# FFV1
# for preservation

Bert Lemmens

**PACKED**
Centre of Expertise in Digital Heritage

Peter Bubestinger

österreichische**mediathek** ⬤
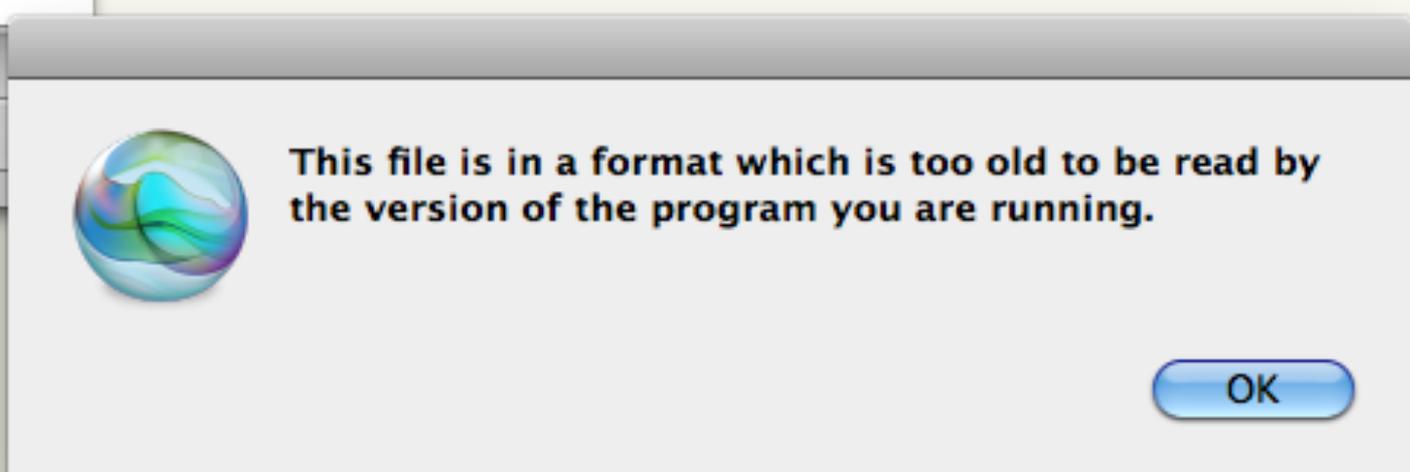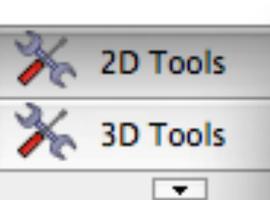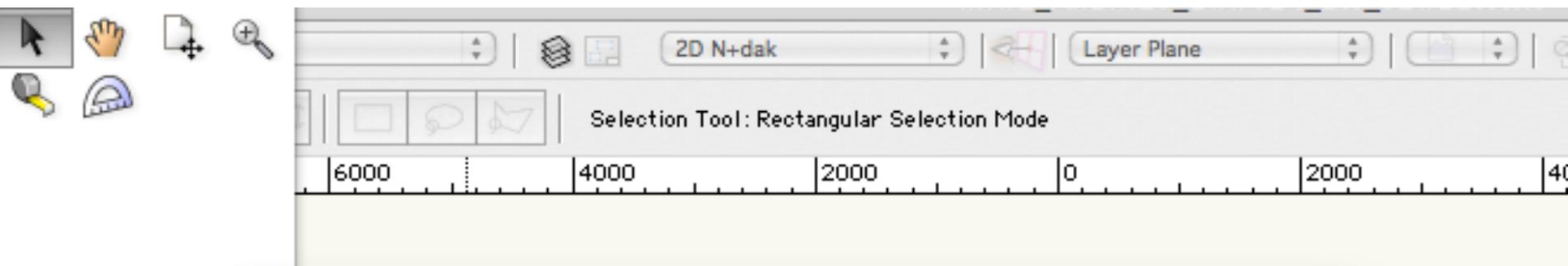audiovisuelles archiv · technisches museum wien

Tessa Fallon

**MediaArea.net**

FIAT-IFTA, Vienna | 8 October 2015

# TOPICS

- PREFORMA challenge brief
- standard specifications selected
- FFV1 for preservation
- FFV1 features
- FFV1 standardisation
- MediaConch conformance checker

2D N+dak

Layer Plane

Selection Tool: Rectangular Selection Mode

6000      4000      2000      0      2000      4

2D Tools

3D Tools

This file is in a format which is too old to be read by the version of the program you are running.

OK

# Something is technically wrong.

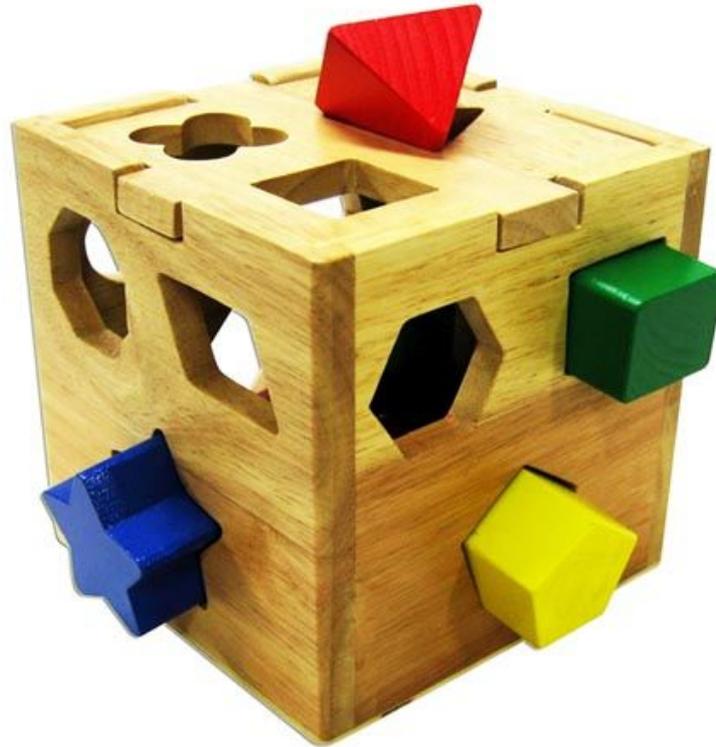Thanks for noticing—we're going to fix it up and have things back to normal soon.

