

Conch

Conformance checking for media files

Jerome Martinez and David Rice
MediaArea.net

About MediaArea

"As a National Library incorporating the National Screen and Sound Archive of Wales, we have to preserve digital audiovisual material in perpetuity. Part of this work is characterising AV files and extracting technical metadata. We found no better tool at this job than MediaInfo, and the support and response from MediaArea SARL has always been excellent."

Key players: Conch Team

Jerome Martinez, Lead/Developer

Dave Rice, Lead/Project Management

Guillaume Roques, Developer

Tessa Fallon, Research & Development: Standardization

Ashley Blewer, Research & Development: Design

Erik Piil, Research & Development: File Formats

Key players: Advisors

- Format advisors
- Institutional advisors
- Standardization advisors

The PREFORMA Challenge

Conformance Checker

Status of formats

- MKV
- FFV1
- LPCM

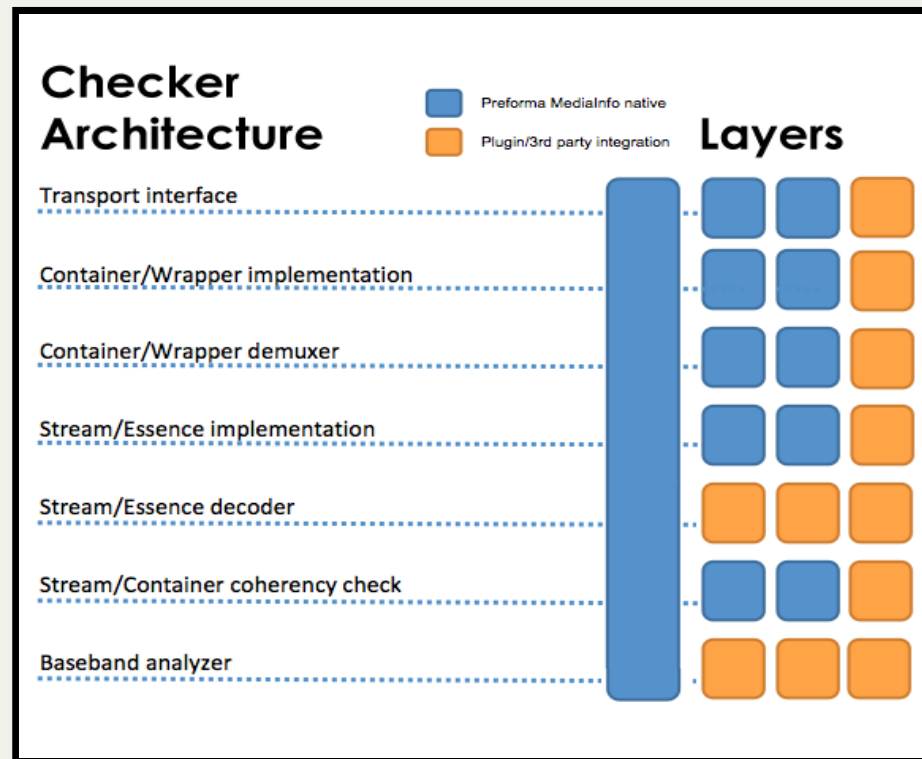
Standardization

- IETF

Implementation

How to label and test a bitstream

Several layers: transport, container, stream



Implementation checker

- MKV: EBMLMaxSizeLength limit, DocTypeVersion, CRC-32...
- FFV1: Missing header, colorspace_type, bits_per_raw_sample, crc_parity...
- Container/Stream cross-check: PCM block duration and bitrate, MKV/FFV1 width...

Policy checker

- Must be BWF: container format must be WAV, audio format must be PCM, sampleRate must be 48000 or 44100, ...
- Custom policy check, e.g. PREFORMA: container format must be MKV, video format must be FFV1, audio format must be PCM...
- Very detailed check: MKV version, FFV1 version and micro_version, color space, bit depth, CRC-32 presence...

