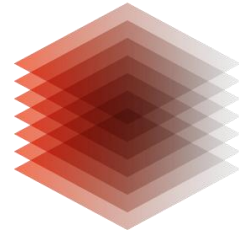LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK

TIB

# Preservation Planning:
# Monitoring the risk of file format obsolescence in TIBs
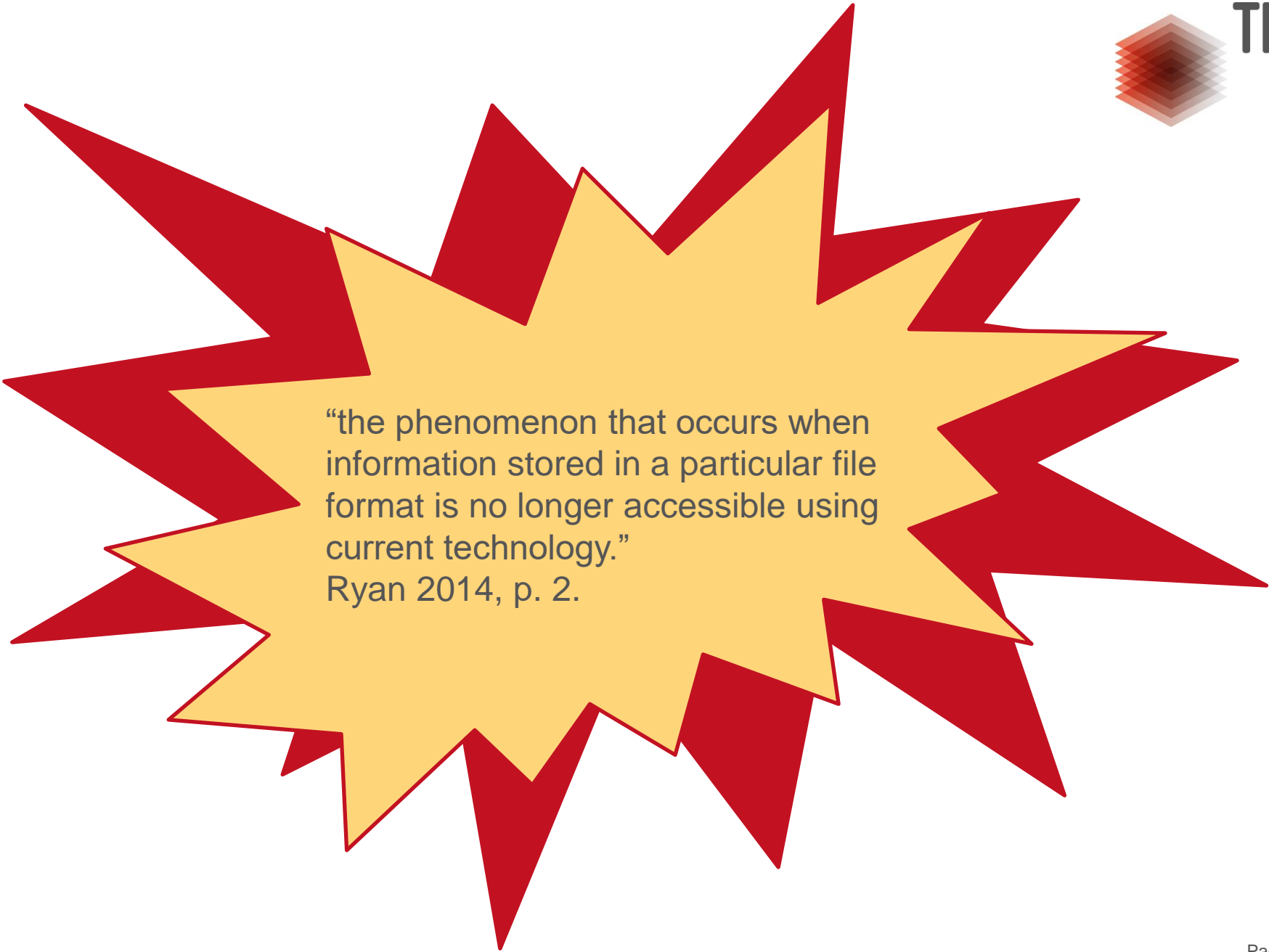# AV holdings

