

1<sup>st</sup> Int. Workshop on Standards and Technologies in  
Multimedia Archives and Records (STAR),  
Lausanne, 2010-04-26/27

# **Professional Archival Application Format at ISO/IEC MPEG**

NTT Communication Science Labs.  
Noboru Harada  
([harada.noboru@lab.ntt.co.jp](mailto:harada.noboru@lab.ntt.co.jp))

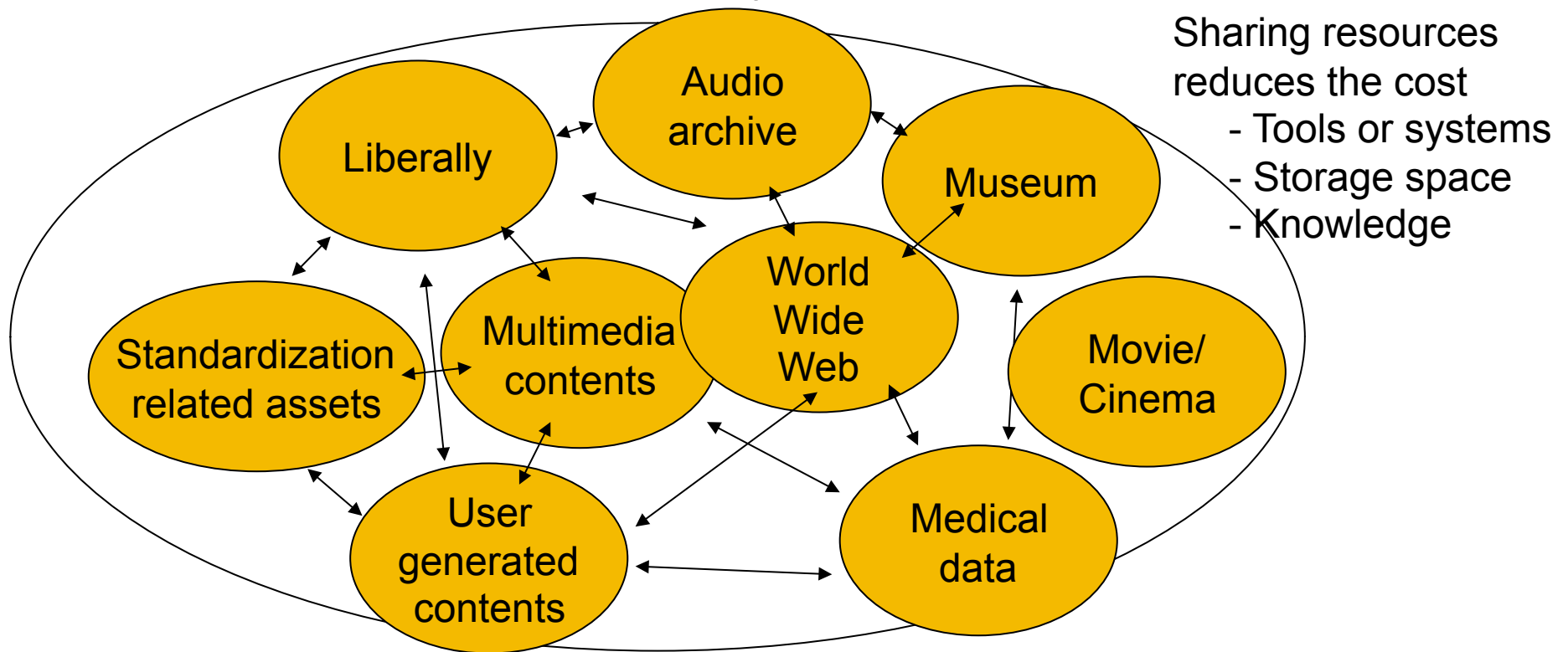
**Why do we need archiving?**  
**Why standard is so important?**

# What does “Archiving” mean?

## Everybody’s thinking of Archiving

World is getting smaller!

“Archive”s will possibly be inter-connected!



