CHECK YOUR STANDARD: PREFORMA AND MEDIACONCH

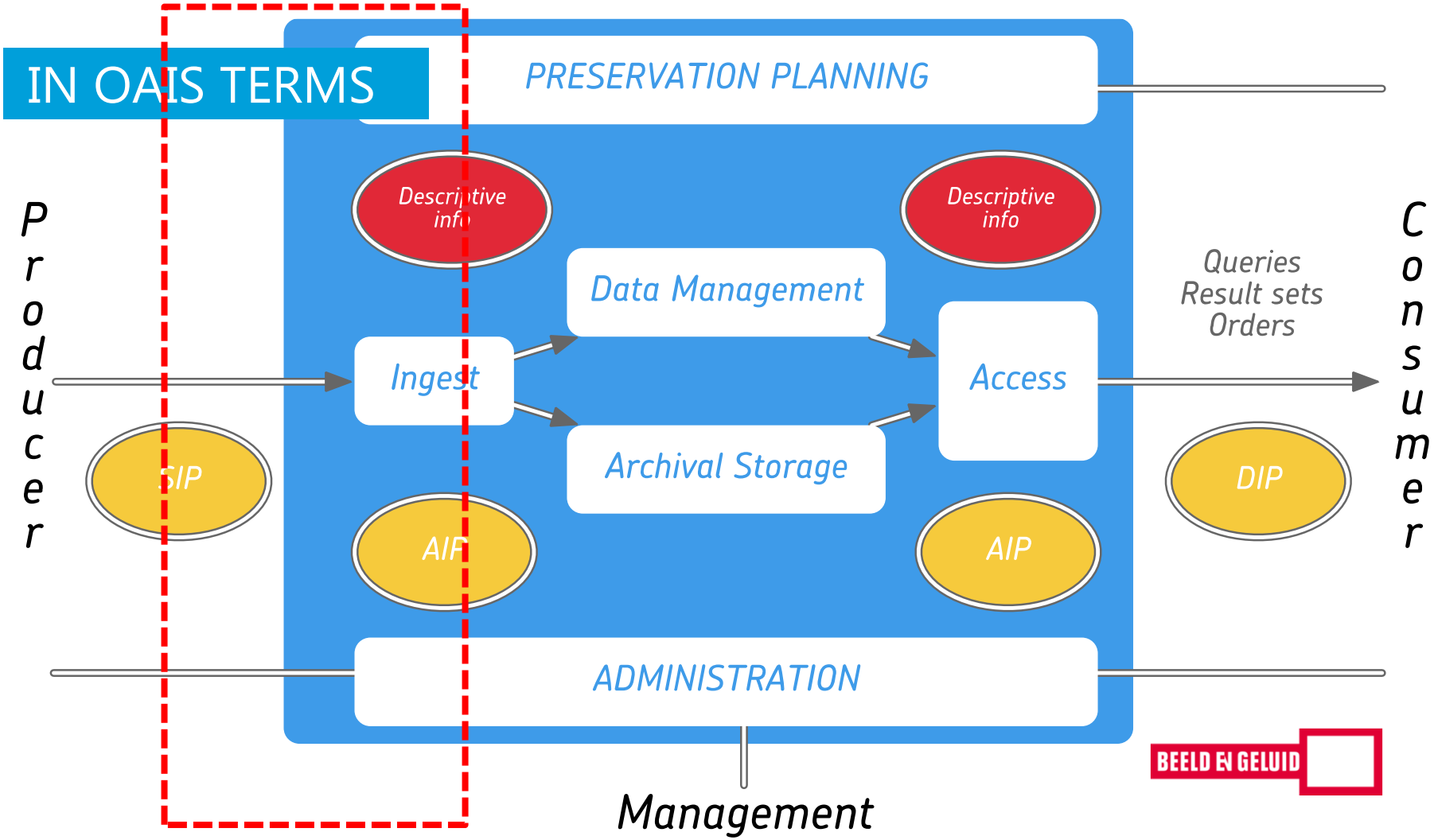
Erwin Verbruggen
Netherlands Institute for Sound and Vision
JTS 2016 – NMS Singapore