Reporter

Input: analysis of the file, policies, options

Computer readable Output: XML, JSON...

Human readable Output: HTML, PDF...

Metadata Fixer

- Add CRC, coherency between container and stream
- Embed additional metadata
- Warning: not all is possible

Standardization of expression format for policy checks

MKV and Archives

User Interface Coverage

Command Line - Checker

```
Testing FFU1.mkv...

Implementation checking errors:
Matroska      | DocTypeVersion too high      | DocTypeReadVersion=4, must be lower than DocTypeVersion=2
FFU1          | CRC not valid                 | CRC is present but is not valid, frame corrupted
x-check       | Width Coherency              | Matroska width=704, FFU1 width=720

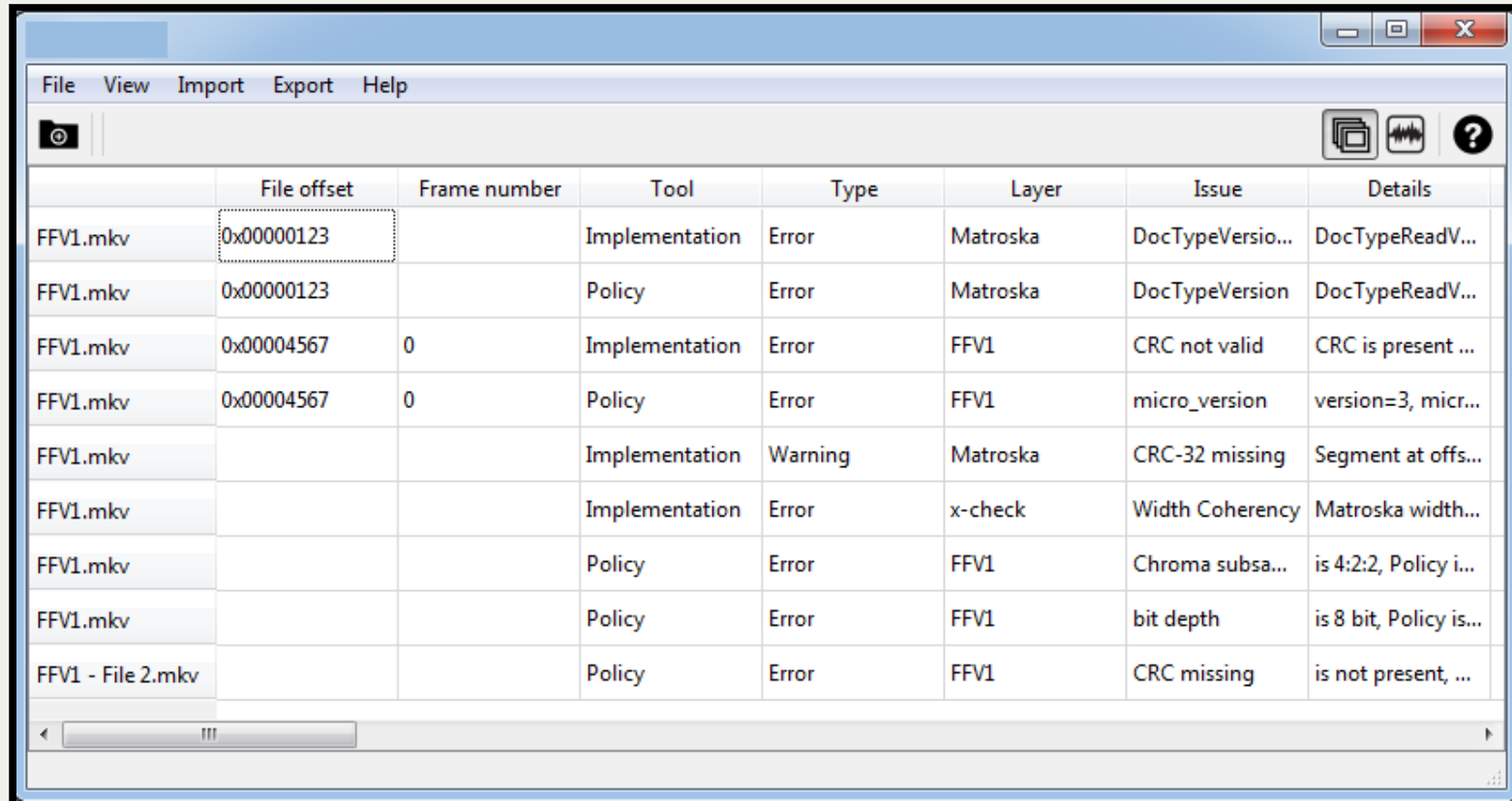
Implementation checking warnings:
Matroska      | CRC-32 missing                | Segment at offset 0x00000123 is missing mandatory CRC-32
FFU1          | micro_version                 | version=3, micro_version=3, this micro-version is pre-standard

Policy checking errors:
Matroska      | DocTypeVersion                | DocTypeReadVersion=4, Policy is DocTypeReadVersion<=3
FFU1          | Chroma subsampling            | is 4:2:2, Policy is 4:4:4
FFU1          | bit depth                     | is 8 bit, Policy is >=10
FFU1          | CRC missing                   | is not presence, Policy is presence mandatory

Total: 3 implementation errors, 2 implementation warnings, 4 policy errors
```