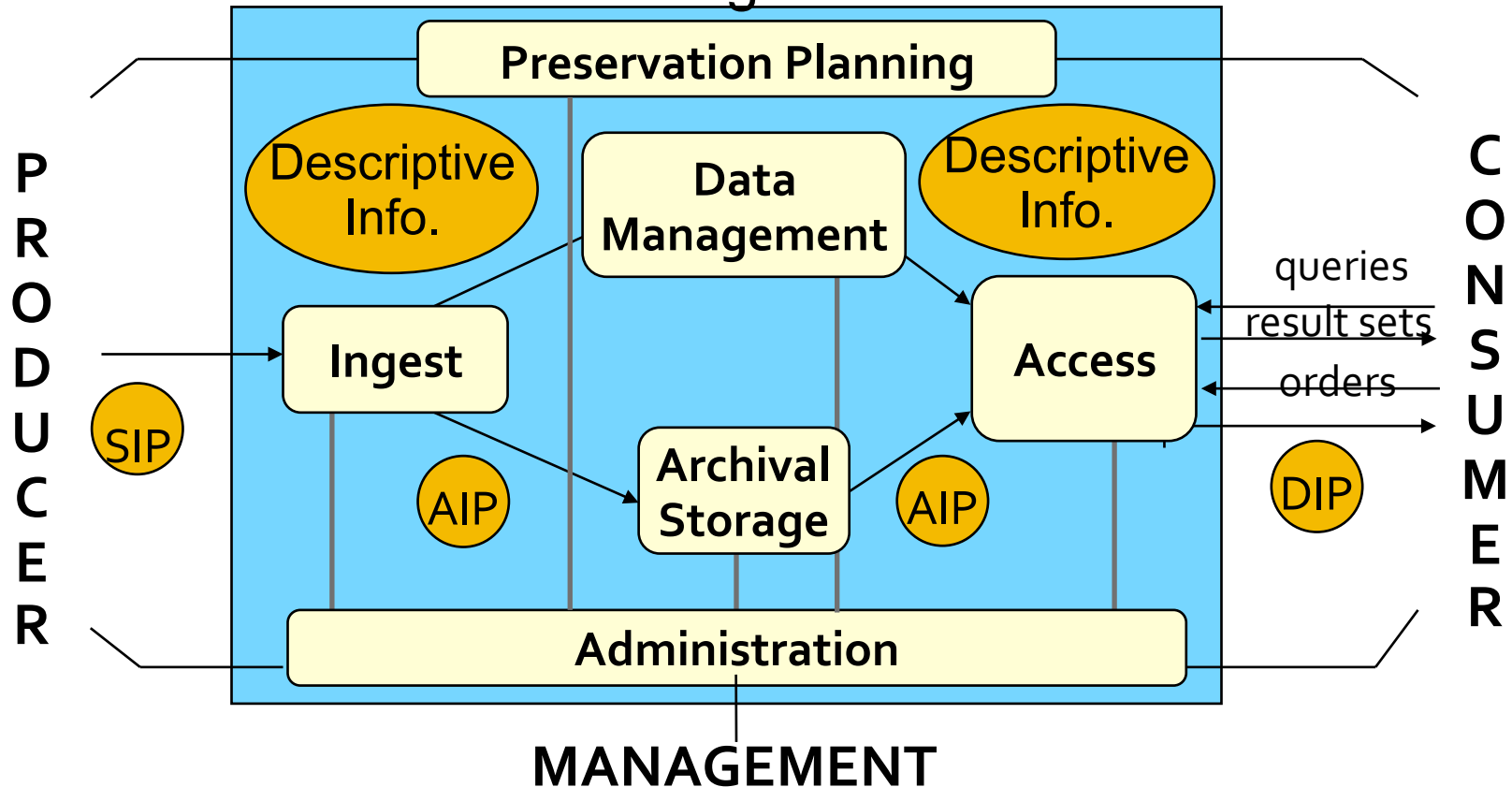
“Archive” has different meaning to each other!

# What should we archive?

- What? → Everything needed to be archived (???)
- How long? → Long enough time (???)
- Where? → ???
- How? → Format?/Media? (???)
- And then? What to do with them?  
→ Digital Black-hole???