LABS SOUND AND VISION
HOME
EXAMPLES
ABOUT
CONTACT
LINKS
Sound and Vision LABS
ME EXPERIMENT WITH OUR
ATIONS, DATASETS AND APIS
DIG IN!
PRESTOCENTRE.EU
OPENIMAGES.EU
BIT.LY/FLOSSINVENTORY
APPLICATIONS
DATASETS
APIS
LABS.BEELDENGELUID.NL

CONFORMANCE CHECKER IN PRESERVATION

- File has been produced according to the specifications of a standard file format
  - File matches the acceptance criteria for long-term preservation by a memory institution
- Reports which properties deviate from the standard specification and acceptance criteria
  - In human and machine readable format
- Performs automated fixes for simple deviations in the metadata of the preservation file.

BEELD EN GELUID


IN OAIS TERMS
PRESERVATION PLANNING
Descriptive info
Descriptive info
Data Management
Ingest
Access
Archival Storage
SIP
AIP
AIP
DIP
Queries
Result sets
Orders
Producer
Consumer
ADMINISTRATION
Management
BEELD EN GELUID

PREFORMA

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

Slide with thanks to Claudio Prandoni

BEELD EN GELUID

REFERENCE STANDARD

For PREFORMA:

1. Format can capture content in an uncompressed or mathematically lossless encoding, and retain as many original properties as possible (ie. bit-depth, resolution, …)
2. Format is Free/Libre (cf EIFv1)
3. Format is well documented
4. Format has been adopted

BEELD EN GELUID

Open standard?	AUDIOVISUAL			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
Open standards	MKV/FFV1 OGG/Dirac	PNG	WebM/VP8 OGG/Theora	PDF/A1 PDF/A3 PDF/A3	PNG
PREFORMA	MKV (?) | OGG | JPEG2000 (?) | FFV1 | Dirac | LPCM (?)			PDF/A	TIFF 6.0 (?)

Slide with thanks to Bert Lemmens

REFERENCE STANDARD

For PREFORMA:

1. Format can capture content in an uncompressed or mathematically lossless encoding, and retain as many original properties as possible (ie. bit-depth, resolution, …)
2. Format is Free/Libre (cf EIFv1)
3. Format is well documented
4. Format has been adopted

BEELD EN GELUID

"FFV1/MKV ARE GAINING TRACTION IN A CORNER OF THE COMMUNITY - WHETHER FOR SIZE, HIPNESS, ANTI-ESTABLISHMENT, SIMPLER-THAN-MXF, "FREE", OR OPEN OR WHATEVER."

George Blood, AMIA List, 2015-12-03

See Session 7: *Review and comparison of FFV1 versus other lossless video codecs* - **Peter Bubestinger**

BEELD EN GELUID

CHALLENGE BRIEF

- Develop an open source conformance checker that:
  - checks if a file complies with standard specifications
  - checks if a file complies with the acceptance criteria of the memory institution
  - reports back to human and software agents
  - performs simple fixes
- Establish an ecosystem around an open source reference implementation that:
  - generates useful feedback for those who control software
  - advances improvement of the standard specification
  - advances development of new business cases for managing preservation files