Merle Friedrich,
London, 24. October 2018
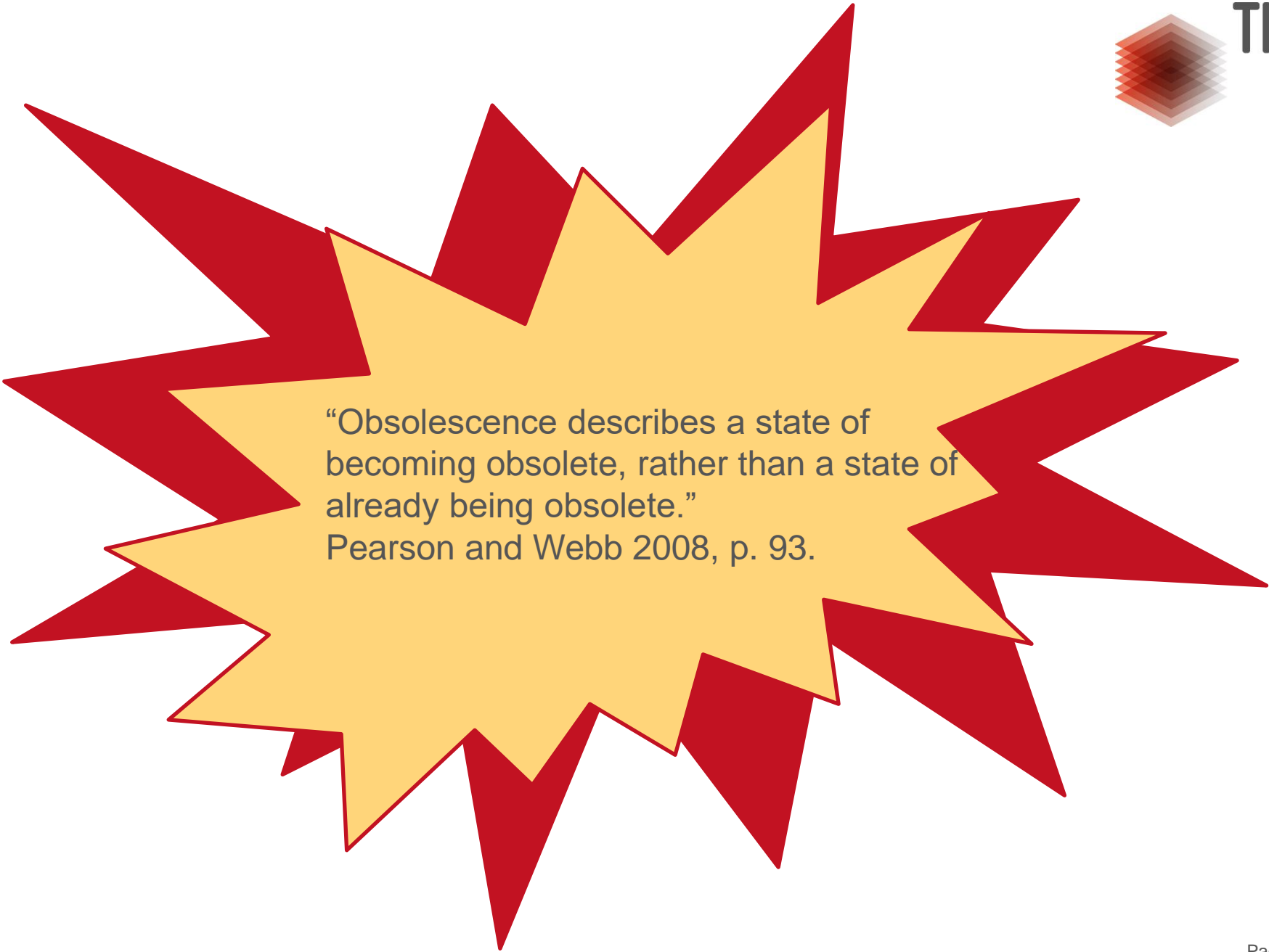No Time to Wait! Rough Consensus and Running Archives