# ISO 14721: Open Archival Information System (OAIS) Reference Model

Helps us to discuss issues using well-defined models and terms



SIP = Submission Information Package

AIP = Archival Information Package

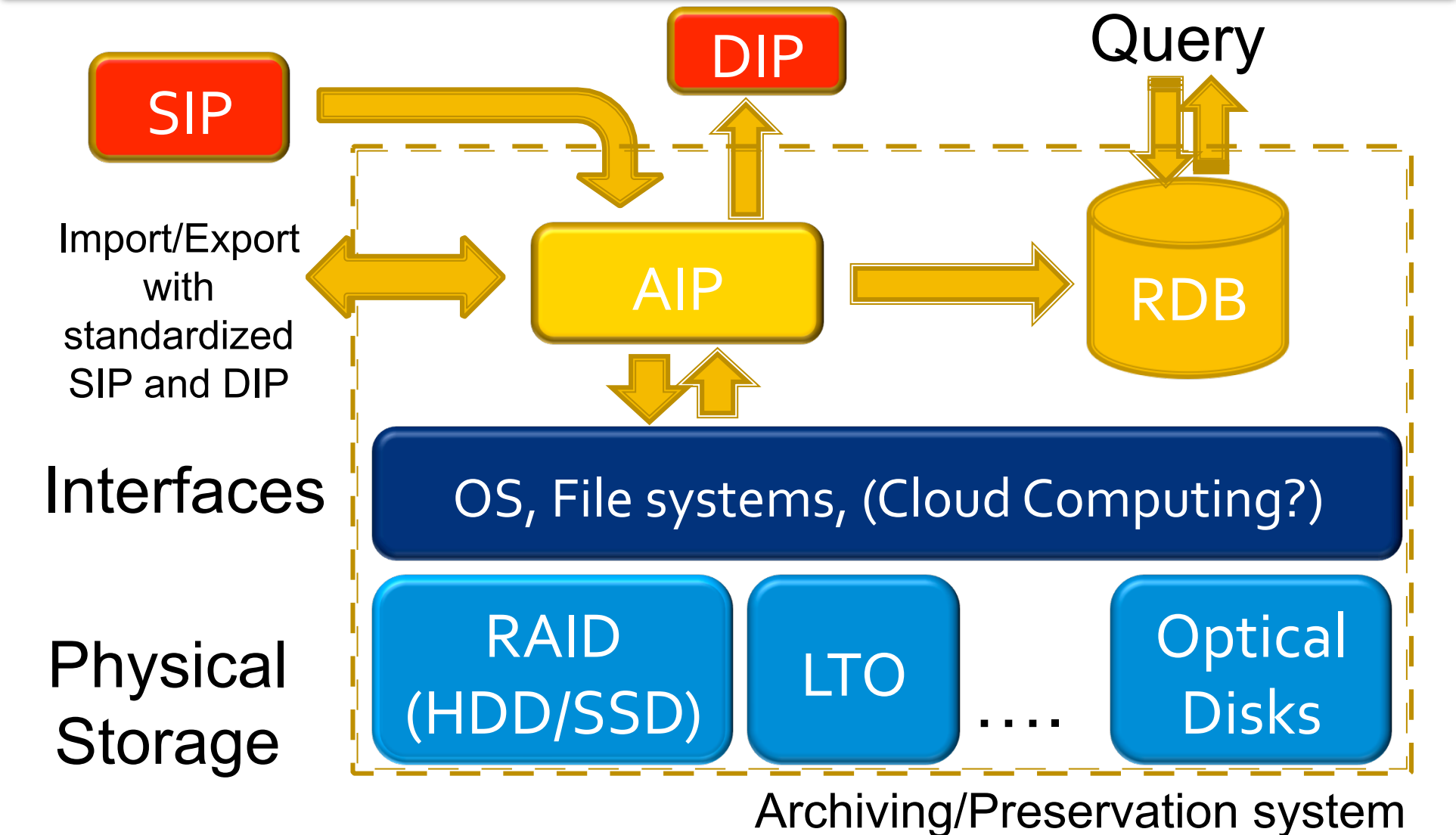
DIP = Dissemination Information Package

} Information Package

# Meta-data is the Key of interoperability (It might cause a headache/nightmare)

- Every archive has its own archiving policy
  - Archiving policy / agreement may be difficult to be generalized because of its historical / cultural reasons
  - Legacy data format should be treated as well as standardized format
  - Multiple byte character set shall be supported
- All digital media contents and meta-information should be treated in general way
- General and standardized approach
  - Open Archival Information System (OAIS) Reference Model (ISO 14721)
  - ISO/IEC 23000-6 Professional Archival Application Format (PA-AF)
- File structure based approach is the must for enabling legacy data formats

**SIP, AIP, DIP should be  
the interface point**



# What is PA-AF?

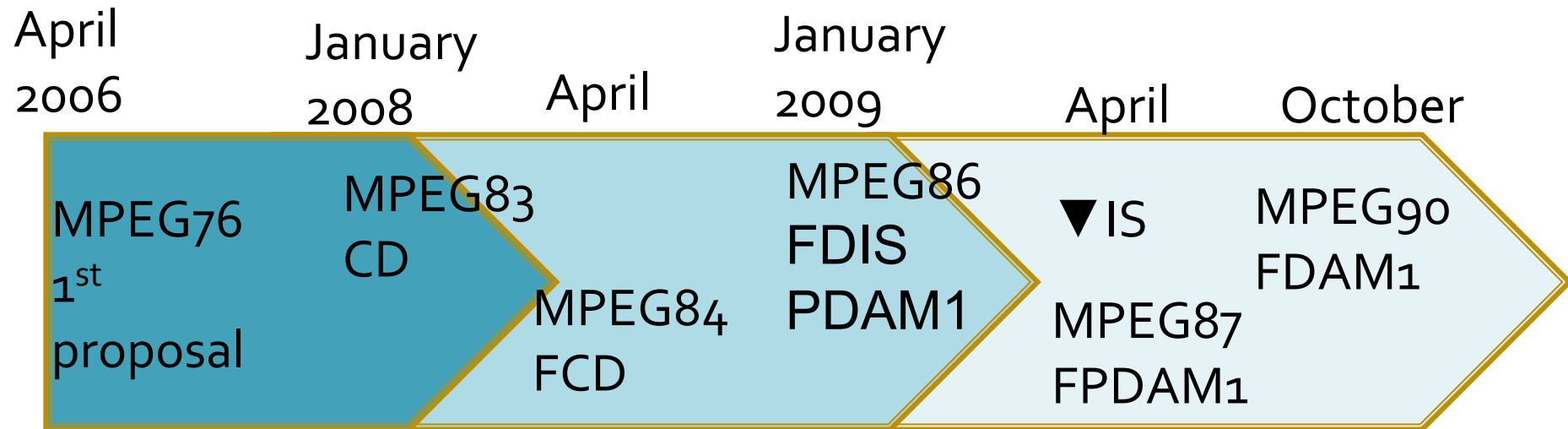
- PA-AF provides a standardized packaging format for digital files.

It specifies

- metadata format to describe the original structure of digital files archived in a PA-AF file
- necessary information to reverse the pre-processing processes applied to digital files prior to archiving them in a PA-AF file
- context information related to archived digital files
- and a file format for carriage of the metadata formats and digital files



# MPEG-A PA-AF standardization



ISO/IEC 23000-6:2009 Professional Archival Application Format

ISO/IEC 23000-6:2010/Amd1 Conformance and reference software for PA-AF

Liaison organizations:

ISO/IEC JTC1 SWG-ARM (Sub Working Group Archival and Retrieval Mechanisms)

ISO/IEC JTC1 SG-DCMP (Study Group on Digital Contents Management and Protection)

ISO/IEC JTC1 SC29/WG1 (JPEG) D-Cinema Ad hoc Group

The Society of Motion Picture and Television Engineers (SMPTE)

