Topic 7 - Data Management, Systems & Migration

Speaker notes

No notes on this slide.

Speaker notes

No notes on this slide.

Overview

- Types of systems for managing files and data: DAMS, MAMS, DPS
- Everything about migration

Speaker notes

No notes on this slide.

MAM? DAM!

- **DAM:** Digital Asset Management
- **MAM:** Media Asset Management
- Collection Management System

Speaker notes

Disclaimer: DAMs or MAMs may be used for preservation, but are sometimes/often not designed or intended for use in a preservation context, but merely to store and handle “digital assets”: from regular office files (documents, images, etc) to managing in-house assets of larger companies. And some of these systems were then “also” used by archives.

The term “Collection Management System” usually indicates that it was more likely intended to be used in a preservation context, such as museums for example - where it may be used beyond digital: To handle physical collections even, like books, chairs, or anything. And files ;)

You may see the term “Collection Management System” being used interchangeably with DAM or MAM by the preservation community.

(Note: The abbreviation “CMS” usually means “**Content** Management System” which is something completely different. In order to save some screen space, I will use the abbreviation “CMS” in these slides however instead of typing “Collection Management System”)

A DAM is usually the generic version of MAM - and sometimes the borders between “is it a MAM? is it a DAM?” are fuzzy and unclear, because they are so closely related.

Typical for “classic” DAMs: Often trimmed and designed for handling “2D material” (documents, images). When it comes to audiovisual, they’re mostly inadequate or not suitable for archive-suitable quality media handling.

MAMs are usually better suited for handling media. For example:

- auto-generating access/preview copies suitable for low-bandwidth/internet/browser access.
- image area annotation
- time based annotation (“markers”)

But watch out! Even systems designed to handle AV media, are often not taking too much care about preservation/archival aspects. It’s not uncommon that “looks/sounds good enough!” is exactly where you’re at.

So please:

- Try *before* you buy!
- Don’t trust sales. Ever.

A short list

of some popular OpenSource CMS:

- AtoM (Access To Memory)
- Omeka
- ResourceSpace

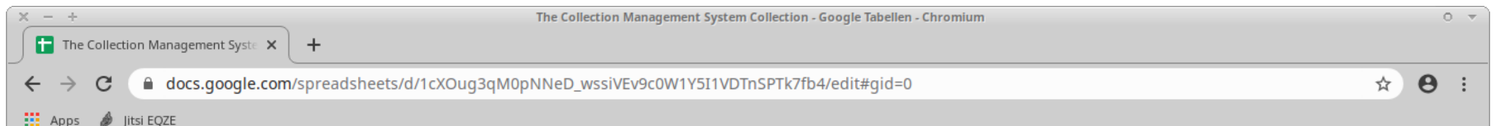
Speaker notes

No notes on this slide.

A loooong list...

of not only OpenSource CMS:

https://bits.ashleyblewer.com/blog/2017/08/09/collection-management-system-collection/



Speaker notes

No notes on this slide.

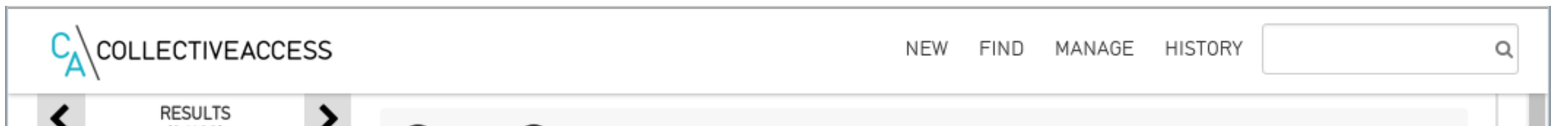
A look into: CollectiveAccess

Speaker notes

No notes on this slide.

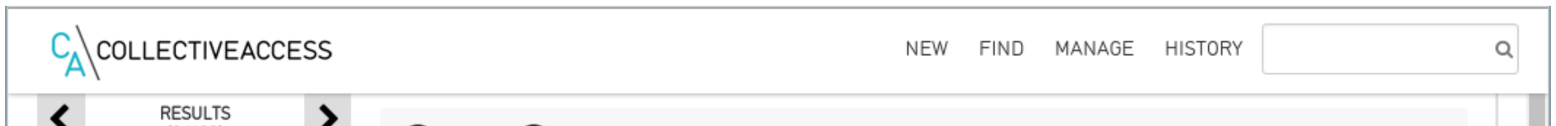
Speaker notes

No notes on this slide.