PREFORMA
**FUTURE MEMORY STANDARDS**

"Empower memory institutions to gain full control over the technical properties of digital content intended for long-term preservation."

*PREFORMA Challenge Brief*

# VALIDATOR



"something that tells you whether
a file is what it claims to be"

# CHALLENGE BRIEF

## #1. Develop
## an open source conformance checker that:

- checks if a file complies with standard specifications
- checks if a file complies with acceptance criteria of memory institutions
- reports back to human and software agents
- perform simple fixes

# CHALLENGE BRIEF

#2. Establish an ecosystem around an open source reference implementation that:

- advances improvement of the standard specification
- advances development of new business cases for managing preservation files
- generates useful feedback for those who control software that implements the specification

# OPEN SOURCE

- aim for establishing a sustainable research and development community, with a wide range of contributors and users from different stakeholder groups.

- ensure long-term availability of the software, beyond the memory institutions and suppliers involved in PREFORMA.

- licenses:
  - **software**: "GPLv3 or later and MPLv2 or later"
  - **digital assets:** Creative Commons CC-BY v4.0 and in open file formats

# PROJECT

- pre-commercial procurement project, co-funded by the European commission under the FP7-ICT Programme

- 2.805.000 euro R&D budget

- 48 months ( 1 Jan 2015 – 31 Dec 2017)

- Riksarkivet, Sweden (Coordinator)

- PROMOTOR, Italy (Technical Coordinator)

- PACKED, Belgium (WP2 Lead – R&D Tender)

# PROJECT

## R&D partners

- Open Preservation Foundation | PDF Association | Digital Preservation Coalition | Dual Lab | KEEP
- EasyInnova | University of Girona | University of Basel
- MediaArea.net

## Technical partners

- PACKED Centre of Expertise in Digital Heritage | Belgium
- Fraunhofer Institute for Digital Media Technology | Germany
- University of Skovde |  Sweden
- University of Padua | Italy

## Memory institutions

- National Archives | Sweden
- Netherlands Instituut for Sound and Vision | Netherlands
- Royal Institute for Cultural Heritage | Belgium
- Greek Film Center | Greece
- Local Government Management Agency | Ireland
- Prussian Cultural Heritage Foundation, Germany
- City of Girona | Spain
- Ministry of Culture | Estonia
- National Library | Sweden

# PRE-COMMERCIAL PROCUREMENT

- procuring R&D services
- developing innovative solutions that address specific public sector challenges and needs
- competition-like procurement method
- phased procurement of development contracts to reduce risk.
- end result: multiple solutions for the same problem, almost ready to bring them on the market