User Interface Coverage

Graphic User Interface - Checker



The screenshot shows a window titled "Graphic User Interface - Checker" with a menu bar (File, View, Import, Export, Help) and a toolbar. The main area contains a table with the following data:

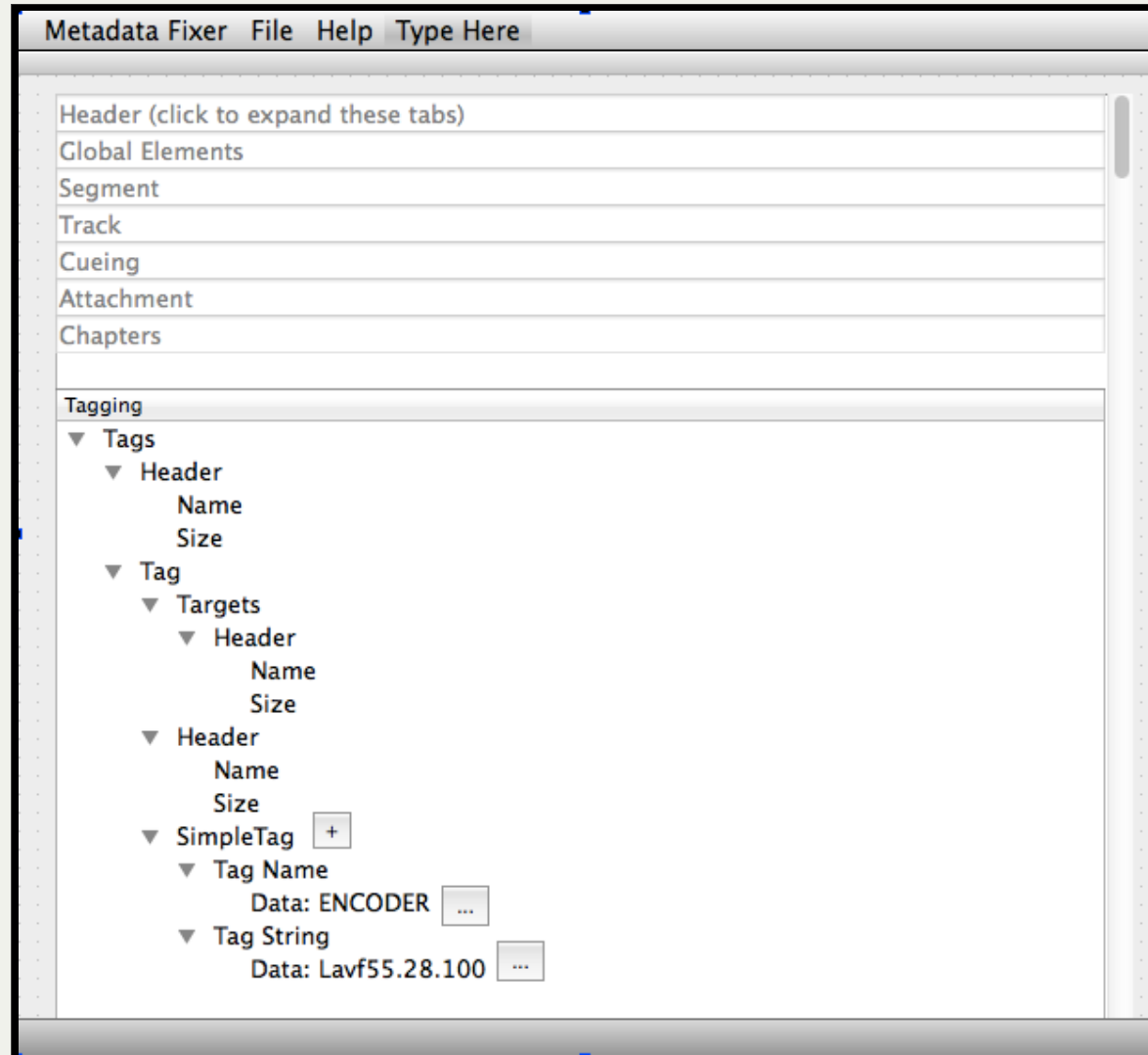
	File offset	Frame number	Tool	Type	Layer	Issue	Details
FFV1.mkv	0x00000123		Implementation	Error	Matroska	DocTypeVersio...	DocTypeReadV...
FFV1.mkv	0x00000123		Policy	Error	Matroska	DocTypeVersion	DocTypeReadV...
FFV1.mkv	0x00004567	0	Implementation	Error	FFV1	CRC not valid	CRC is present ...
FFV1.mkv	0x00004567	0	Policy	Error	FFV1	micro_version	version=3, micr...
FFV1.mkv			Implementation	Warning	Matroska	CRC-32 missing	Segment at offs...
FFV1.mkv			Implementation	Error	x-check	Width Coherency	Matroska width...
FFV1.mkv			Policy	Error	FFV1	Chroma subsa...	is 4:2:2, Policy i...
FFV1.mkv			Policy	Error	FFV1	bit depth	is 8 bit, Policy is...
FFV1 - File 2.mkv			Policy	Error	FFV1	CRC missing	is not present, ...