BEELD EN GELUID

Slide with thanks to Claudio Prandoni

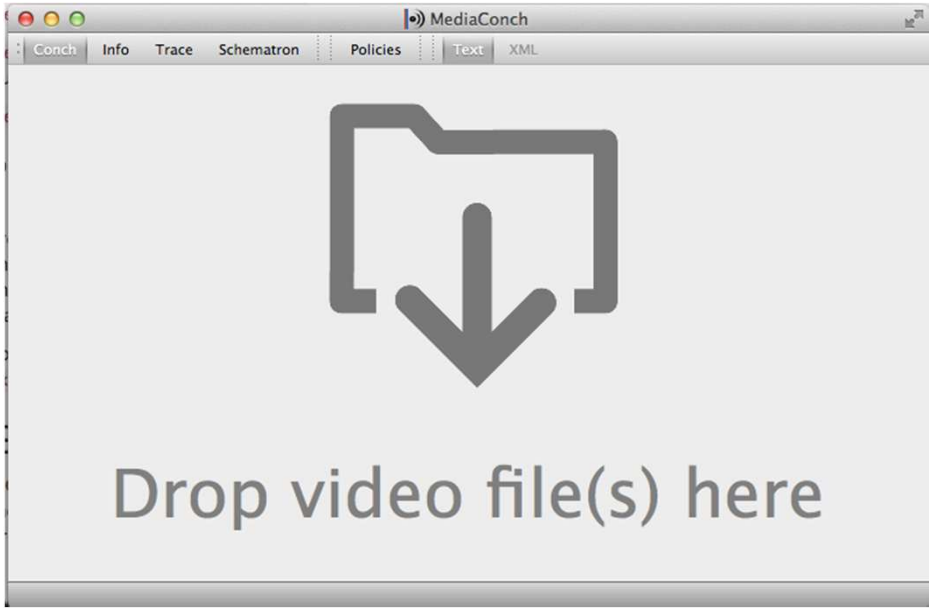
THE PREFORMA OUTCOMES

- VeraPDF Consortium:
  - Open Preservation Foundation
  - PDF Association
  - Digital Preservation Coalition
- PDF/A Validation

- TI/A Standard Initiative:
  - Digital Humanities Lab / University of Basel
  - Agents Research Lab / University of Girona
  - Easy Innova
- DPF Manager