# PRE-COMMERCIAL PROCUREMENT
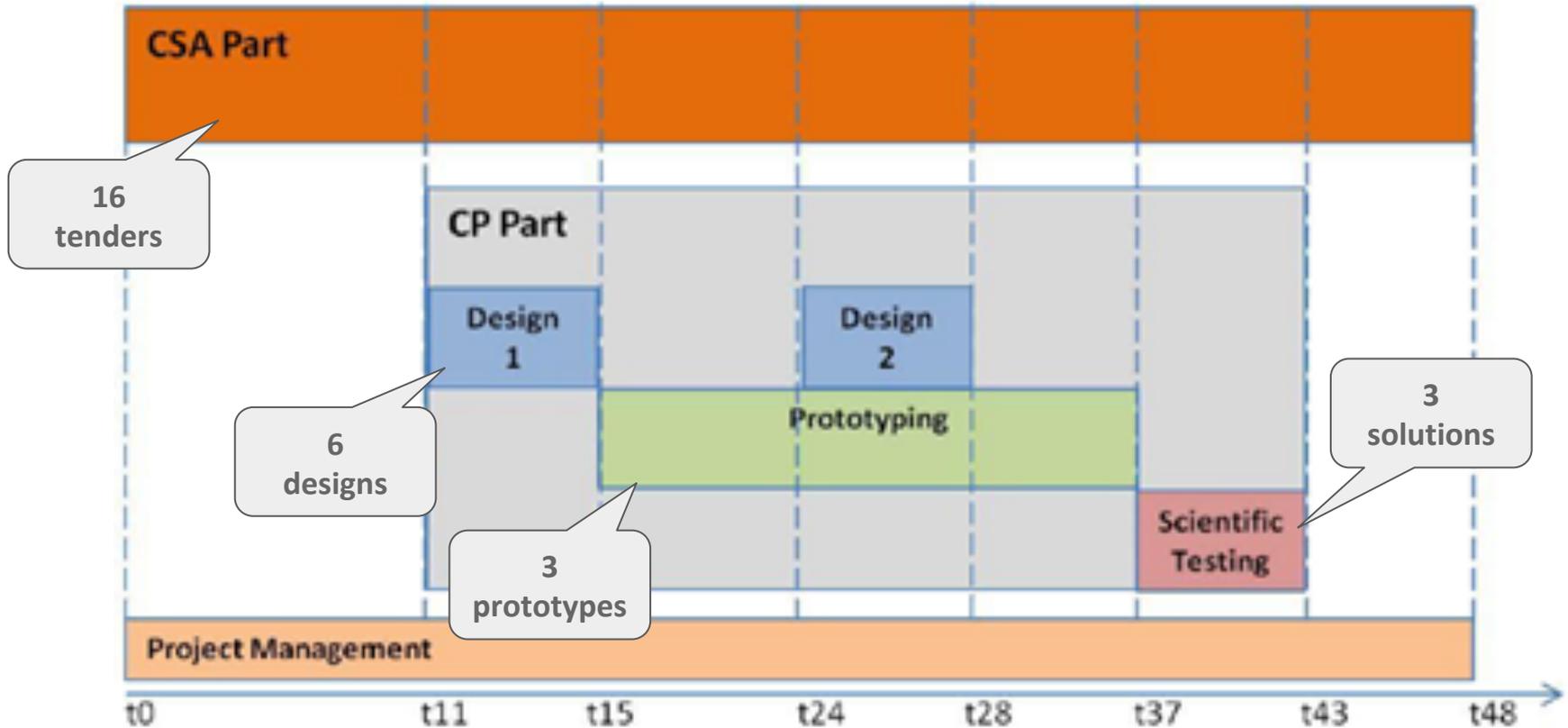
# STANDARD SPECIFICATIONS?

MXF | MPEG | IMX | XDCAM | HD422 | DPX | DCP | JPEG2000 | MOV | MPEG2 | AVI | MPEG4 | AVC | PDF 1.4 | PDF/A1 | TIFF 6.0 | JPEG | RAW | AS-07 | MPEG-AF | PDF | MKV | FFV1 | Dirac | PNG | WebM | VP8 | OGG | Theora | PDF/A2 | PDF/A3 | LPCM

# PRESERVATION FILE FORMAT?

- a life-cycle that is as long as possible

- capture a digital copy of the original (digital or analog) document in a quality that is high enough to substitute the original in case it is damaged or destroyed
- intended for storing content in a trusted repository.

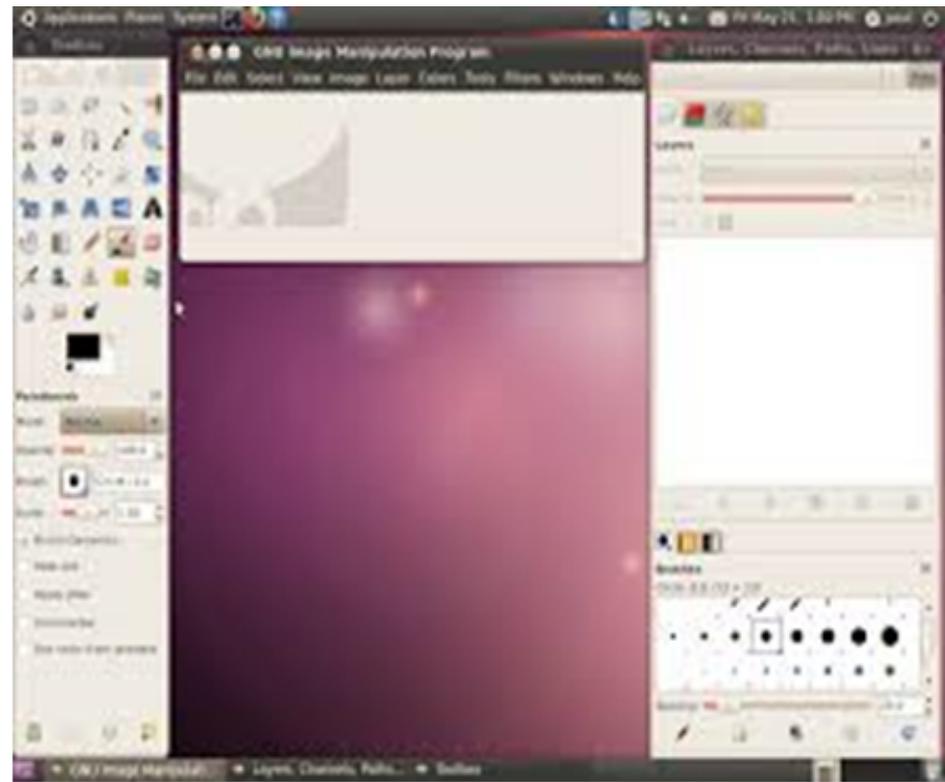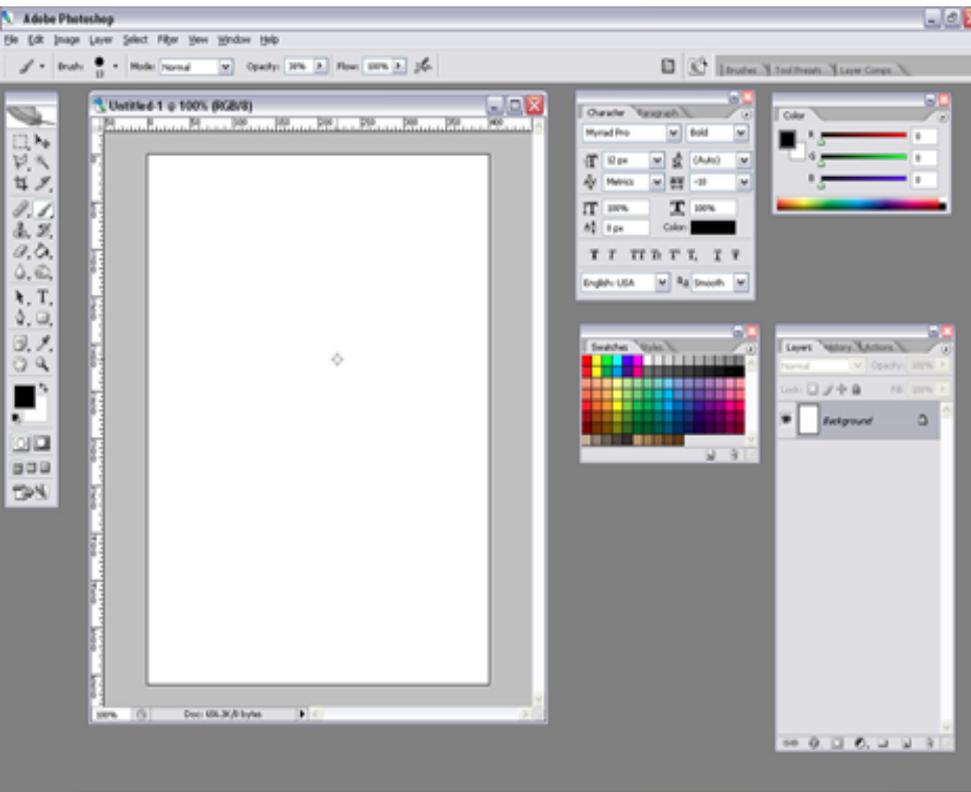# #1. CAPTURES UNCOMPRESSED OR LOSSLESS CONTENT
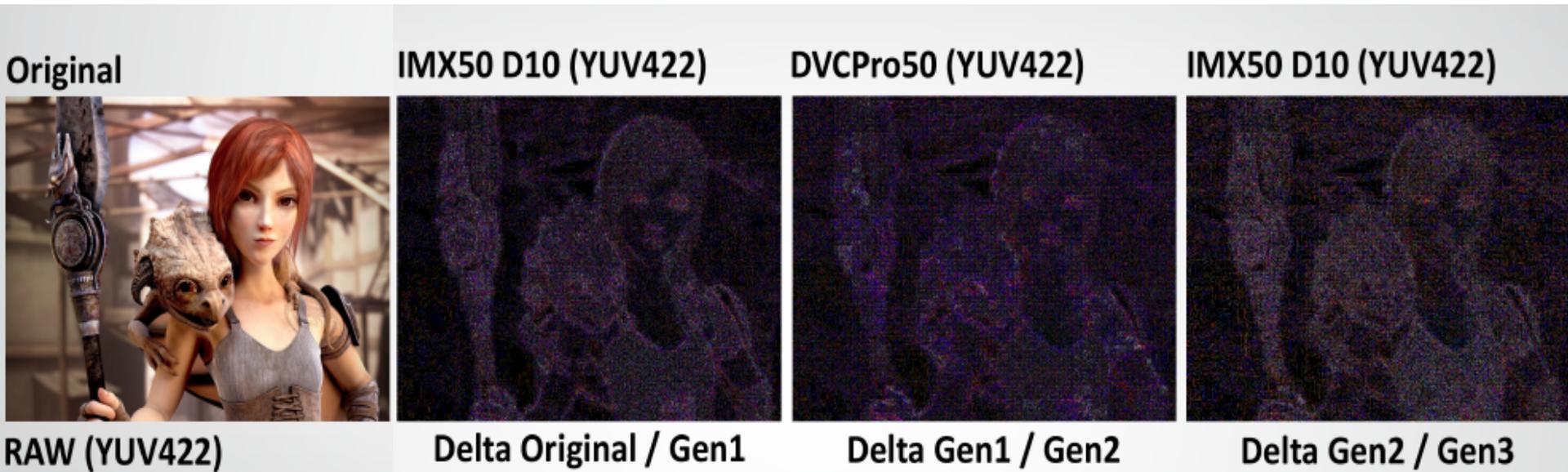
# #2. FREE / LIBRE