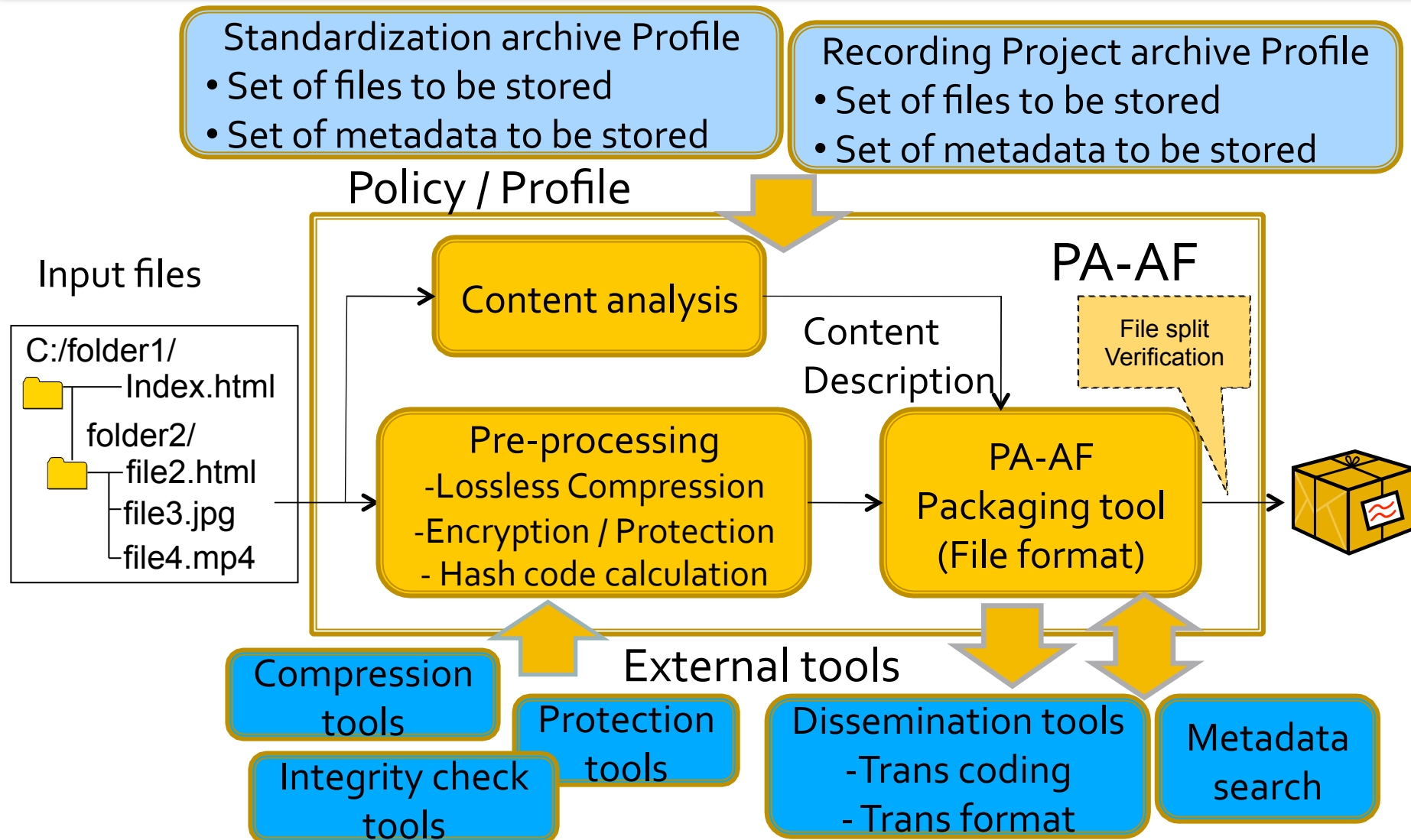
ISO TC20/SC13 The Consultative Committee for Space Data Systems (CCSDS)

and Data Archive Ingest (DAI) working group

# MPEG-A Professional Archival Application Format (PA-AF)

- ISO/IEC 23000-6 Professional Archival Application Format
- Build with MPEG tools and other standards
- General Packaging Format
  - Preserve hierarchical folder structure and attributes
  - Any Meta-information can be contained as XML files (MPEG-7, Dublin core, METS, AES, etc.)
  - Policy / coding scheme independent
  - Cross-platform support (OS, file system, multiple-byte character set)
  - Support file sub-division based on the package file size
- Can serve as Information Packages defined in OAIS Reference Model (SIP, AIP, DIP)

# Scope of Professional Archival Application Format (PA-AF)



# MPEG tools applied to PA-AF

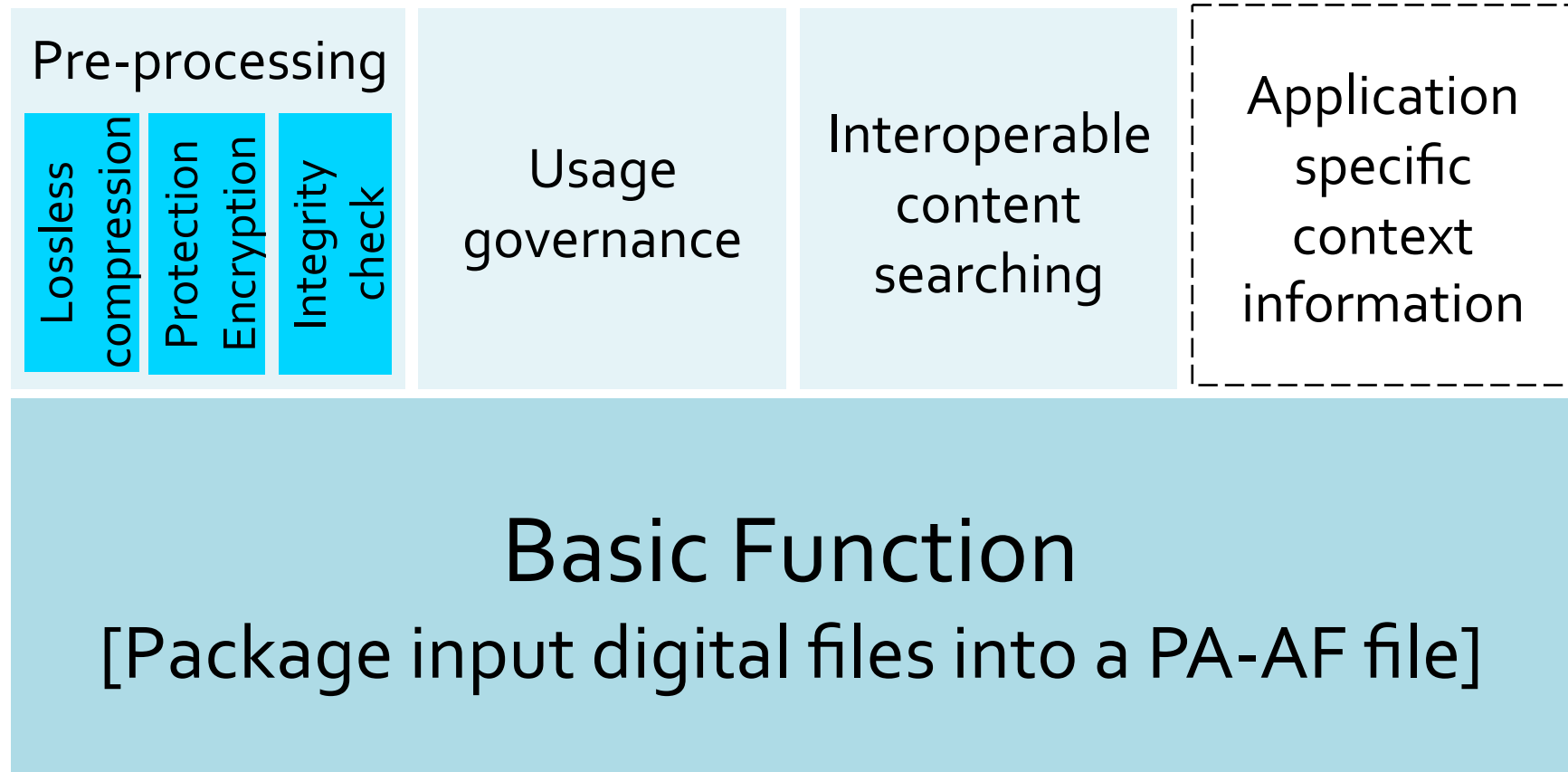
- MPEG-21 File Format (derivative of ISO-base file format)
- MPEG-21 Digital Item Declaration (DIDL) 2<sup>nd</sup> Ed.
- MPEG-21 Digital Item Identifier
- MPEG-7 Multimedia Description Schemes
- MPEG-21 Intellectual Property Management and Protection(IPMP) Base Profile
- MPEG-21 Rights Expression Language (REL)