Leibniz
Association

# Agenda

1. **Obsolescence**

2. **Normalization – yes or no?**

3. **OAIS – Preservation Planning**

4. **Monitor the Risk of Obsolescence**
   - **View Paths**

5. **Migration**

**References**

"the phenomenon that occurs when information stored in a particular file format is no longer accessible using current technology."
Ryan 2014, p. 2.

"Obsolescence describes a state of becoming obsolete, rather than a state of already being obsolete."
Pearson and Webb 2008, p. 93.

file format that is **at risk** to become **inaccessible** by our **designated community**

# Normalization

| Amount | Container | Video-Codec | Audio-Codec |
|---|---|---|---|
| 7384 | MPEG-4 | AVC | AAC, Version 4 |
| 1062 | WebM, Format : WebM | VP8 | Vorbis |
| 963 | MPEG-PS | MPEG Video, Version 2 | MPEG Audio, Version 1 |
| 765 | MPEG-4 | AVC | AAC |
| 706 | MPEG-PS | MPEG Video, Version 1 | MPEG Audio, Version 1 |
| 557 | MPEG-4 | MPEG-4 Visual | AAC, Version 4 |
| 421 | MPEG-4 | MPEG-4 Visual | |
| 210 | WebM, Version 2 | VP8 | Vorbis |
| 175 | MPEG-TS | AVC | AC-3 |
| 97 | MPEG-PS | MPEG Video, Version 2 | |
| 93 | AVI | AVC | PCM |
| 46 | MPEG-PS | MPEG Video, Version 2 | AC-3, Format : DVD-Video |
| 44 | Flash Video | VP6 | MPEG Audio, Version 1 |
| 37 | MPEG-PS | MPEG Video, Version 2 | MPEG Audio, Version 1, Format : MPEG Audio, Version 1, Format : RLE |
| 33 | Flash Video | VP6 | MPEG Audio, Version 2 |
| 31 | MPEG-4 | AVC | |
| 22 | MPEG-PS | MPEG Video, Version 2 | MPEG Audio, Version 1, Format : DVD-Video |
| 30 | MPEG-4 | MPEG-4 Visual | AAC |
| … | … | … | … |

# Normalization

**Advantages**

Limited set of formats in the archive

- validation
- preservation planning
- less tools

**Disadvantages**

- migration neccessary by producer
    - undocumented changes of material
    - delivering material is unattractive