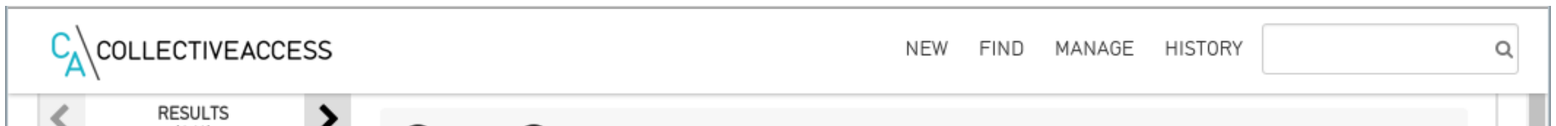
Speaker notes

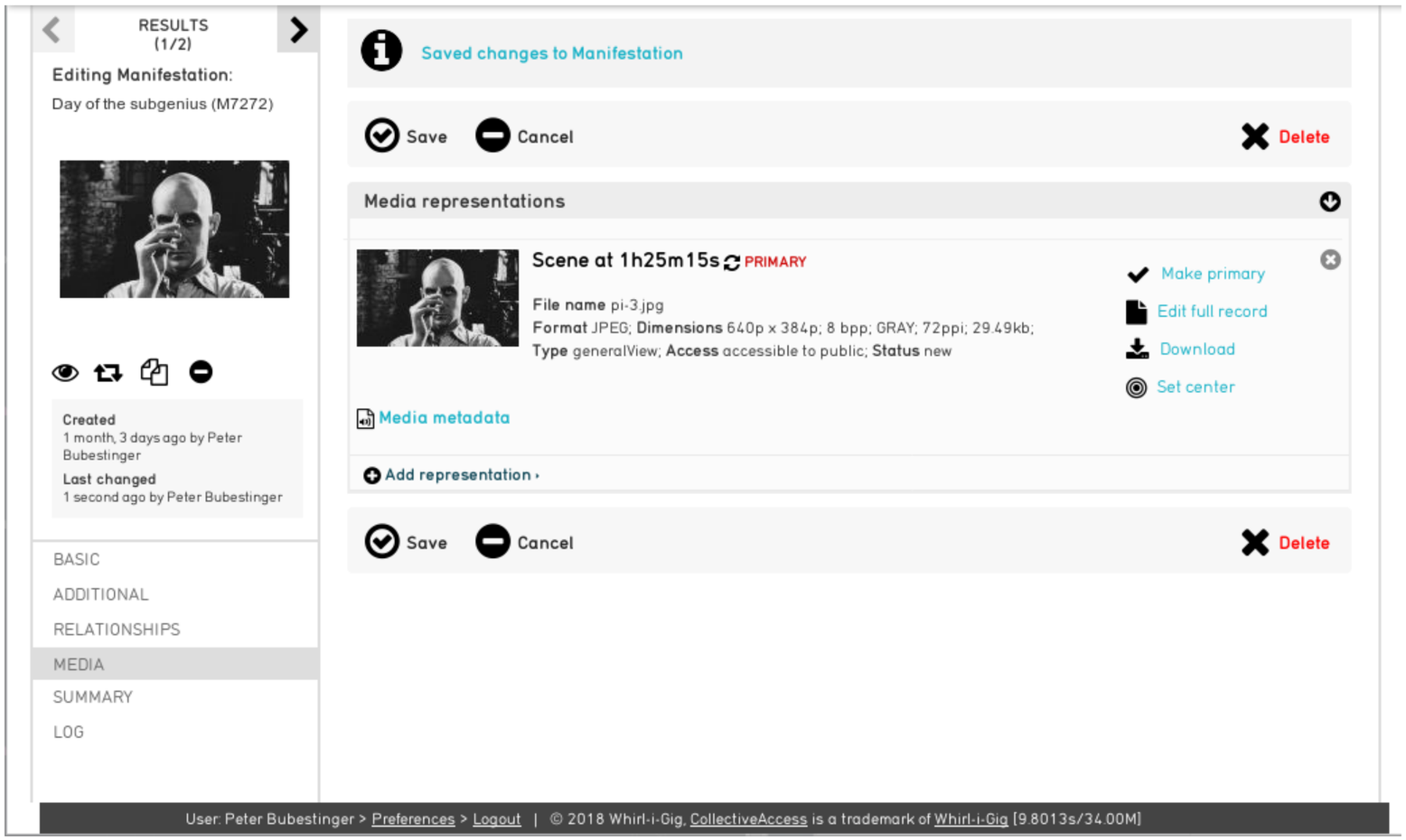
No notes on this slide.