Other pre-processing tools

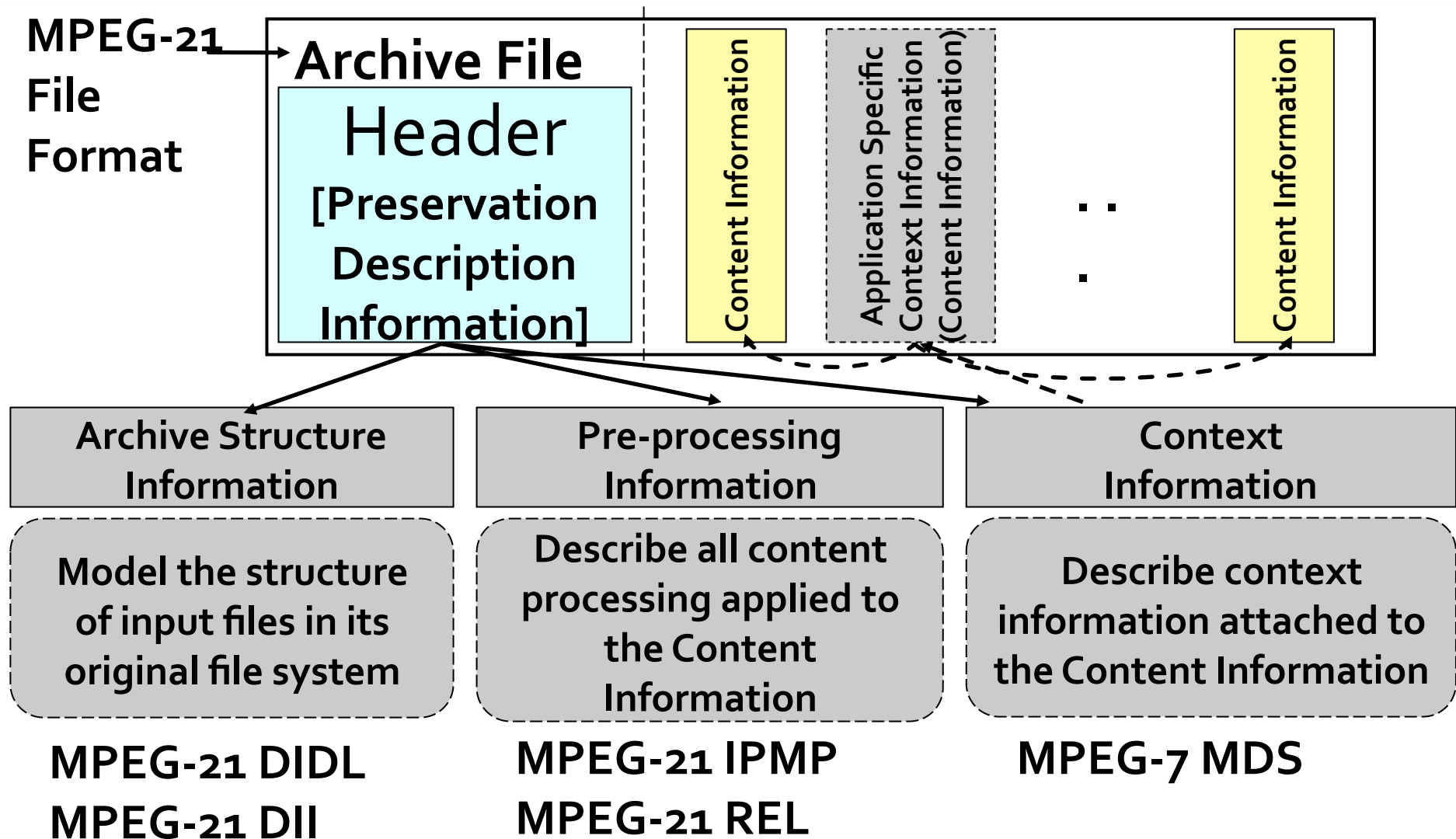
- MPEG-4 Audio Lossless Coding (ALS)
- OpenSSL, Gzip, etc.