User Interface Coverage

Graphic User Interface - Checker Debug

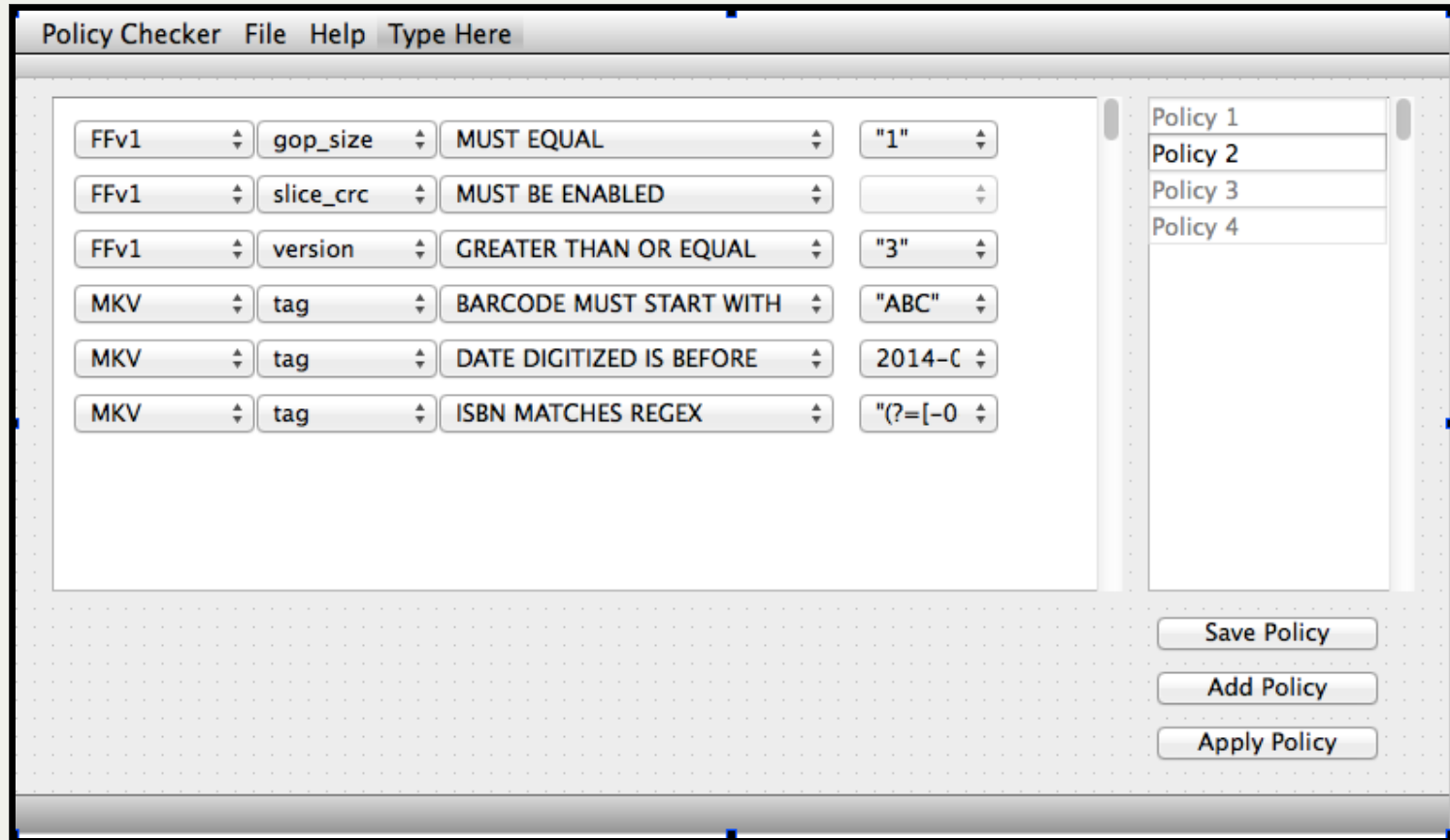
User Interface Coverage

Graphic User Interface - Fixer



User Interface Coverage

Graphic User Interface - Policy Checker



User Interface Coverage

Web Interface

Automation