Speaker notes

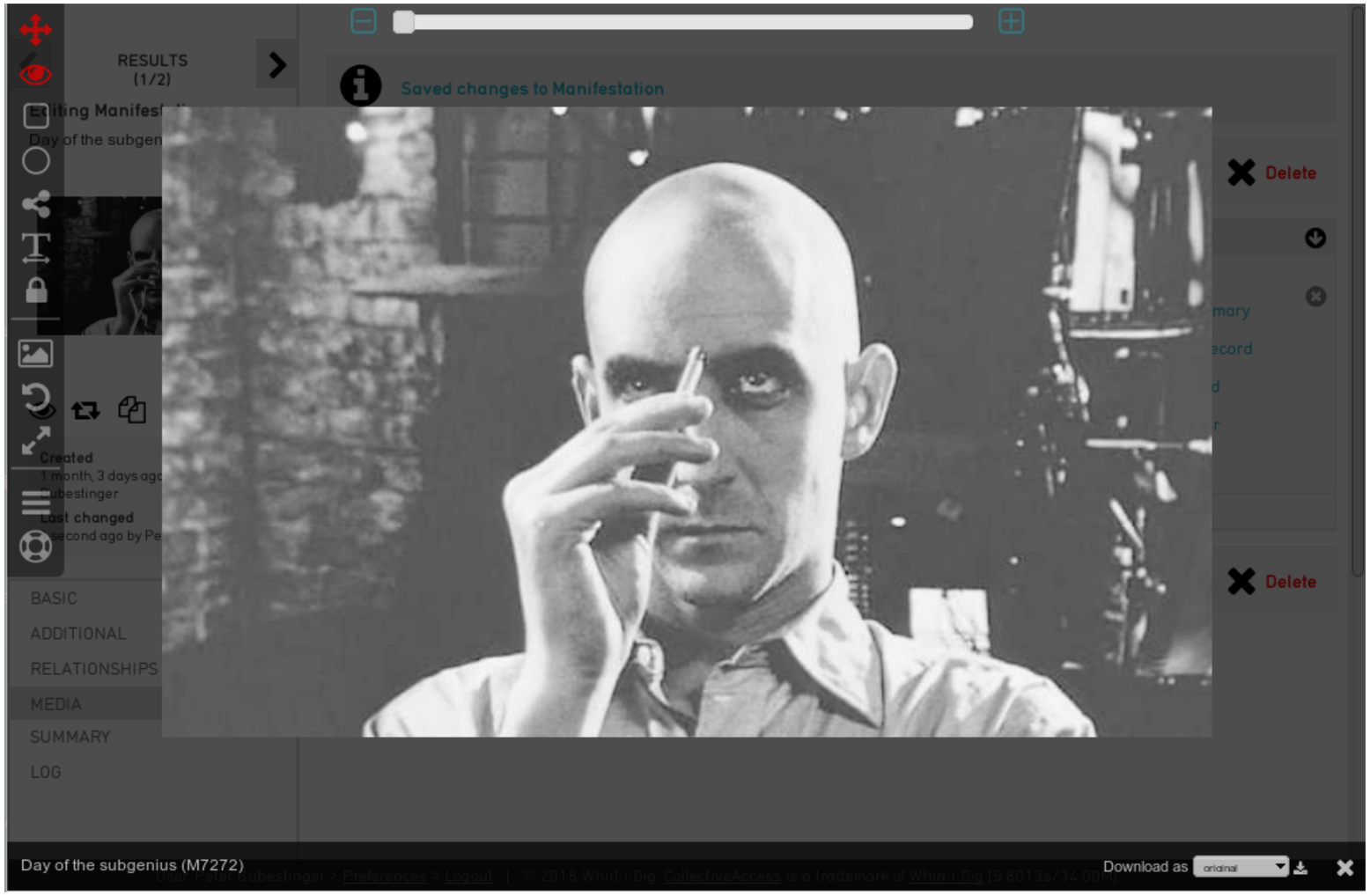
No notes on this slide.





Speaker notes

No notes on this slide.



Speaker notes

No notes on this slide.

DPS: Digital Preservation System

There are several workflows with individual tasks to be performed for preservation of digital objects.

DPS can help/improve organizing, monitoring and maintaing these tasks.

Speaker notes

For example from Ingest SIP to AIP:

- Fixity creation
- Filename documentation & detox
- Virus checking
- Adding metadata
- Create preview images
- …
- Adjust structure to in-house rules
- … A popular example of such a system is “Archivematica”

archivematica

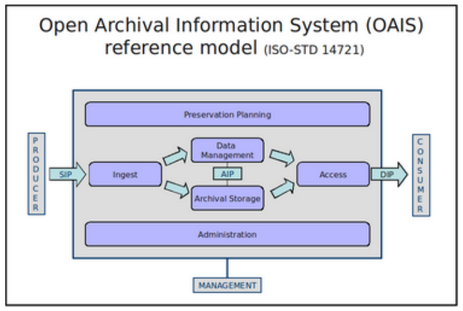
Home | Downloads | Documentation | Community | Development | News | Wiki | Demo

Archivematica 1.9.1 is our latest release.