MEDIACONCH

MediaConch



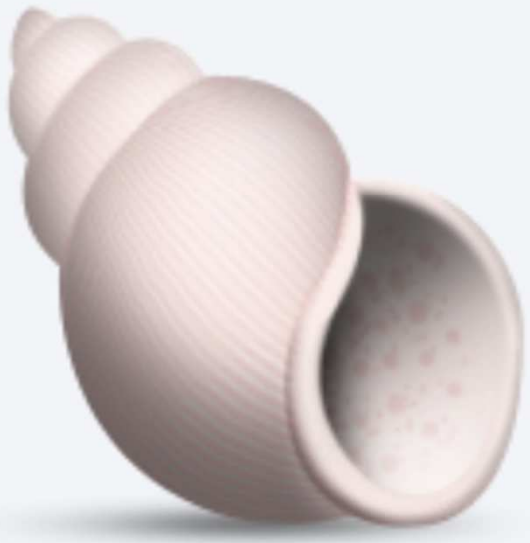
BEELD EN GELUID

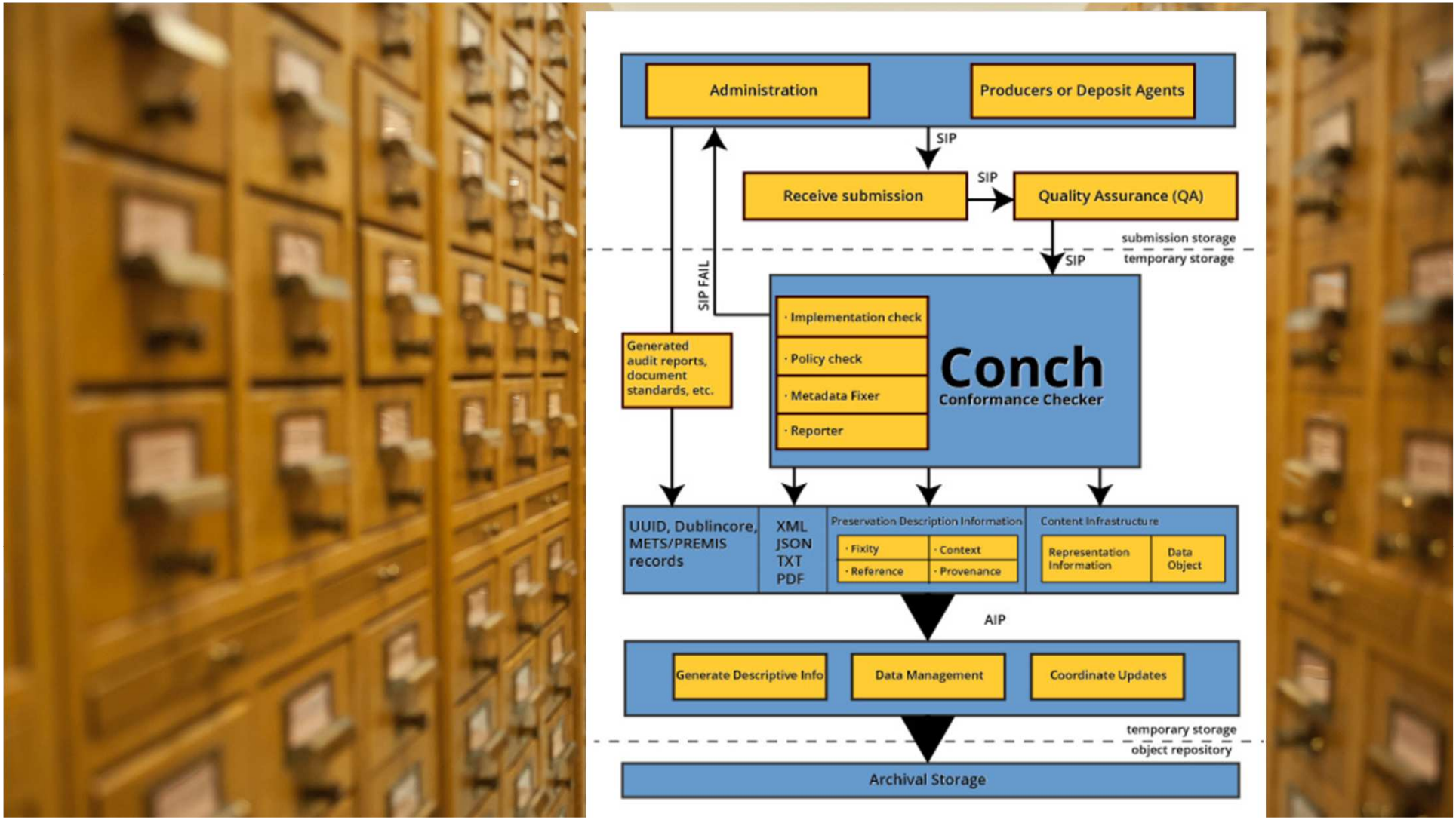