- The standard is adopted and will be **maintained by a not-for-profit organization**, and its ongoing development occurs on the basis of an open decision-making procedure available to all interested parties (consensus or majority decision etc.).

- The standard has been published and the standard specification document is **available either freely or at a nominal charge**. It must be permissible to all to copy, distribute and use it for no fee or at a nominal fee.

- The **intellectual property** - i.e. patents possibly present - of (parts of) the standard is made irrevocably **available on a royalty-free basis**.

- There are no **constraints on the re-use** of the standard

*European Interoperability Framework for Pan-European eGovernment Service (version 1.0 2004)*

# #3. WELL DOCUMENTED / STANDARD

# #4. ADOPTED BY USERS & SERVICE PROVIDERS

# WHAT DID THE 'EXPERTS' SAY?

| | AUDIO\|VISUAL | | | TEXT | IMAGE |
|---|---|---|---|---|---|
| | broadcast | film | Consumer | | |
| **PREFORMA stakeholders** | MPEG-IMX (MXF/MPEG2) <br> XDCAM HD422 (MXF/MPEG4) | DPX <br> DCP (MXF/JPEG2000) | MOV/MPEG2 <br> AVI/MPEG2 <br> MPEG/MPEG2 <br> MPEG/MPEG4 -AVC | PDF 1.4 <br> PDF/A1 | TIFF 6.0 <br> JPEG <br> JPEG2000 <br> RAW |
| **Industry standards** | AS-07 (MXF/MPEG2) (MXF/JPEG2000) <br> FIMS (MXF/MPEG2) | DCDM (TIFF 6.0) <br> DCP (MXF/JPEG2000) <br> IMF (MXF/MPEG4) | MPEG-AF | PDF | JPEG2000 <br> TIFF |
| **Developers** | MKV/FFV1 <br> OGG/Dirac | PNG | WebM/VP8 <br> OGG/Theora | PDF/A1 <br> PDF/A3 <br> PDF/A3 | PNG |

# WHAT DID THE GUIDELINES SAY?

| #1. uncompressed/ lossless? | AUDIO\|VISUAL | | | TEXT | IMAGE |
|---|---|---|---|---|---|
| | broadcast | film | Consumer | | |
| **PREFORMA stakeholders** | MPEG-IMX (MXF/MPEG2) XDCAM HD422 (MXF/MPEG4) | DPX DCP (MXF/JPEG2000) | MOV/MPEG2 AVI/MPEG2 MPEG/MPEG2 MPEG/MPEG4 -AVC | PDF 1.4 PDF/A1 | TIFF 6.0 JPEG JPEG2000 RAW |
| **Industry standards** | AS\|07 (MXF/MPEG2) (MXF/JPEG2000) FIMS (MXF/MPEG2) | DCDM (TIFF 6.0) DCP (MXF/JPEG2000) IMF (MXF/MPEG4) | MPEG-AF | PDF | JPEG2000 TIFF |
| **Developers** | MKV/FFV1 OGG/Dirac | PNG | WebM/VP8 OGG/Theora | PDF/A1 PDF/A3 PDF/A3 | PNG |