- or migration neccessary by archive
    - work load
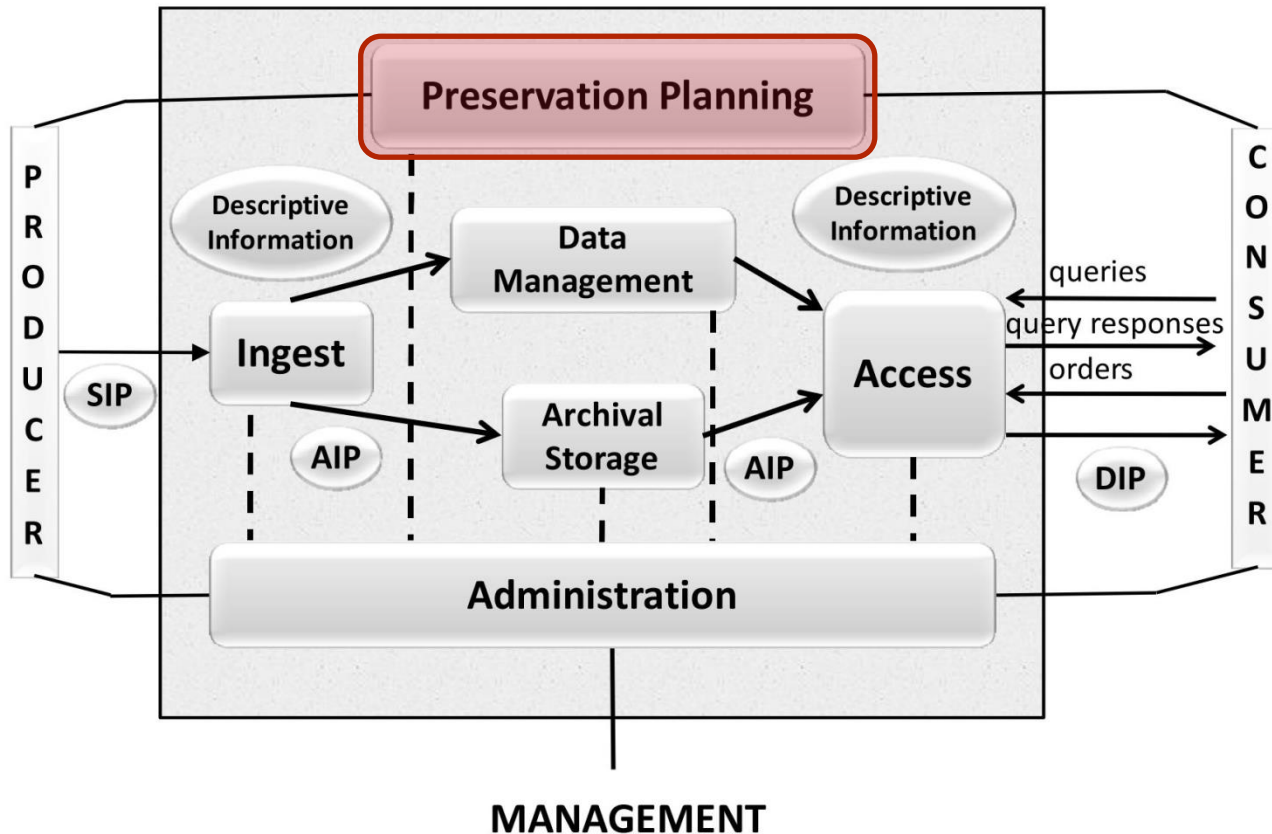    - keeping original (authenticity!) -> more storage space needed

**Sometimes we have to take everything we can get**
**… just the way it is**

obsolete

# OAIS – Preservation Planning



**Figure 4-1: OAIS Functional Entities**

# Watch

Technology Watch

Community Watch

Preservation Watch

# Monitor the Risk of Obsolescence

Indicators for Obsolescence:

Ryan (2014): Who's afraid of file format obsolescence?

