# Open source

- GPLv3+ and MPLv2+
- Relies on MediaInfo (metadata extraction tool)
- Use well-known open source libraries: Qt, sqlite, libevent, libxml2, libxslt, libexslt...

# Supported formats

- Priorities for the implementation checker
  - Matroska
  - FFV1
  - PCM
- Can accept any format supported by MediaInfo for the policy checker
  - MXF + JP2k
  - QuickTime/MOV
  - Audio files (WAV, BWF, AIFF...)
  - ...

# Supported formats

Can be expanded

- By plugins
  - Support of PDF checker: VeraPDF plugin
  - Support of TIFF checker: DPF Manager plugin
  - You use another checker? Let us know
- By internal development
  - More tests on your preferred format is possible
  - It depends on you!

# Versatile

Several input formats are accepted

- FFV1 from MOV or AVI
- Matroska with other video formats
- (Work in progress) Extraction of a PDF or TIFF attachment from a Matroska container and analyze with a plugin (e.g. VeraPDF and DPF Manager)
- ...

# Versatile

Input can be from:

- Files (local/network)
- FTP/FTPS/SFTP
- HTTP/HTTPS
- Amazon S3

# Versatile

Binaries are provided for:

- Windows
- Mac

Homebrew users: "brew install mediaconch", that's all!

- Linux (Ubuntu, Debian, Fedora, OpenSUSE...)

Ubuntu 16.04 and Debian Testing/9 users:

"apt-get install mediaconch", that's all!

(it is in the official distros repository)

- Embedded devices? Doable

(we tested it on a Raspberry Pi )

- Can be ported on other distros (BSD...)

# Standardization

- Matroska is widely used but not (yet) standardized
- FFV1 is gaining increasing usage in preservation contexts but is not (yet) standardized



# CELLAR: IETF workgroup

- Open standards group
- Goal to IETF-standardize Matroska/FFV1/FLAC
- A lot of progress, especially with Matroska/EBML specs
- <https://datatracker.ietf.org/wg/cellar/charter/>

# Worldwide

- 2 project leaders
  - Jérôme Martinez (Digital Media Analysis Specialist, France)
  - Dave Rice (Archivist, USA)
- Presentations worldwide
  - IASA, France
  - FIAT/IFTA, Austria
  - FOSDEM, Belgium
  - AMIA, USA
  - Code4Lib, USA
  - JTS, Singapore
  - (3-6 October 2016) IPRES, Switzerland
  - (25-29 September 2016) IASA, USA

# Matroska research corpus

- We analyze all Matroska files from archive.org
- Interface with some statistics of Matroska elements usage (e.g. files with CRC-32 elements...)

<https://mediaarea.net/MediaConchCorpus/>

# What's next?

Still under development but already usable  
(PREFORMA prototyping phase up to end 2016)


- Better handling of huge collections
- Better user interface
- Statistics
- Standardize Matroska and FFV1
- More conformance tests
- Integration in Archivematica
- Fixer

# And after PREFORMA sponsorship?

It depends on you!

- This is open source
- Driven by user requests
- Everyone can develop or sponsor a development
- Potential features:
  - Integration of test of your preferred format  
(MXF? doable. JP2k? doable. WAV? doable...)
  - Integration of other checkers  
(BWF MetaEdit? QCTools?)
  - Better integration in your workflow
  - ...

# Example (Online)

 **MediaConch** | Home | Checker | Policies | Display | Help ▾ | Logged in as ashley ▾

## Check files

[Check by file upload](#) | [Check online files](#) | [Check server files](#)
































**Policy**  ▾ | **Display**  ▾

[Check files](#)

### Results

× Close all results

Show  entries | Search:

| Files                                  | Implementation | Policy | MediaInfo   | MediaTrace  | Status   |   |
|--|----------------|--------|---|---|----------|---|
| ffv1_0.mov                             |                |        |       |       | Analyzed |    |
| ffv1_test_pixfmt-yuva444p_coder-1_1... |                |        |       |       | Analyzed |    |
| ffv1_test_pixfmt-bgr0_coder-0_level... |                |        |   |   | Analyzed |  |
| ffv1_test_pixfmt-gbrp14le_coder-1_1... |                |        |   |   | Analyzed |  |
| ffv1_0.mkv                             |                |        |   |   | Analyzed |  |
| ffv1_test_pixfmt-yuv422p_coder-0_le... |                |        |   |   | Analyzed |  |
| ffv1_test_pixfmt-yuv420p_coder-        |                |        |   |   | In queue |  |