# WHAT DID OUR (SWEDISH) LEGAL ADVISER SAY?

| #2. open standard? | AUDIO\|VISUAL | | | TEXT | IMAGE |
|---|---|---|---|---|---|
| | broadcast | film | Consumer | | |
| **PREFORMA stakeholders** | MPEG-IMX (MXF/MPEG2) XDCAM HD422 (MXF/MPEG4) | DPX DCP (MXF/JPEG2000) | MOV/MPEG2 AVI/MPEG2 MPEG/MPEG2 MPEG/MPEG4 -AVC | PDF 1.4 PDF/A1 | TIFF 6.0 JPEG JPEG2000 RAW |
| **Industry standards** | AS\|07 (MXF/MPEG2) (MXF/JPEG2000) FIMS (MXF/MPEG2) | DCDM (TIFF 6.0) DCP (MXF/JPEG2000) IMF (MXF/MPEG4) | MPEG-AF | PDF | JPEG2000 TIFF |
| **Developers** | MKV/FFV1 OGG/Dirac | PNG | WebM/VP8 OGG/Theora | PDF/A1 PDF/A3 PDF/A3 | PNG |

# WHAT DID OTHER COLLECTIONS SAY?

| #4. adoption? | AUDIO\|VISUAL | | | TEXT | IMAGE |
|---|---|---|---|---|---|
| | broadcast | film | Consumer | | |
| **PREFORMA stakeholders** | MPEG-IMX (MXF/MPEG2) XDCAM HD422 (MXF/MPEG4) | DPX DCP (MXF/JPEG2000) | MOV/MPEG2 AVI/MPEG2 MPEG/MPEG2 MPEG/MPEG4-AVC | PDF 1.4 PDF/A1 | TIFF 6.0 JPEG JPEG2000 RAW |
| **Industry standards** | AS\|07 (MXF/MPEG2) (MXF/JPEG2000) FIMS (MXF/MPEG2) | DCDM (TIFF 6.0) DCP (MXF/JPEG2000) IMF (MXF/MPEG4) | MPEG-AF | PDF | JPEG2000 TIFF |
| **Developers** | MKV/FFV1 OGG/Dirac | PNG | WebM/VP8 OGG/Theora | PDF/A1 PDF/A3 PDF/A3 | PNG |

# WHAT WE BECAME AWARE OF…

- standard formats <> multiple implementations
- consensus > based on short term implementation needs, not on long term preservation needs
- too often we copy format guidelines/policies without checking if they are applicable
- (technically) open $\neq$ (free/libre) open format
- technically appropriate formats > little support/adoption

# WHAT WE HAVE CHOSEN (EVENTUALLY…)

TEXT (strengthen the consensus)

- ISO 32000-1:2008 (PDF 1.7)
- ISO 19005-1:2005 (PDF/A-1)
- ISO 19005-2:2011 (PDF/A-2)
- ISO 19005-3:2012 (PDF/A-3)

IMAGE (improve the consensus)

- ISO 12234-2:2001 (TIFF/EP)
- ISO 12369:2004 (TIFF/IT)

AUDIOVISUAL (uncharted territory...)

- OGG / MKV
- **FFV1** / **Dirac** / **ISO 15444-1 (JPEG2000 core coding system)**
- LPCM

# ABOUT THE AV SPECIFICATIONS

- #1. Lossless or uncompressed
- #2. Open format (cf.EIFv1)

MUST requirement!

- #3. Documentation
- #4. Adoption

Which design has the best idea to solve the shortcomings?

- JPEG2000
  - RAND licensing file format
  - permissive licensing software
  - documentation

- FFV1
  - standardisation
  - documentation

# WHERE ARE WE NOW?

- https://githu.com/verapdf (PDF/A)

- https://github.com/EasyInnovaSL/DPFManager (TIFF)

- https://github.com/MediaArea/MediaConch (MKV|FFV1|LPCM)

- second public release: end October 2015 (Quarterly)

- open source workshop, Stockholm: 7 April 2016

# FFV1 for preservation?

**Here be dragons...**

# FFV1 for preservation.
**Here be ~~dragons~~… questions!**

- FFV1: A different approach?

- Features?

- Lossless codec performance comparison?

- Licensing?

- Sustainability?

- Accessibility & Interoperability?

# A different approach?
**FFV1 is different by design**

- Format is Free/Libre/Open since day 1

- Reference implementation =
  Production implementation

- Open/transparent development =
  Features defined by the actual users

- Only one mode: Lossless.