What is Archivematica?

Archivematica is a free and open-source digital preservation system that is designed to maintain standards-based, long-term access to collections of digital objects. Archivematica is packaged with the web-based content management system AtoM for access to your digital objects.

Open source OAIS



Archivematica provides an integrated suite of free and open-source tools that allows users to process digital objects from ingest to archival storage and access in compliance with the Open Archival Information System (OAIS) functional model and other digital preservation standards and best practices.

All of the Archivematica code is released under a GNU Affero General Public License and Archivematica documentation is released under a Creative Commons Attribution-ShareAlike 4.0 International License.

Lowering the barriers to best-practice digital preservation

The goal of the Archivematica project is to give archivists and librarians with limited technical and financial capacity the tools, methodology and confidence to begin preserving digital information today. The project has conducted a thorough OAIS use case and process analysis to synthesize the specific, concrete steps that must be carried out to comply with the OAIS functional model from Ingest to Access. Through deployment experiences and user feedback, the project has expanded even beyond OAIS to address analysis and arrangement of transferred digital objects into SIPs and allow for archival appraisal at multiple decision points

Archivematica 1.9.1

[Table of contents]

CONTENTS

SEARCH

Open the general index or type your search in the search box.

AVAILABLE PROJECTS

Archivematica

- Version 1.9.1 (stable)
- Version 1.8.1 (legacy)
- Version 1.7.2 (legacy)
- Version 1.6.1 (legacy)
- Version 1.5 (legacy)
- Version 1.4 (legacy)

Archivematica Storage Service

- Version 0.14.1 (stable)
- Version 0.13.0 (legacy)
- Version 0.12.0 (legacy)
- Version 0.11.1 (legacy)
- Version 0.10 (legacy)
- Version 0.9 (legacy)

Demo installation (Sandbox)

Speaker notes

No notes on this slide.

österreichischemediathek
audiovisuelles archiv · technisches museum wien

Suche

Digitalisierung

Menü

Digitalisierung > DVA Profession engl.

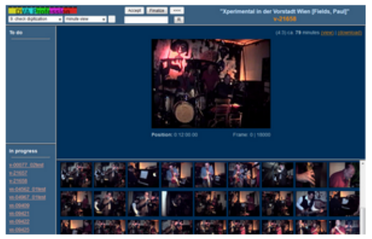
DVA-Profession

DVA-Profession („Digital Video Archive – Profession “) is a workflow management system for the digitization of tape based video recordings, designed for the purpose of digital long term preservation.

Auf dieser Seite:

Summary

The process of digitization constitutes a massive change of the analog original – but in the long term only a digital representation of the analog tape can be sustained for the future.
Respecting this massive intervention, accurate documentation of every step in the process of conversion is necessary.
For public memory institutions a precise documentation of every operation guarantees the authenticity of the original analog source in the digital age – a necessity for future research respecting the verification of sources.
DVA-Profession manages the whole digitization process and offers precise documentation of all tasks in the process (XML METS). Access copies and preview images of the digitized videos are generated automatically in the process. DVA-Profession provides the video operator with analysis data of the digital video files which are used for accurate quality control of the digitization. MD5 hashcodes are generated in the workflow to offer full control over the integrity of the produced files. When the digitization process is complete and the quality control shows a success, the created files are accurately written to their final storage location.
The archive master files and metadata produced by DVA-Profession constitute a solid foundation for successful digital long term preservation.
Based on the metadata created in the DVA-workflow the whole history of the digitization and the signal chain

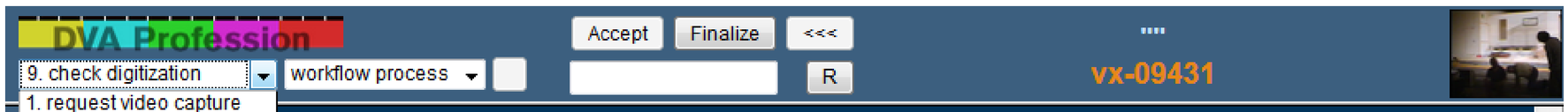


DVA Screenshot Minute View

Speaker notes

No notes on this slide.