MediaConch

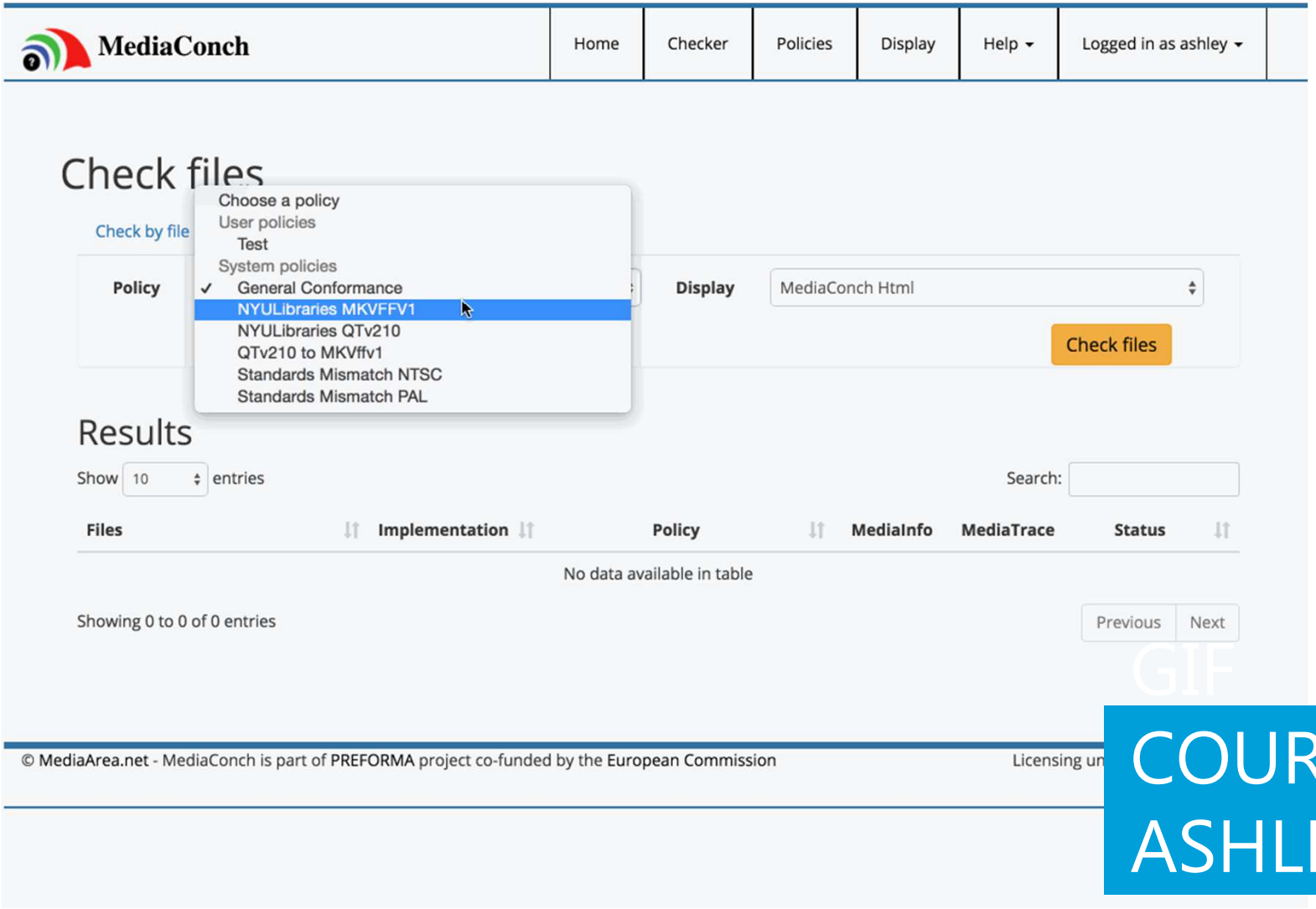
Home | About | Checker | Policies | Login | Register