Periodical checks

Detection of new files

Database side

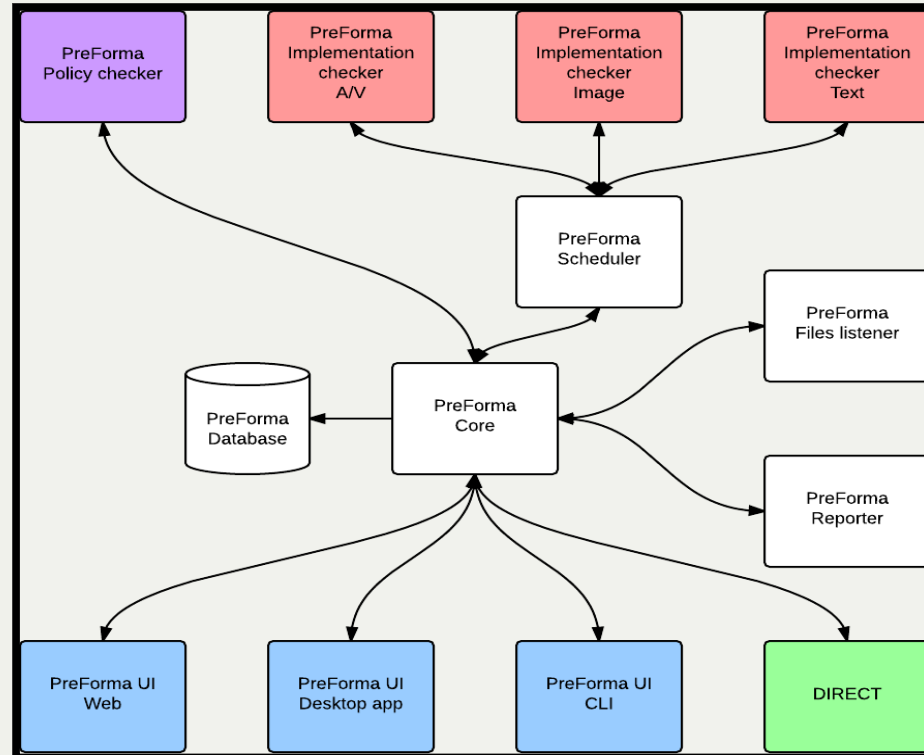
Configuration

Analysis of the files

Policies

The Core

Interaction between all other elements

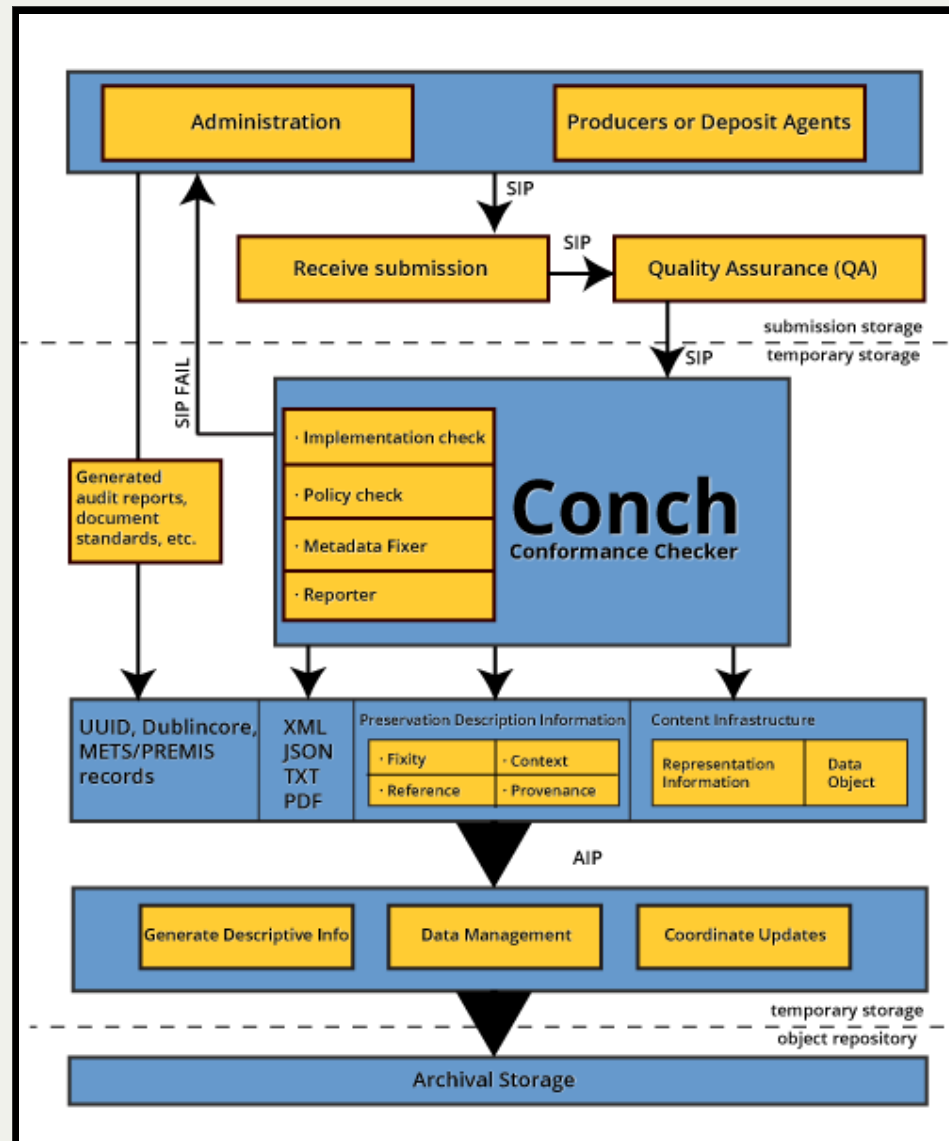


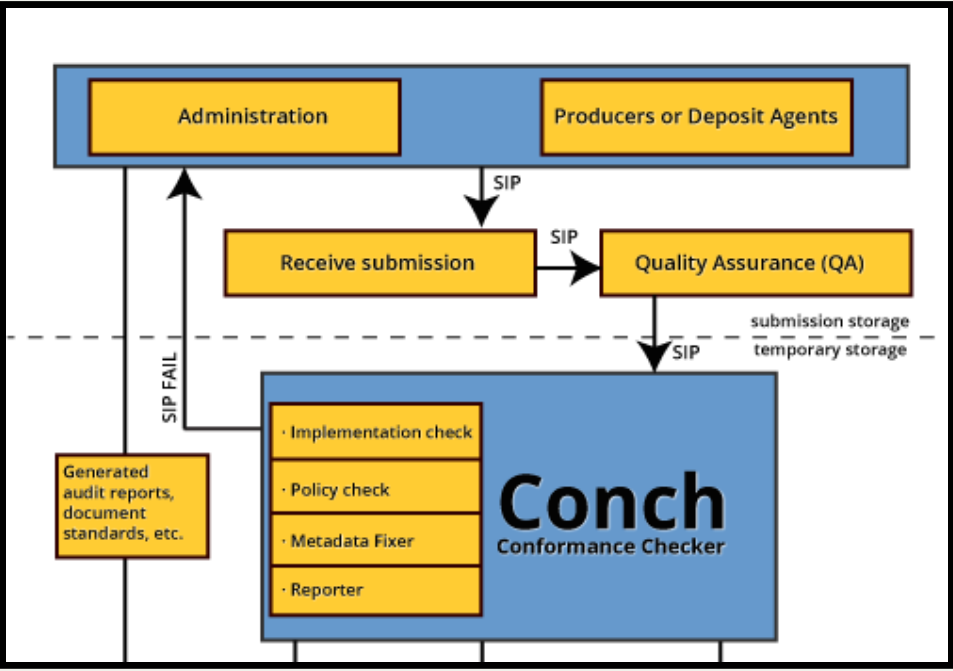
Optimization priorities