Workflow management



Speaker notes

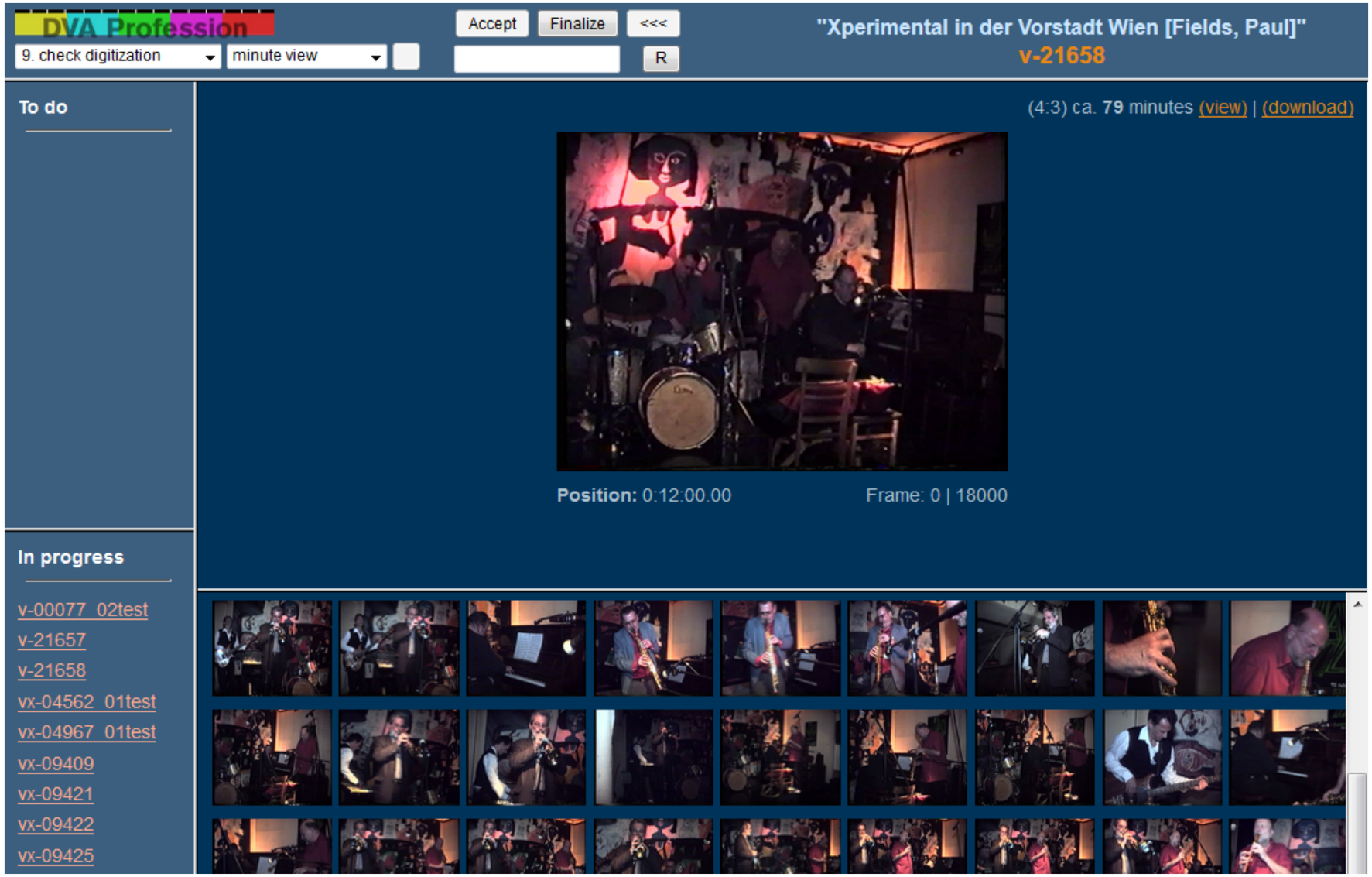
No notes on this slide.

Quality control

Speaker notes

No notes on this slide.

Other stuff



Speaker notes

No notes on this slide.

Comments?

Questions?