MediaConch

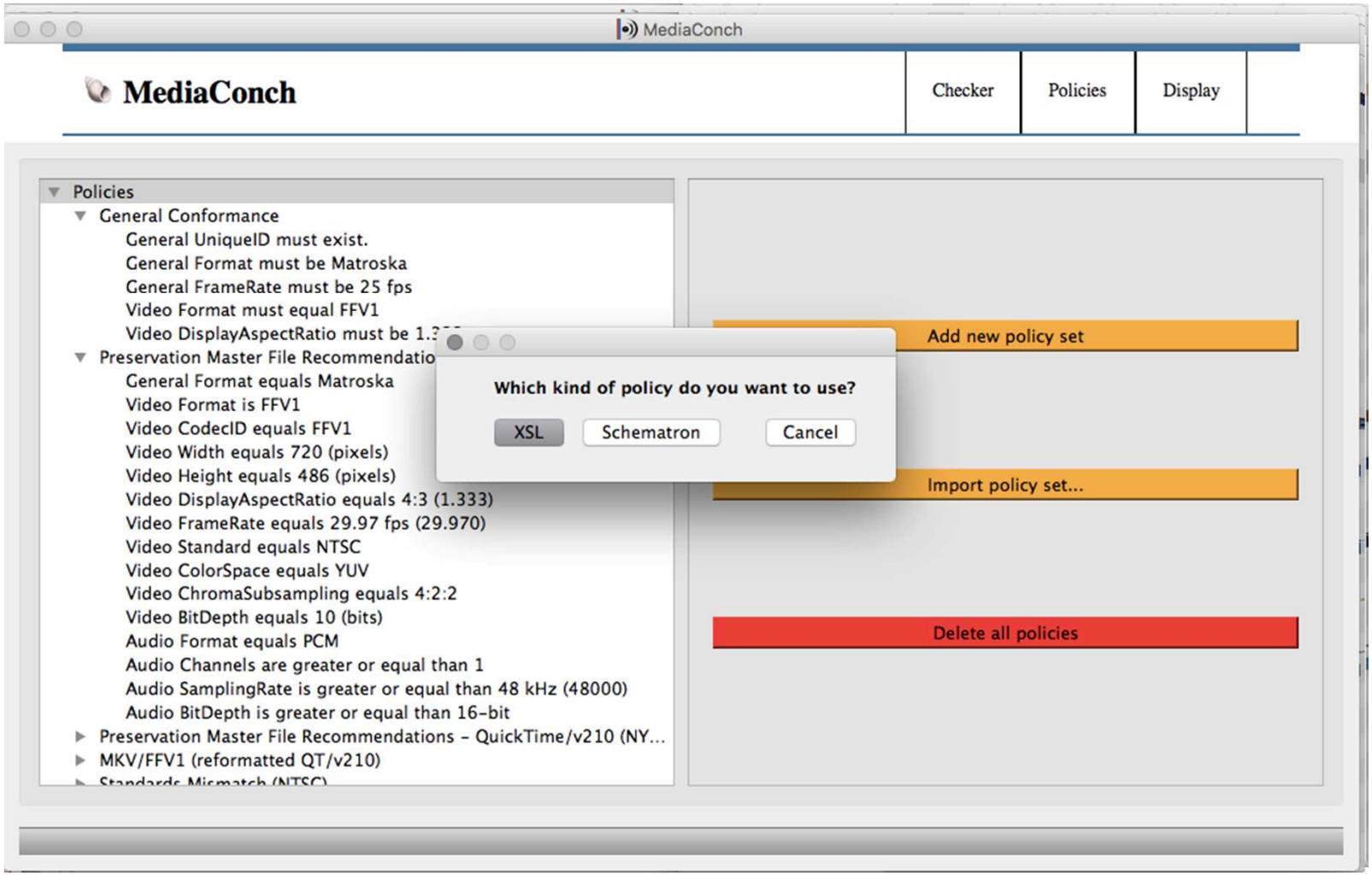
MediaConch is an extensible, open source software project consisting of an implementation checker, policy checker, reporter, and fixer that targets preservation-level audiovisual files (specifically Matroska, Linear Pulse Code Modulation (LPCM) and FF Video Codec 1 (FFV1)) for use in memory institutions, providing detailed and batch-level conformance checking via an adaptable and flexible application program interface accessible by the command line, a graphical user interface, or a web-based shell. Conch is currently being developed by the MediaArea team, notable for the creation of open source media checker software, MediaInfo. Furthermore, the MediaArea team is dedicated to the further development of the standardization of the Matroska and FFV1 formats to ensure their longevity as a recommended digital preservation file format.