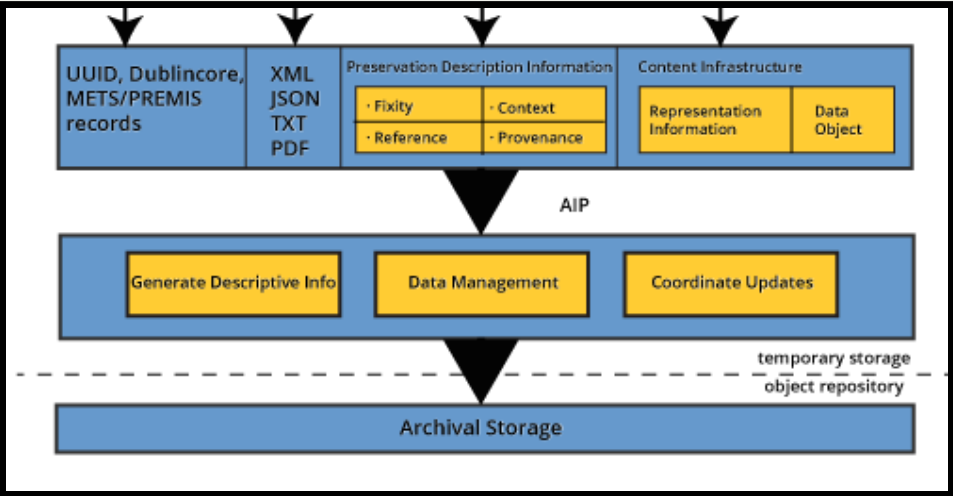
Data Exchange

OAIS

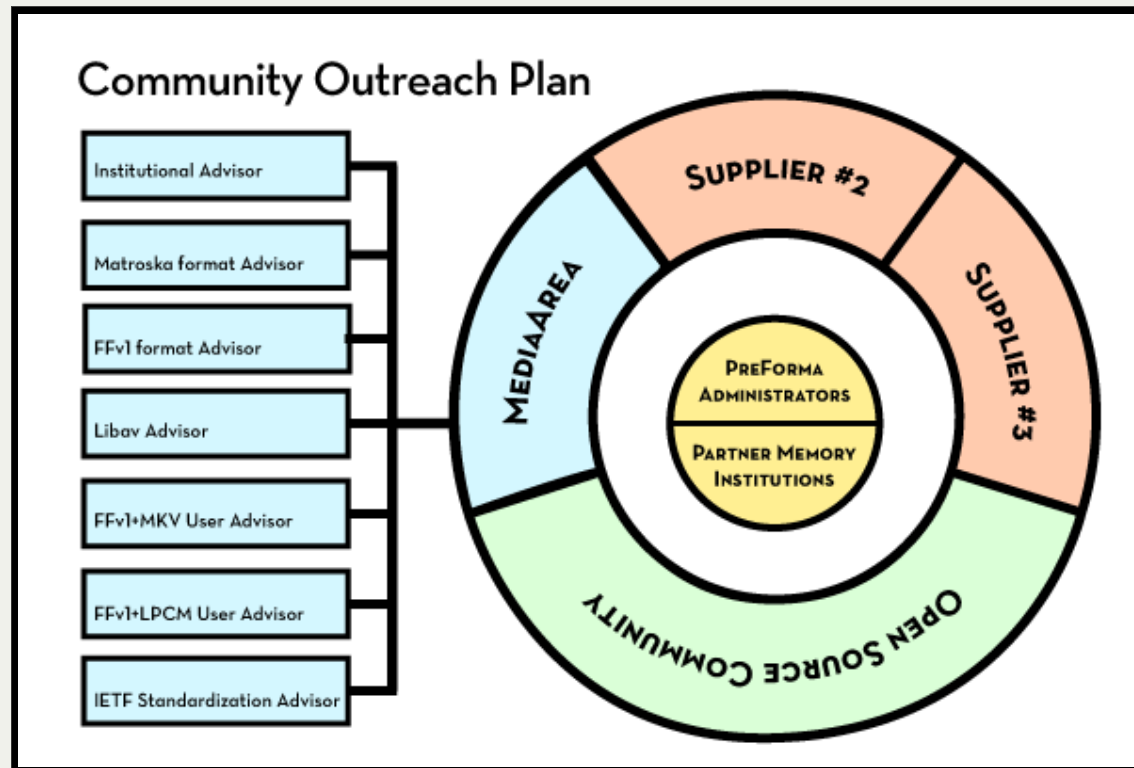
OAIS workflows







Outreach/Community



Involvement,
stakeholders,
and the open source
community

Development Approach

- Github
- Agile development
- Stipended collaboratives
- Community involvement

Development Approach

- Cross platform
- Multiple developer interfaces (bindings)
- Multiple end user interfaces

Future Plans

Support, new feature requests

MediaArea Plans

Integration of other MediaArea's products

Integration & Community

Thanks!

??? Questions ???