# Basic and enhanced functionalities of PA-AF

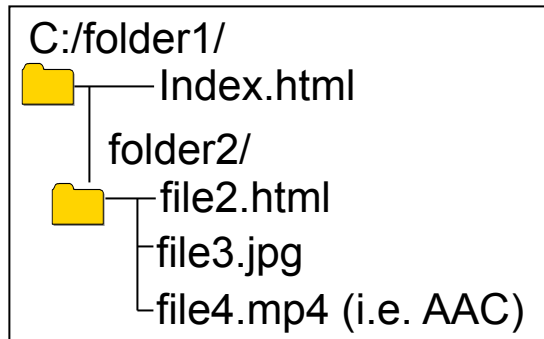
Archives can decide which level of services should be applied based on their own archiving policy



# PA-AF file structure

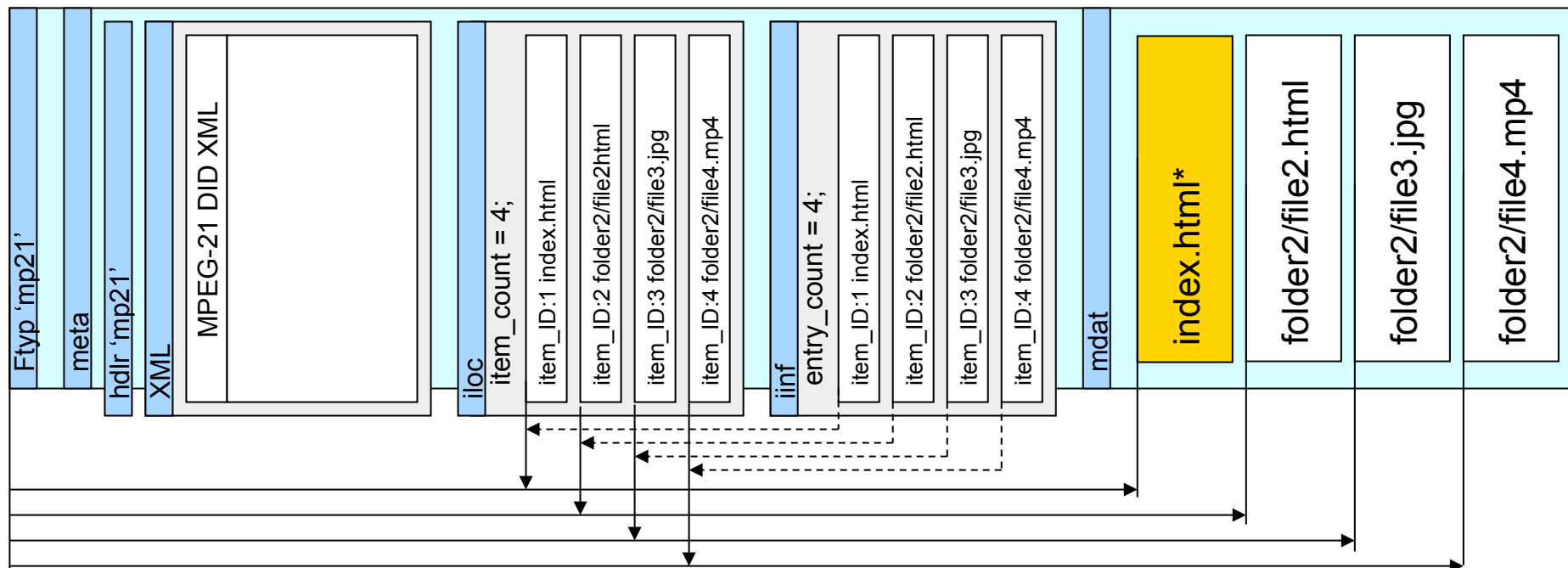


# Archiving HTML files using PA-AF

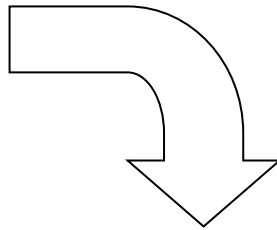
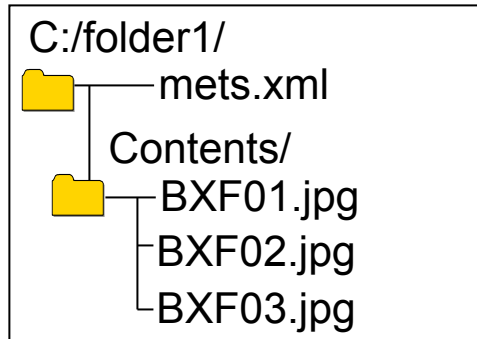


\*Index.html can be treated as an application specific context information

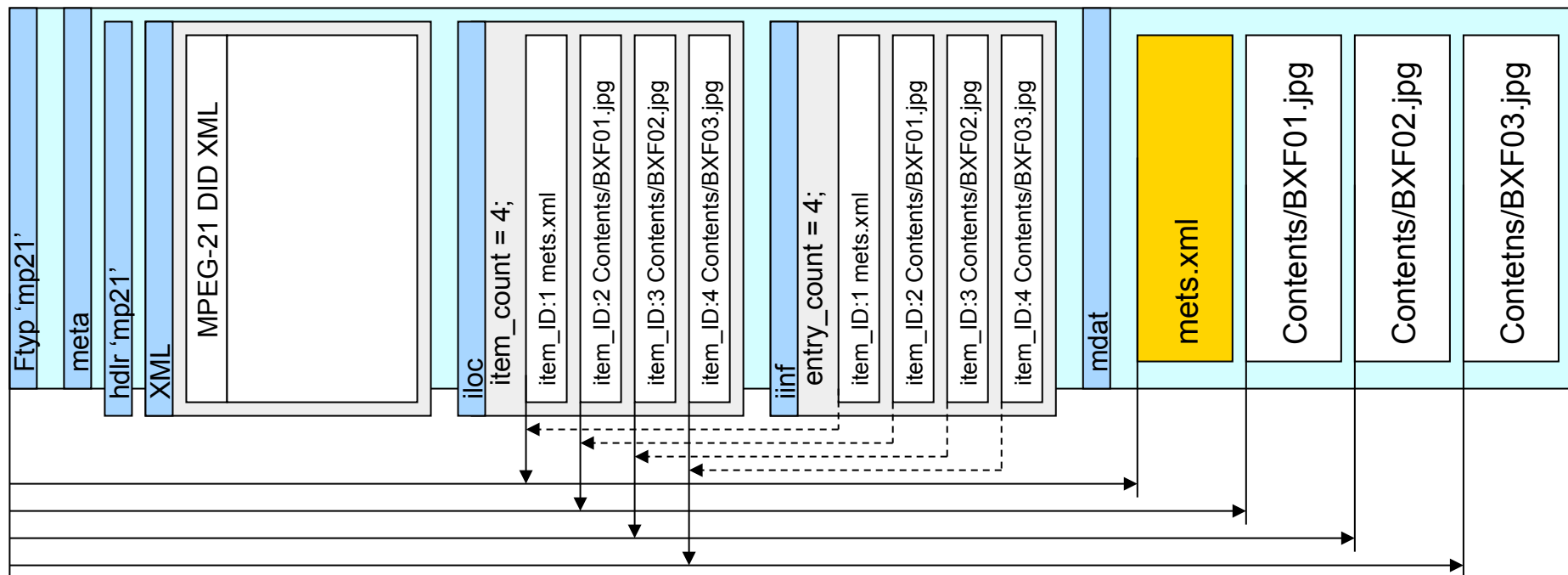
## MPEG-21 file format for the Professional Archival Application Format