# Only one mode?
**Reconsidering the "All-in-One" approach...**

# The All-in-One approach
## Jack of all trades - Master of none?

# The All-in-One approach
**Or: The right tool for the right job?**

# The All-in-One approach
## File formats that can do everything?

# The All-in-One approach
## Or: Selecting what you really need?

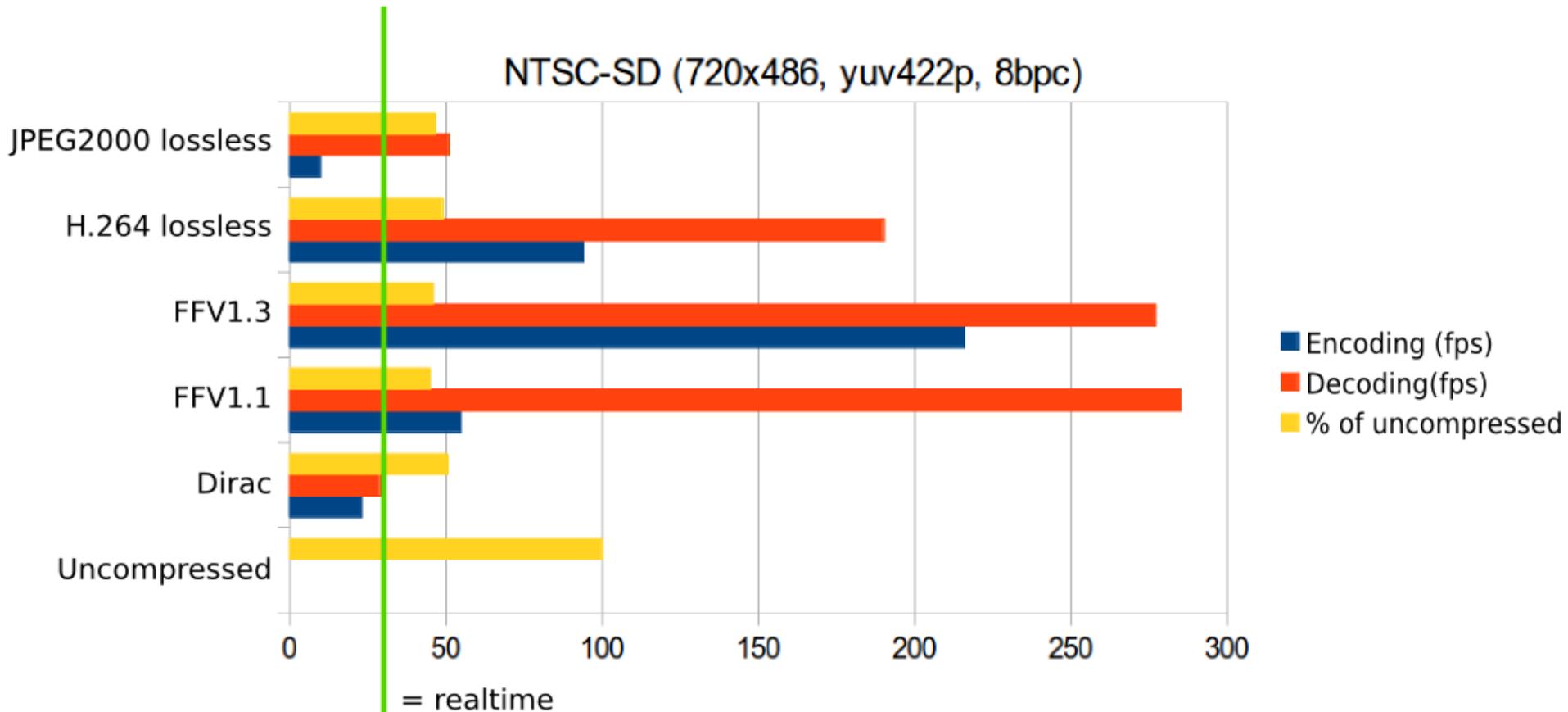# FFV1 Features

**Current status**

# FFV1 Features

**Current status**

- Color spaces: YUV, RGB, Gray, (XYZ) (linear / logarithmic)
- Bits-per-component (bpc):
  - YUV: up to 16 bits
  - RGB: up to 14 bits
- Multithreading
- CRC checksums (error correction/concealment)
- Alpha channel
- Aspect ratio

# Performance comparison
## Speed and compression ratio



NTSC-SD (720x486, yuv422p, 8bpc)

Legend:
- Encoding (fps)
- Decoding(fps)
- % of uncompressed

= realtime

# FFV1 License

**Free/Libre Software: Why it matters?**

# License

**Free Software / Open Source: It's a feature!**

- GNU Lesser General Public License (LGPL)
- Free/Libre Software licensing =
- 4 Freedoms:
  1. **Use**
  2. **Study**
  3. **Share**
  4. **Improve**

# License
**Free Software / Open Source: It's a feature!**

- Archiving institutions:
  - Common interests
  - Common challenges
  - Common solutions!

# License
**Free Software / Open Source: It's a feature!**

- License freedom #4: **Improve**
  - Mediathek initiated & paid improvement:
  - Multithreading (=faster)
  - CRCs
    (=error detection/correction/concealment)
- Others contributed too:
  - Dave Rice
  - NOA
  - and many more...

# License
**Free Software / Open Source: It's a feature!**

- Free choice of developers / support
- Direct communication with original authors
- Improvements committed back to FFmpeg

- **Result:**
  - FFV1 version 3 (FFV1.3)
  - Optimized for long-term preservation

# Sustainability

**Why the license matters...**

# Sustainability

## Archiving the source code?

Free Software Freedoms + Copy of source code:

- It's like archiving your recorder/replayer

- ...including schematics and construction parts.

= **No artificial restrictions** preventing reading/writing the format or **adapting it to future conditions**.

# Accessibility

**Shouldn't great stuff be available for everyone?**

# Accessibility

**FFV1 comes with FFmpeg/LibAV by default**

- FFmpeg / LibAV is Free Software

- Providing command line tools and program libraries (to include it in other applications)

- Free Software = Implementation available for everyone: Users, developers, vendors

- No black box: Problems can be solved, rather than hacking workarounds

- Collaboration encouraged

# Accessibility
## Benefits of FFV1 in FFmpeg / LibAV

- for Users:
    - Works with over hundreds of applications that are built on FFmpeg/LibAV libraries.
- for Vendors / Developers:
    - Easy to add FFV1 support by using FFmpeg/LibAV libraries.
    - No proprietary development necessary.
- for Archives:
    - Commandline/libraries allow integration in automated (existing) workflows

# Interoperability
## by design

- Reference implementation = Free Software
- Reference = Production implementation = FFmpeg
- No gap between paper (theory) and code (practice)
- No proprietary (closed) implementation-variations (black box)
- No vendor dependency

**Almost done...**
Link collection coming up

# Link collection
## For further reading

- FFV1 on Wikipedia:
  https://en.wikipedia.org/wiki/FFV1

- Projects using FFmpeg:
  https://trac.ffmpeg.org/wiki/Projects

- Video archiving FAQ / codec comparisons / FFV1 quick start guide:
  http://download.das-werkstatt.com/pb/mthk/info/

- Digital A/V media tech-basics explained:
  http://xiph.org/video/vid1.shtml

# Standardization of FFV1

*"I am very disappointed that jpeg2000 is a 'standard' and ffv1 is not :/ "*

"...fear of the format being changed arbitrarily."

"...we are strongly drawn to 'capital-S' (official) standards"

**"Any idea why FFV1 isn't more commonly acknowledged as an archival option?"**

"Cons: Codec is not yet standardized on paper."