- conducted and evaluated expert interviews
- compared 21 indicators

*"the only factor that can be considered a direct cause to file format endangerment is **rendering software available**. Secondary factors to this are specifications available and community/3rd party support"*
*Ryan (2014), p.210*
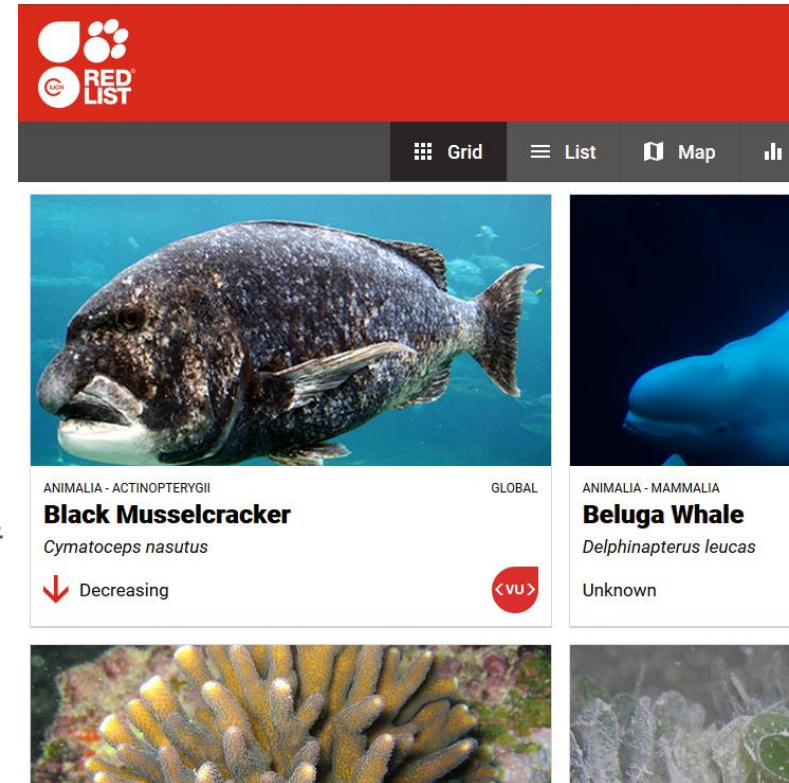
# View Paths

View paths for the majority of file formats in TIBs  AV holdings

| Format | View Path |
|---|---|
| Container: MPEG-4<br>Video: AVC<br>Audio: AAC, Version 4 | Intel® Core™ i7-8700  – Windows 10 – Windows Media Player Version 12 |
| | Intel® Core™ i5-3470 – Windows 8.1 Pro – VLC Media Player Version 2.2.2 |
| Container: WebM<br>Video: VP8<br>Audio: Vorbis | Intel® Core™ i7-8700  – Windows 10 – MPC-HC Version 1.7.13 |
| | Intel® Core™ i5-3470 – Windows 8.1 Pro – VLC Media Player Version 2.2.2 |
| Container: MPEG-PS<br>Video: MPEG Video, Version 2<br>Audio: MPEG Audio, Version 1 | Intel® Core™ i7-8700  – Windows 10 – Windows Media Player Version 12 |
| | Intel® Core™ i5-3470 – Windows 8.1 Pro – VLC Media Player Version 2.2.2 |

Concept View Paths: Steenbakkers, Johan F. (2005): Digital Archiving in the Twenty-First Century: Practice at the National Library of the Netherlands