Speaker notes

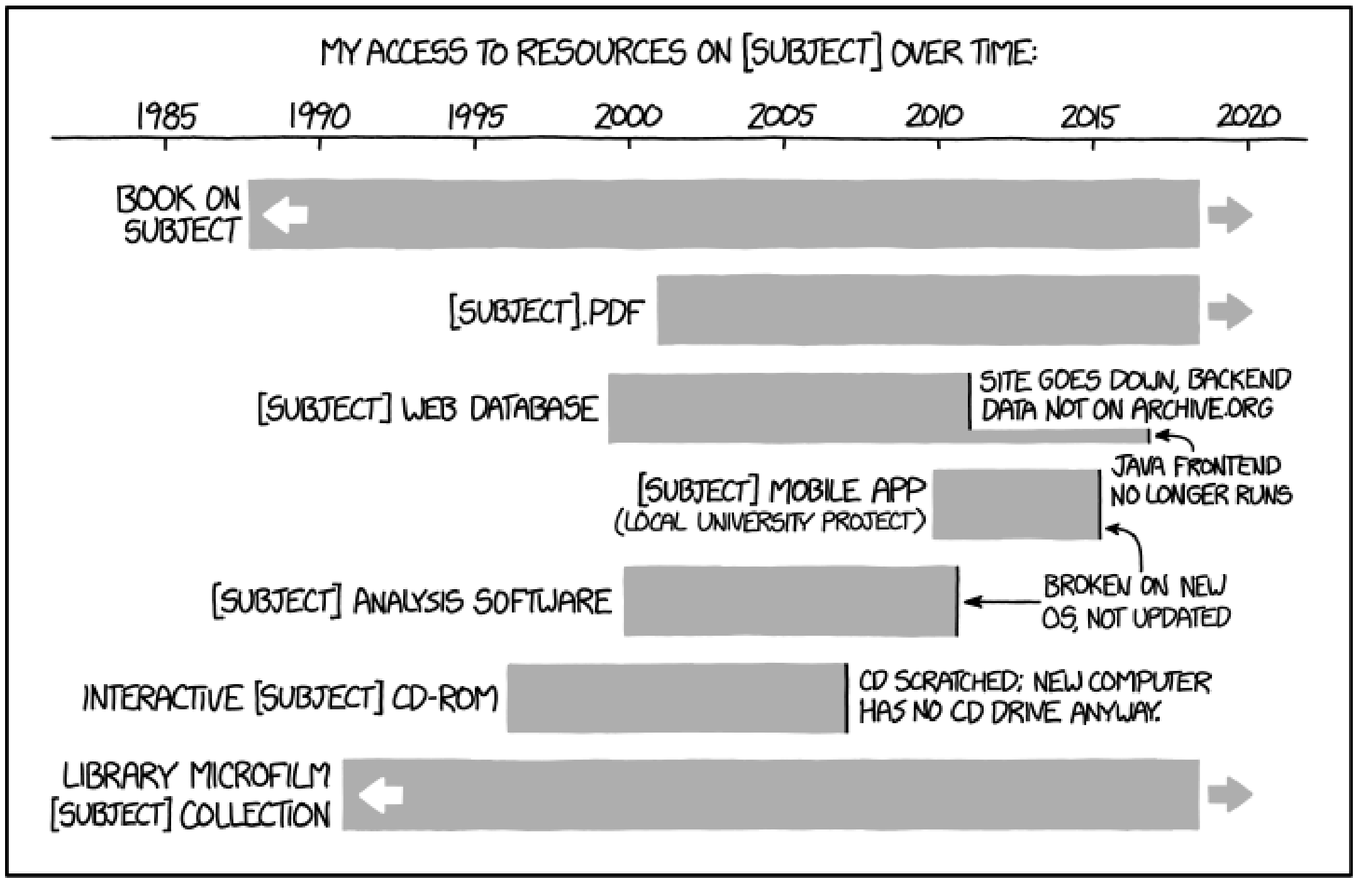
No notes on this slide.

Now for something completely different...

Speaker notes

No notes on this slide.

Migration

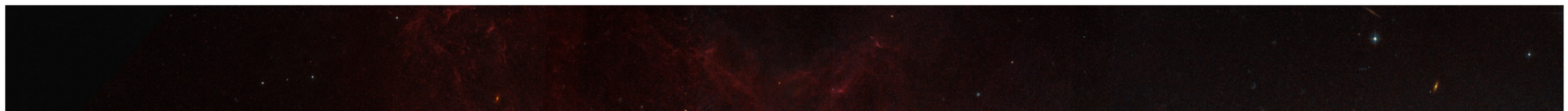


IT'S UNSETTLING TO REALIZE HOW QUICKLY DIGITAL RESOURCES CAN DISAPPEAR WITHOUT ONGOING WORK TO MAINTAIN THEM.

Speaker notes

No notes on this slide.

How long?



Speaker notes

No notes on this slide.

Eternal Migration

- There is no final carrier.
- There is no evergreen format.

Therefore fact: Any data must sooner or later be migrated.

Speaker notes

No notes on this slide.

Migration Types

- Storage
- Format
- Software / platform / environment

Speaker notes

No notes on this slide.