"I think there has been an understandable tendency to hold back on making decisions on preservation options, pending an authoritative declaration of legitimate target formats."

# Standardization

- Development status prior to PREFORMA
- Impetus for standardization
  - Credibility
  - Transparency
  - Sustainability
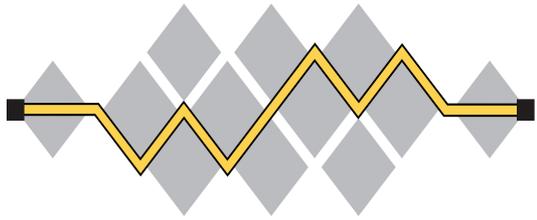- Advantages of standardization
  - ...see above.

# Finding An Appropriate Standards Body

Criteria

- non-profit organization
- open source + no pay wall
- recognized standards body
- support of developer communities

ISO? SMPTE? Or…

# Internet Engineering Task Force

- History
- Structure
- Real Time Applications & Infrastructure Area
  - Dispatch Working Group
- How does FFV1 fit in?

# IETF MISSION STATEMENT

**Mission Statement**

The mission of the IETF is to make the Internet work better by producing high quality, relevant technical documents that influence the way people design, use, and manage the Internet.

The IETF will pursue this mission in adherence to the following cardinal principles:

**Open process** - any interested person can participate in the work, know what is being decided, and make his or her voice heard on the issue. Part of this principle is our commitment to making our documents, our WG mailing lists, our attendance lists, and our meeting minutes publicly available on the Internet.

**Technical competence** - the issues on which the IETF produces its documents are issues where the IETF has the competence needed to speak to them, and that the IETF is willing to listen to technically competent input from any source. Technical competence also means that we expect IETF output to be designed to sound network engineering principles - this is also often referred to as"engineering quality".

**Volunteer Core** - our participants and our leadership are people who come to the IETF because they want to do work that furthers the IETF's mission of "making the Internet work better".

**Rough consensus and running code** - We make standards based on the combined engineering judgement of our participants and our real-world experience in implementing and deploying our specifications.

**Protocol ownership** - when the IETF takes ownership of a protocol or function, it accepts the responsibility for all aspects of the protocol, even though some aspects may rarely or never be seen on the Internet. Conversely, when the IETF is not responsible for a protocol or function, it does not attempt to exert control over it, even though it may at times touch or affect the Internet.

https://www.ietf.org/about/mission.html

# IETF Process

- Community support and participation
- Practitioners and developers
- Requests for comment
- Review by consensus
- Transparent, democratic process

| Proposal | Working group formation | Drafts, more drafts, even more drafts | Publication |
|----------|-------------------------|---------------------------------------|-------------|

**Review/ discussion**

# CELLAR!

## Codec Encoding for LossLess Archiving and Realtime transmission

- FFV1 + Matroska + FLAC
- Community and developer support
- Next steps: IESG review and feedback

Charter:

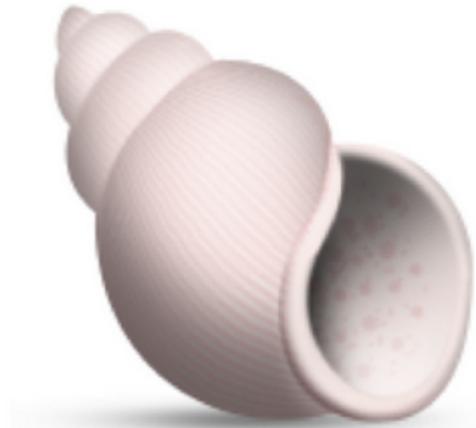https://datatracker.ietf.org/doc/charter-ietf-cellar/

# IETF LINKS

Mailing list:

https://www.ietf.org/mailman/listinfo/dispatch

Charter:

https://datatracker.ietf.org/doc/charter-ietf-cellar/

Working group name: CELLAR

CELLAR mailing list to come! Follow DISPATCH until then…

# MediaConch

An open source audiovisual
file conformance checker

# MediaArea Team

Project Leaders:

Jérôme Martinez (Digital Media Specialist)

Dave Rice (Archivist)

Other Members:

Guillaume Roques (Developer)

Florent Tribouilloy (Developer)

Ashley Blewer (Archivist)
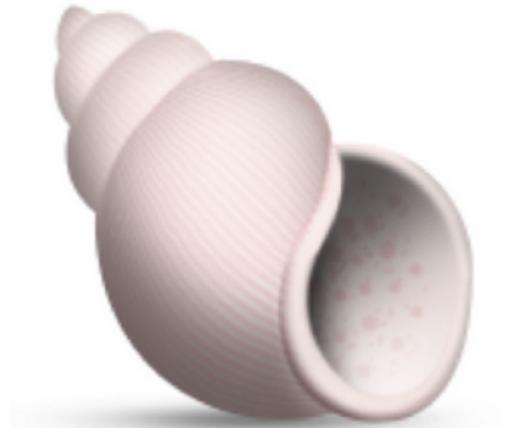
Tessa Fallon (Archivist)

Erik Piil (Archivist)