# View Paths

# Outlook - Migration

Migrate to a suitable format

Adoption

Transparency

Metadata Support

Technological dependencies

Disclosure
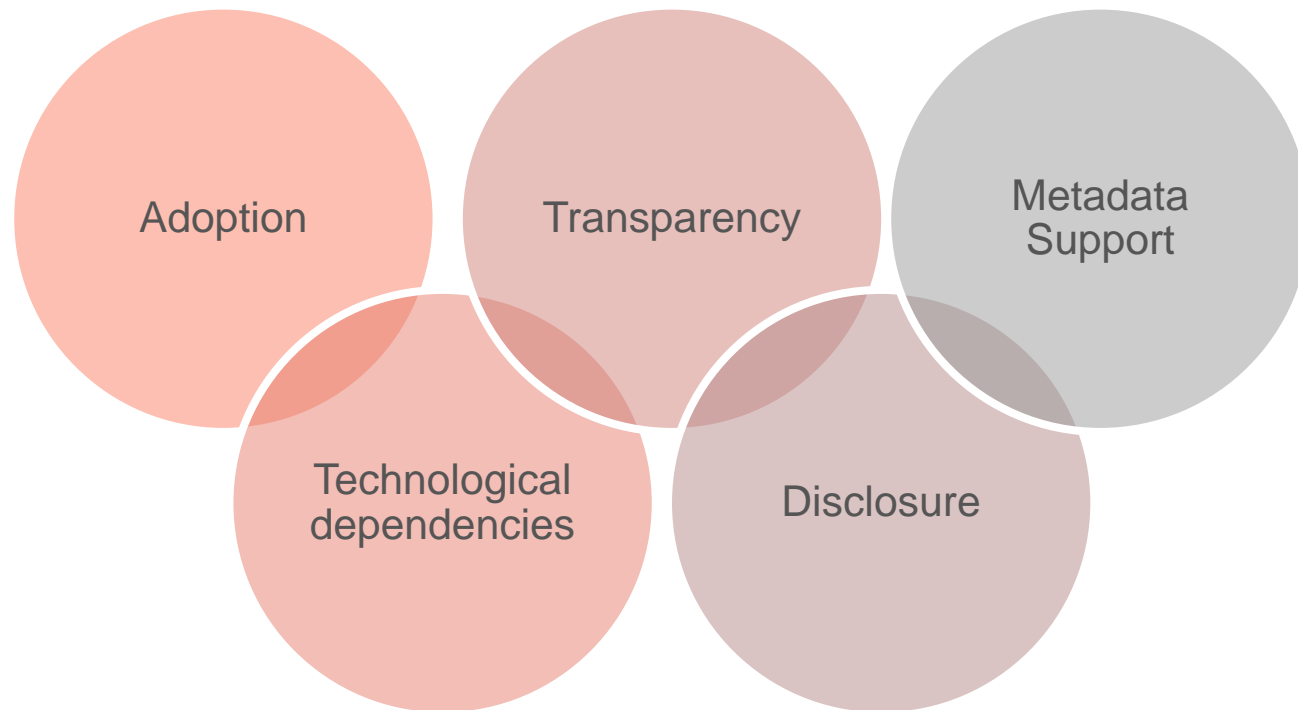
According to Todd, Malcolm, 2009. File formats for preservation:
Technology Watch Report.

# Outlook - Migration

Preservation Planning with Rosetta

- extract technical metadata upon ingest
- automated risk report (based on view paths)
- creating sets on the base of risk report (and additional technical metadata, collection, rights…)
- define evaluation criteria
- define tool for migration (internal or external)
- run test and analyse results (automated or manual)
- sign-off preservation plan (or refine)

Future work:

develop a migration plugin for our archival software (Rosetta)
to migrate to ffv1 / mkv (with ffmpeg)

# References

**Pearson, David; Webb, Colin (2008)**: Defining File Format Obsolescence: A Risky Journey. In *IJDC* 3 (1), pp. 89–106. DOI: 10.2218/ijdc.v3i1.44.

**Recommended Practice CCSDS 650.0-M-2, June 2012**: Reference Model for an Open Archival Information System (OAIS). Available online at https://public.ccsds.org/pubs/650x0m2.pdf.

**Ryan, Heather M. (2014)**: Who's afraid of File Format Obsolescence? Evaluating File Format Endangerment Levels and Factors for the Creation of a File Format Endangerment Index. University of North Carolina, Chapel Hill. School of Information and Library Science. Available online at https://cdr.lib.unc.edu/record/uuid:86e05605-97e3-40de-927a-ce5f5a6fa4d0.

**Steenbakkers, Johan F. (2005)**: Digital Archiving in the Twenty-First Century: Practice at the National Library of the Netherlands. In *Library Trends* 54 (1), pp. 33–56. DOI: 10.1353/lib.2006.0010.

**Todd, Malcolm (2009)**: File formats for preservation. Technology Watch Report. Available online at https://www.dpconline.org/component/docman/?task=doc_download&gid=375.

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY
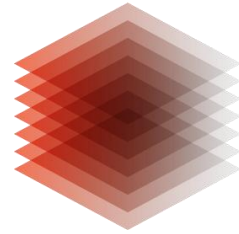
TIB

# MORE INFORMATION
## www.tib.eu

**Contact**
Merle Friedrich
T +49 511 762-14274, merle.friedrich@tib.eu

Leibniz
Association