Migration Types

Or more generically speaking:

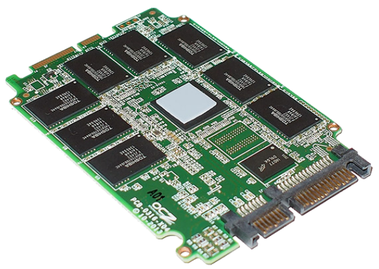
- Hardware
- Software
- File

Speaker notes

No notes on this slide.

Device Media



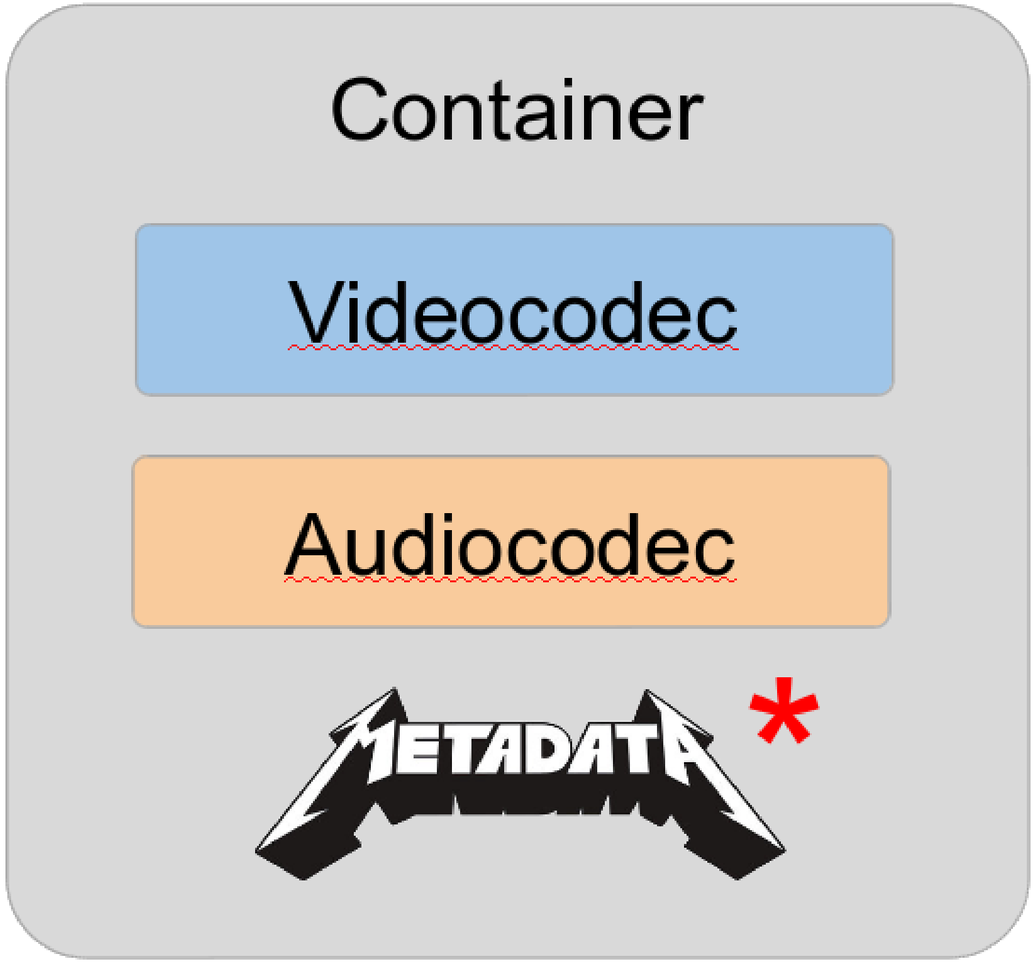



…or other media types (holograms, DNA, etc)

Speaker notes

No notes on this slide.

Data Format(s)



Speaker notes

Not all files may include non-AV data. But most do.

Depending on what data that is, it may involve different formats. Even if just plain text descriptive metadata: What about encoding?

Speaker notes

No notes on this slide.

Device Format Migration

1. Plan your migration
2. Copy the data
3. Check integrity of copy
4. Cross your fingers…

Speaker notes

If everything was planned and executed well, there should be no major issues. However, don’t be surprised to encounter some (hopefully little) things that you either haven’t anticipated, or simply couldn’t have expected or known in the first place.

That should be the exception though - not the rule ;)

Migration Planning

- Consider which changes are needed.
- Evaluate when, how and who.
- Make sure you have a valid backup.
- Schedule possible downtime (and impact on work).
- Impact on IT-administration/access?
- Estimated duration until migration is finished?