# FFV1! What is that?

## PCM

# Media Conch Components

# Check files

✕ Close all results

## Check by file upload

### Policy, Schematron or XSL

**Choose a policy**

Choose a policy ▲▼

Or upload a Schematron (.sch) or a XSL (.xsl) file

Choose File   No file chosen

**File (max 128M)**

Choose File   No file chosen

Check file

## Check online files

## Check server files

## Result for Test2_2.mkv ✕

This file is: **NOT VALID**

Policy report

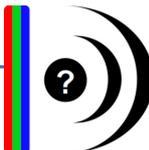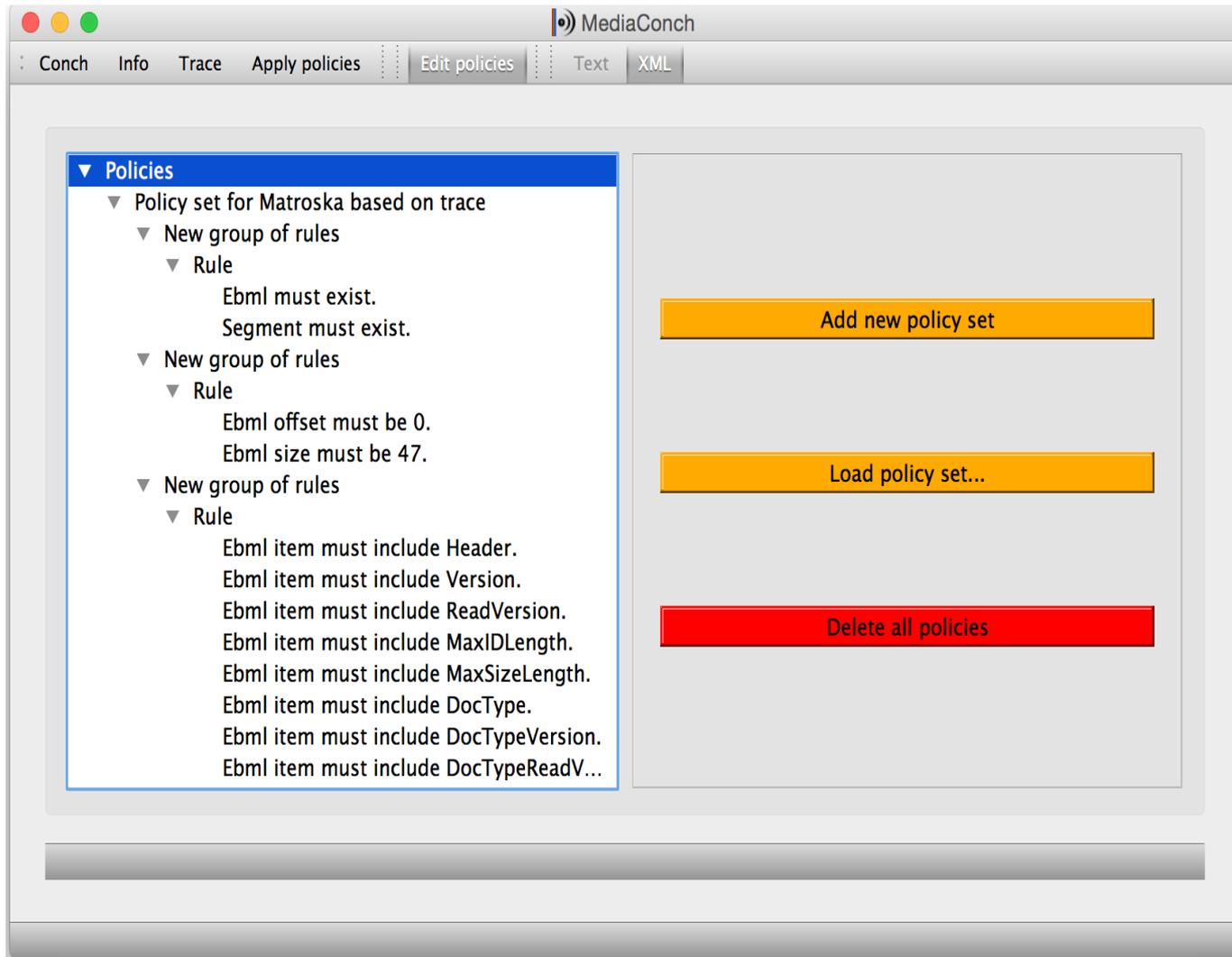Implementation report

MediaInfo XML

MediaInfo Trace

# MediaTrace sample output

<block offset="1920" name="Television information" size="128">    <data offset="1920" name="SMPTE time code">4294967295</data>    <data offset="1924" name="SMPTE user bits">4294967295</data>    <data offset="1928" name="Interlace" moreinfo="2:1 interlace">255</data>    <data offset="1929" name="Field number">255</data>    <data offset="1930" name="Video signal standard" moreinfo="Undefined">0</data>    <data offset="1931" name="Zero">255</data>    <data offset="1932" name="Horizontal sampling rate (Hz)">0.000</data>    <data offset="1936" name="Vertical sampling rate (Hz)">0.000</data>    <data offset="1940" name="Temporal sampling rate or frame rate (Hz)">0.000</data>    <data offset="1944" name="Time offset from sync to first pixel (ms)">0.000</data>    <data offset="1948" name="Gamma">0.000</data>    <data offset="1952" name="Black level code value">0.000</data>    <data offset="1956" name="Black gain">0.000</data>    <data offset="1960" name="Breakpoint">0.000</data>    <data offset="1964" name="Reference white level code value">0.000</data>    <data offset="1968" name="Integration time (s)">0.000</data>    <data offset="1972" name="Reserved for future use">(76 bytes)</data></block>

# Downloads

## 15.08 Release Notes

### CLI and GUI

This release of MediaConch is a bug fixes release. Focused on the trace feature.

### Online

MediaConchOnline features the ability to create an account; to create, import, and develop checker policies; and to validate files.

### Historical Release Notes

15.07

15.06

15.05

### Snapshots

You can test ongoing developments for MediaConch by downloading our daily builds.

# All downloads

| Operating system | Link |
|---|---|
| Windows | Official releases |
| Mac | Official releases |
| debian | Official releases |
| ubuntu | Official releases |
| Linux Mint | Official releases |
| redhat | Official releases |
| CentOS | Official releases |
| fedora | Official releases |
| openSUSE | Official releases |
| SUSE | official releases |
| from source (for other distributions or/and CPU) | official releases |

# A business case for open source development…

- Open source development is part of MediaArea's business plan
- Integration of open source tools into flagship product, MediaInfo
- Combination of subscriptions and paid punctual support, such as bug corrections and new feature requests

# To be considered:

What are critical factors in considering FFV1 for adoption?

How can we ensure that development will continue?

How should we frame the conversation between archives and vendors?

**Questions?**