# Archiving METS files as application specific context information



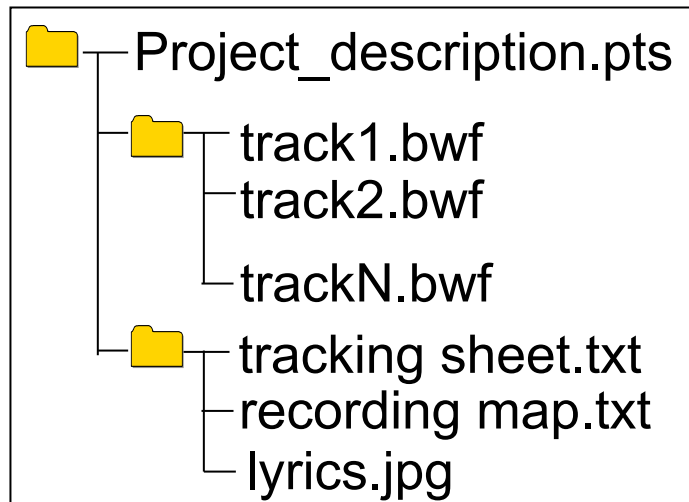
## MPEG-21 file format for the Professional Archival Application Format



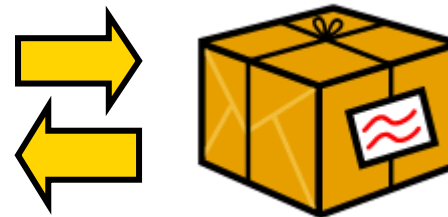


# Example exchange format for audio based on PA-AF with MPEG-4 ALS

- A music project often contains non-audio files
- Folder-structure-based archiving is suitable for delivery or storage of a recorded music project