Administration
Producers or Deposit Agents
SIP
Receive submission
SIP
Quality Assurance (QA)
submission storage
temporary storage
SIP
SIP FAIL
· Implementation check
· Policy check
· Metadata Fixer
· Reporter
Conch
Conformance Checker
Generated audit reports, document standards, etc.
UUID, Dublincore, METS/PREMIS records
XML
JSON
TXT
PDF
Preservation Description Information
· Fixity
· Context
· Reference
· Provenance
Content Infrastructure
Representation Information
Data Object
AIP
Generate Descriptive Info
Data Management
Coordinate Updates
temporary storage
object repository
Archival Storage



COURTESY ASHLEY


MediaConch
MediaConch
Checker
Policies
Display
Policies
General Conformance
General UniqueID must exist.
General Format must be Matroska
General FrameRate must be 25 fps
Video Format must equal FFV1
Video DisplayAspectRatio must be 1.3
Preservation Master File Recommendatio
General Format equals Matroska
Video Format is FFV1
Video CodecID equals FFV1
Video Width equals 720 (pixels)
Video Height equals 486 (pixels)
Video DisplayAspectRatio equals 4:3 (1.333)
Video FrameRate equals 29.97 fps (29.970)
Video Standard equals NTSC
Video ColorSpace equals YUV
Video ChromaSubsampling equals 4:2:2
Video BitDepth equals 10 (bits)
Audio Format equals PCM
Audio Channels are greater or equal than 1
Audio SamplingRate is greater or equal than 48 kHz (48000)
Audio BitDepth is greater or equal than 16-bit
Preservation Master File Recommendations - QuickTime/v210 (NY...
MKV/FFV1 (reformatted QT/v210)
Which kind of policy do you want to use?
XSL
Schematron
Cancel
Add new policy set
Import policy set...
Delete all policies