# Example (Plugins)

## Check files

Check by file upload

Check online files

Check server files

Policy

Policy Set Example

Display

Choose a display

Verbosity

Default level

Check files

## Results

× Close all results

Show 10 entries

Search:

| Files                                  | Implementation | Policy               | MediaInfo | MediaTrace | Status   |
|--|----------------|----------------------|-----------|------------|----------|
| ffv1_test_pixfmt-yuv444p10le_coder-... | ✓ Valid        | ✓ Policy Set Example |           |            | Analyzed |
| ffv1_test_pixfmt-yuva422p_coder-1_1... | ✓ Valid        | ✓ Policy Set Example |           |            | Analyzed |
| ffv1_test_pixfmt-yuva444p_coder-1_1... | ✓ Valid        | ✓ Policy Set Example |           |            | Analyzed |
| veraPDF test suite 6-1-10-t01-pass-... | ✓ Valid        | N/A                  |           |            | Analyzed |
| train1.tif                             | ✗ Not valid    | N/A                  |           |            | Analyzed |

Showing 11 to 15 of 15 entries

Previous 1 2 Next



# Example (Plugins)



Implementation report

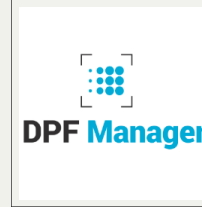
```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<cliReport xmlns="http://www.verapdf.org/ValidationProfile">
  <itemDetails size="41437">
    <name>veraPDF test suite 6-1-10-t01-pass-a.pdf</name>
  </itemDetails>
  <validationResult flavour="PDF_A_1_B" totalAssertions="476" isCompliant="true">
    <assertions/>
  </validationResult>
</cliReport>
```

Download implementation report Close

Show 10 entries Search:

| Files                                  | Implementation | Policy               | MediaInfo | MediaTrace | Status   |
|--|----------------|----------------------|-----------|------------|----------|
| ffv1_test_pixfmt-yuv444p10le_coder-... | ✓ Valid        | ✓ Policy Set Example |           |            | Analyzed |
| ffv1_test_pixfmt-yuva422p_coder-1_1... | ✓ Valid        | ✓ Policy Set Example |           |            | Analyzed |
| ffv1_test_pixfmt-yuva444p_coder-1_1... | ✓ Valid        | ✓ Policy Set Example |           |            | Analyzed |
| veraPDF test suite 6-1-10-t01-pass-... | ✓ Valid        | N/A                  |           |            | Analyzed |
| train1.tif                             | ✗ Not valid    | N/A                  |           |            | Analyzed |

# Example (Plugins)

A screenshot of a software interface showing a dialog box titled "Implementation report". The dialog box contains XML metadata for a file named "train1.tif". The XML structure is as follows:

```
<?xml version="1.0" encoding="UTF-8"?>
<globalreport>
  <individualreports>
    <report>
      <file_info>
        <name>train1.tif</name>
        <fullpath>train1.tif</fullpath>
      </file_info>
      <tiff_structure>
        <ifdTree>
          <ifdNode>
            <number>0</number>
            <isimg>yes</isimg>
            <imagetype check_ifd0="typ">Main image</imagetype>
            <image_representation>strips[1]</image_representation>
            <photometric>0</photometric>
            <hasSubIfd>no</hasSubIfd>
            <hasExif>yes</hasExif>
            <hasXMP>yes</hasXMP>
            <hasIPTC>yes</hasIPTC>
            <tags>
              <tag>
                <name>NewSubfileType</name>
                <id>254</id>
                <value>0</value>
              </tag>
            </tags>
          </ifdNode>
        </ifdTree>
      </tiff_structure>
    </report>
  </individualreports>
</globalreport>
```



# Stay in touch

MediaArea: <https://mediaarea.net>, @MediaArea\_net

MediaConch: <https://mediaarea.net/MediaConch>, @MediaConch

Jérôme Martinez: [jerome@mediaarea.net](mailto:jerome@mediaarea.net)

Slides: <https://mediaarea.net/Events>

License: CC BY