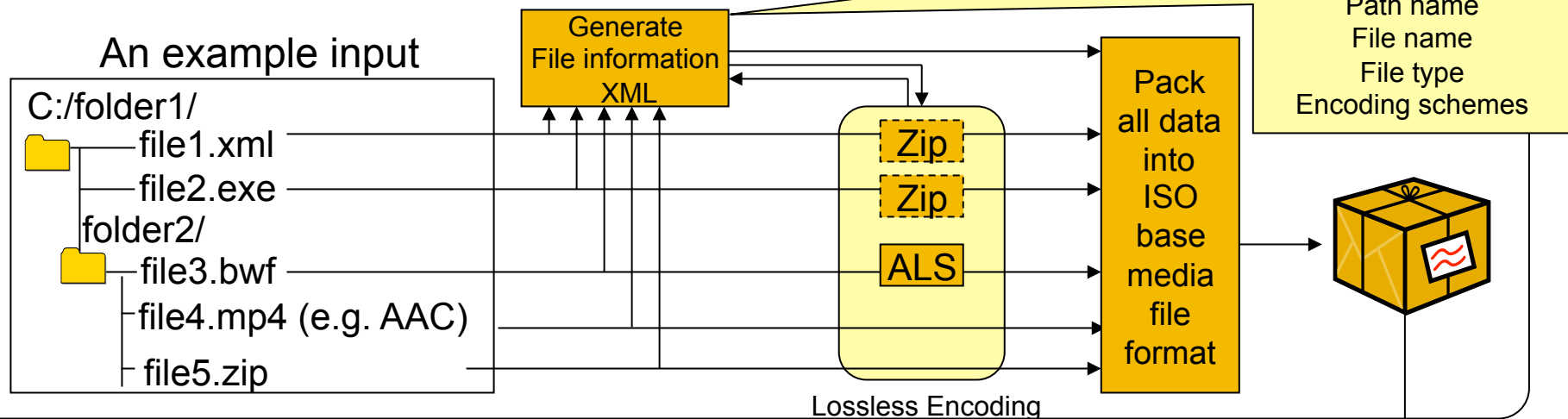
Lossless compression  
and archiving



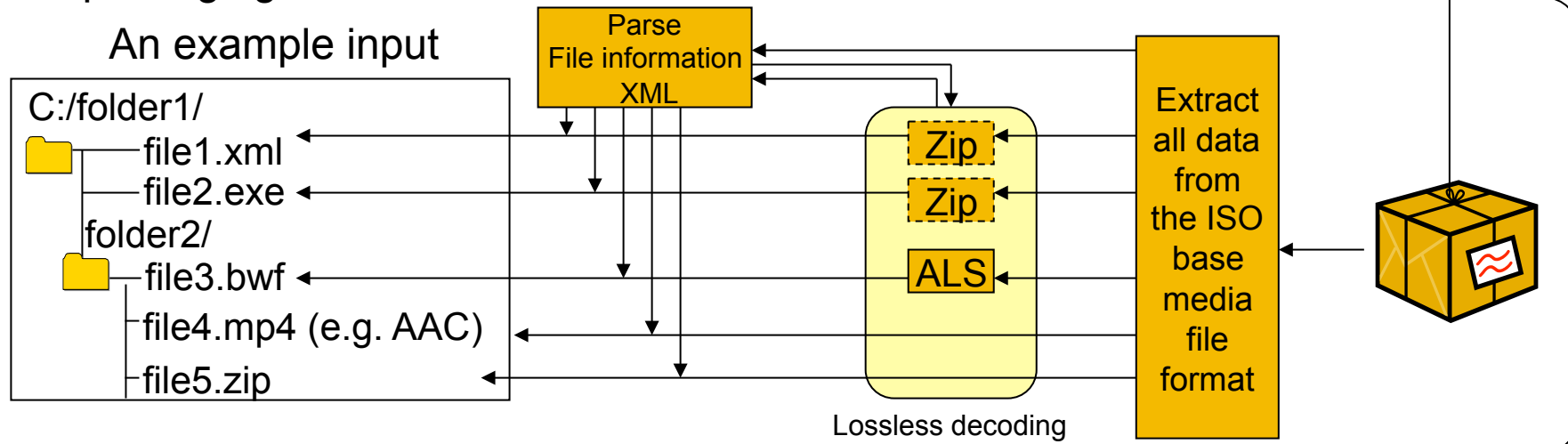
Complete reconstruction  
of original folder structure

# Example of Audio Archiving Tool

## Packaging Tool for the PA-AF

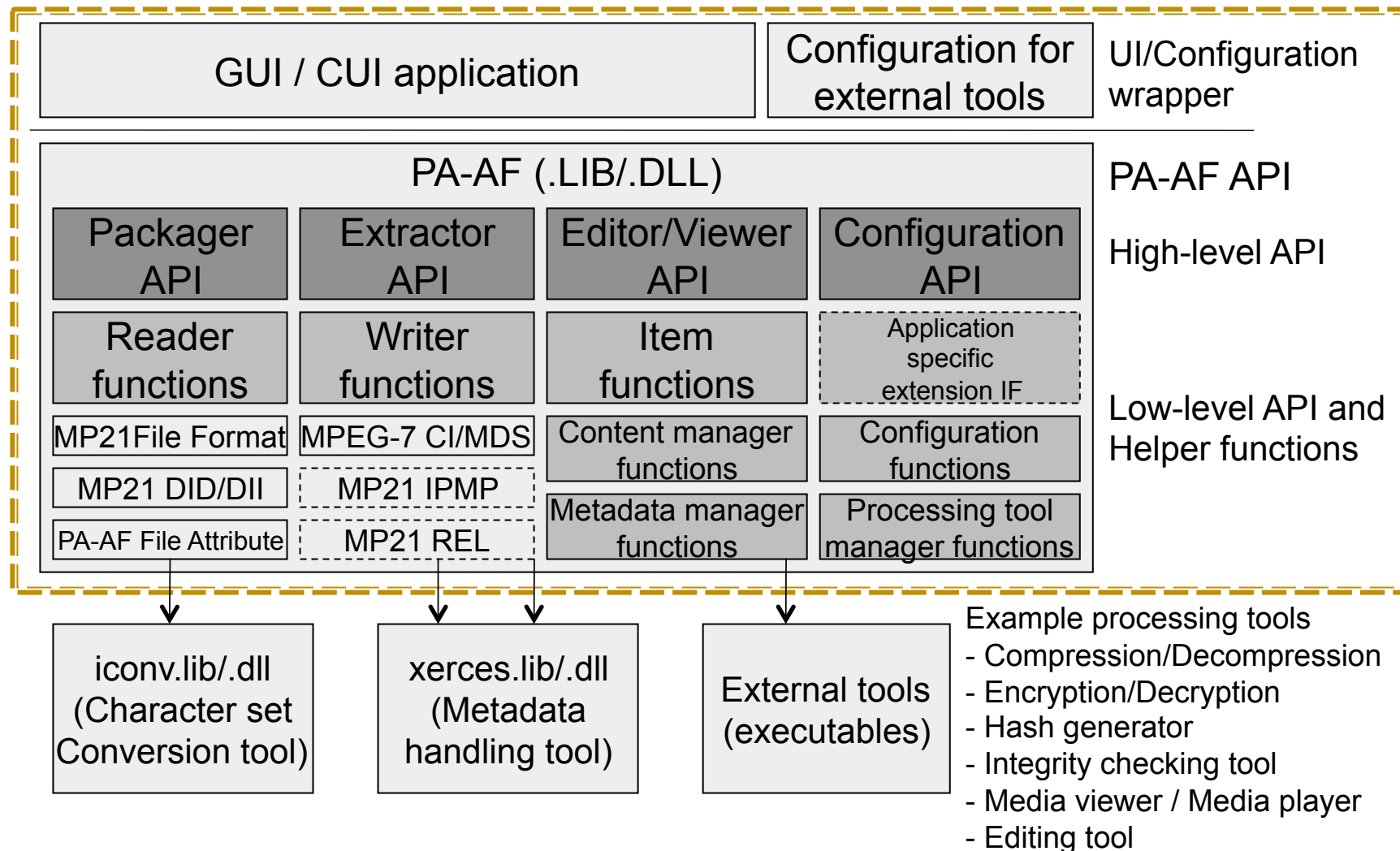


## Un-packaging Tool for the PA-AF



# Overview of the PA-AF reference software and API

PA-AF reference software



# Demo

- PA-AF reference software (GUI)
- Web archiving system for standardization related accets

# Conclusion

- PA-AF is a standardized packaging format for archiving digital files.
- PA-AF can serve as SIP, AIP, and DIP in the OAIS reference model
- Sharing the resources can reduce cost.
- MPEG-A PA-AF software is available at  
<http://www.brl.ntt.co.jp/cs/mrl/en/research.html>  
<http://mccb-pc3.kaist.ac.kr/paaf>