CELLAR

Codec Encoding for LossLess Archiving and Realtime transmission

Project Advisors

CHAIRS:
TESSA FALLON
TIM TERRIBERRY

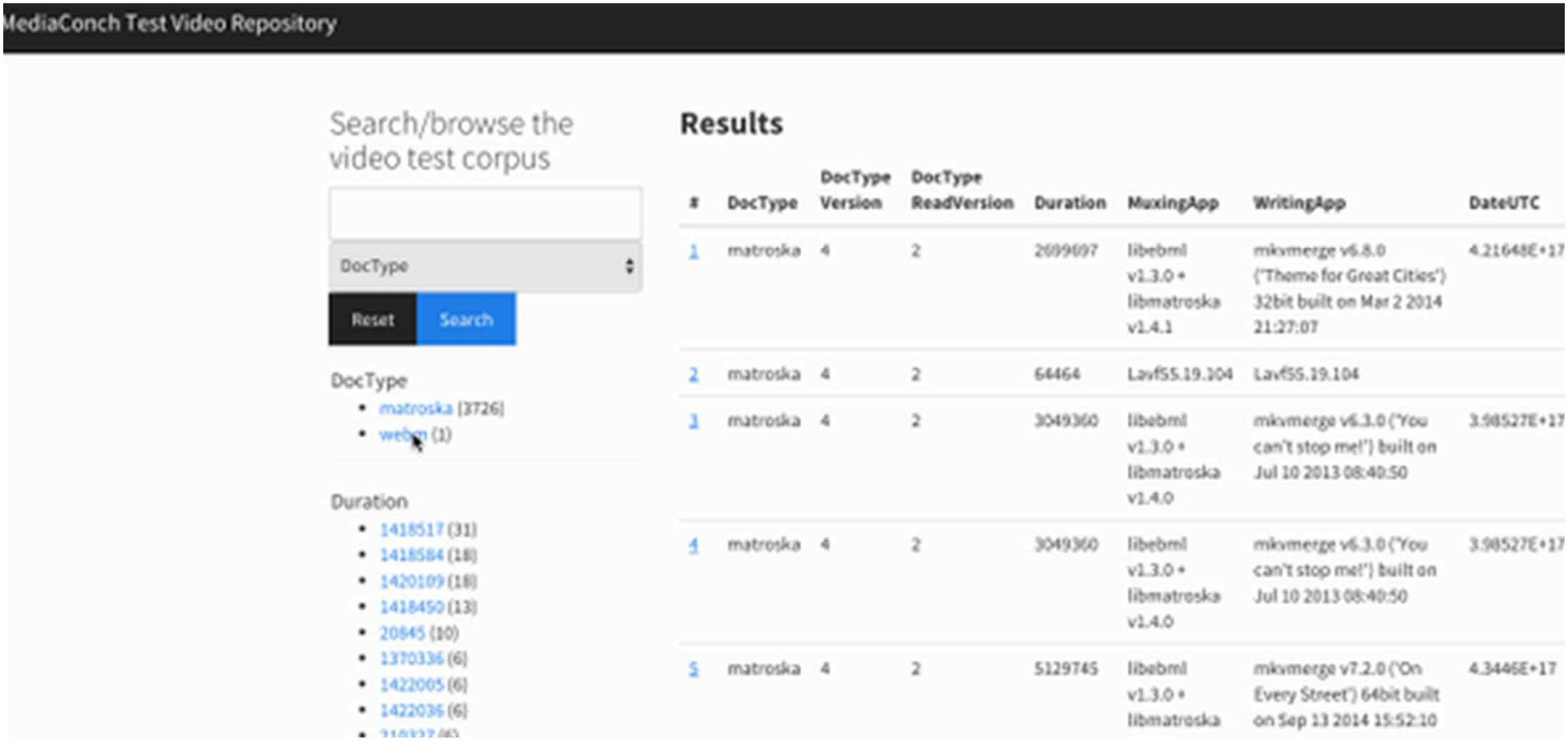
BEELD EN GELUID

A LOOK FORWARD

- MediaConch development:
  - –Asynchronous mode
  - – UI Improvements
  - – Statistics
  - – Standardization
  - – More conformance tests
- PREFORMA Runs until 2017 – user base in place
  - –Including at Sound and Vision?

BEELD EN GELUID

TEST CORPUS


MediaConch Test Video Repository
Search/browse the video test corpus
DocType
Reset
Search
DocType
matroska [3726]
webm (1)
Duration
1418517 (31)
1418584 (18)
1420109 (18)
1418450 (13)
20845 (10)
1370336 (6)
1422005 (6)
1422036 (6)
Results
DocType DocType Version DocType ReadVersion Duration MuxingApp WritingApp DateUTC
1 matroska 4 2 2699697 libebml v1.3.0 + libmatroska v1.4.1 mkvmerge v6.8.0 ('Theme for Great Cities') 32bit built on Mar 2 2014 21:27:07 4.21648E+17
2 matroska 4 2 64464 Lavf55.19.104 Lavf55.19.104
3 matroska 4 2 3049360 libebml v1.3.0 + libmatroska v1.4.0 mkvmerge v6.3.0 ('You can't stop me!') built on Jul 10 2013 08:40:50 3.98527E+17
4 matroska 4 2 3049360 libebml v1.3.0 + libmatroska v1.4.0 mkvmerge v6.3.0 ('You can't stop me!') built on Jul 10 2013 08:40:50 3.98527E+17
5 matroska 4 2 5129745 libebml v1.3.0 + libmatroska mkvmerge v7.2.0 ('On Every Street') 64bit built on Sep 13 2014 15:52:10 4.3446E+17

WHAT YOU CAN DO

- CELLAR participation
- QC for all interfaces
- Bug tracking, reporting
- Documentation
- Your weird video errors

BIT.LY/MEDIACONCHDOCS

GITHUB.COM/MEDIAAREA/MEDIACONCH

BEELD EN GELUID


WORKSHOP
PREFORMA
SERVING THE CULTURAL HERITAGE
OPEN SOURCE PRESERVATION WORKSHOP
7 APRIL 2016
Kungliga Biblioteket
Humlegården, 102 41 Stockholm
opensourceworkshop.preforma-project.eu


THANK
YOU!
PREFORMA
EVERBRUGGEN@BEELDENGELUID.NL
@ERWINVERB | @BENGLABS
MEDIAAREA.NET/MEDIACONCH
@MEDIACONCH
Kungliga Biblioteket
PREFORMA-PROJECT.EU