Speaker notes

No notes on this slide.

Software / platform / environment

Might require:

- Data format migration
- Reorganizing file structure (relocate, rename)
- Change of physical equipment
- Testing to avoid regressions
- etc.

Speaker notes

No notes on this slide.

Things can happen…

- Silent regression
- Unseen (meta)data changes
- Domino-effect: Forced updates of other things
  I’m loving it! ®
- etc.

Speaker notes

No notes on this slide.

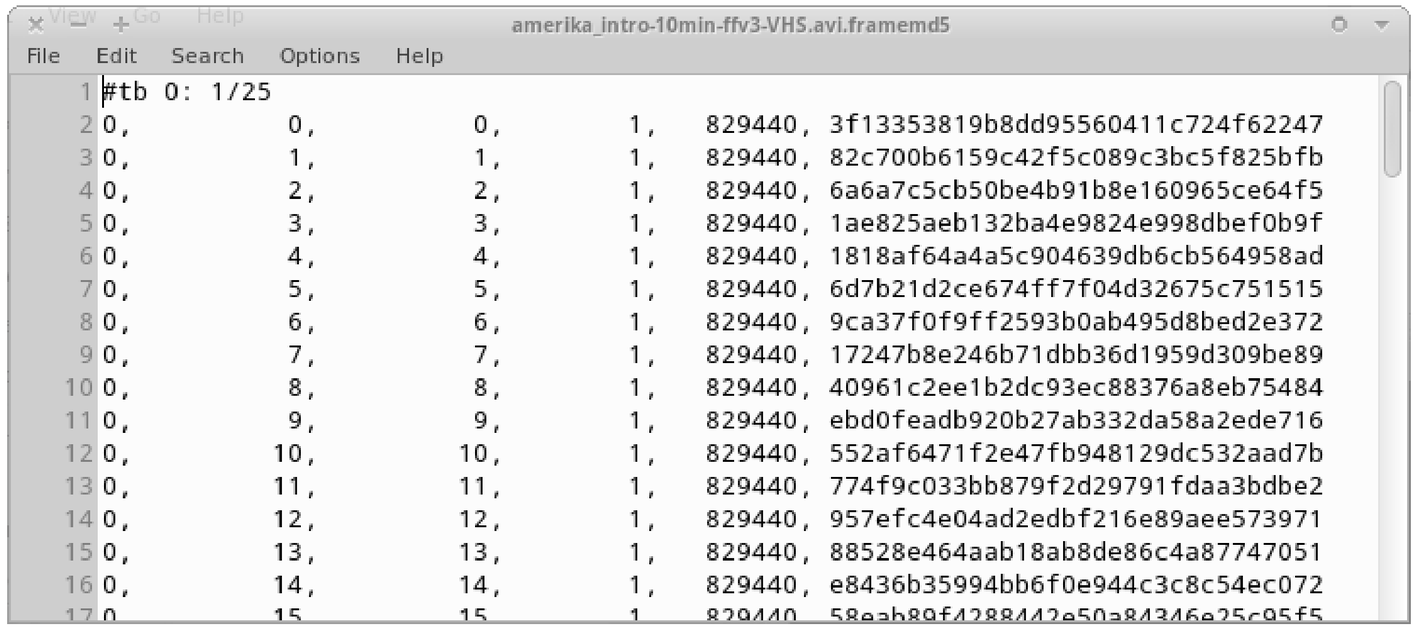
Media Format Migration

How to check integrity of format/codec migration?

Speaker notes

No notes on this slide.

FrameMD5



Speaker notes

FrameCRC/FrameMD5: One CRC/MD5 hashcode for each frame - or group of audio samples.

Eternal Migration

> *“After migration is before migration”*

- Embrace the concept of “Eternal migration”
- Try considering how to get out of a technology before, or while you’re using it.
- Find *your* timing sweet spot

Speaker notes

Sounds live a neverending chore, but it’s like brushing your teeth or washing yourself: If you integrate it in your daily routine, it’s not a big thing at all. And once you’ve done a few migrations of “whatever”, you’ll get the hang of it.

Obsolescence monitoring

- So, when *is* a good time to migrate?
- What could happen if you wait too long?
- Which vital components might become obsolete?

Speaker notes

Have someone in house that keeps an eye on technology news, and please: Speak and exchange yourself with peers in the community!

Don’t listen too much to broadcast/production regarding preservation. (unless you’re in that business)

Migration Summary

- Keep “Eternal Migration” in mind
- Consider migrations *before* you buy
- Migration + integrity checks = BFF
- Ask for documentation!
- Archive the sourcecode / schematics
- Monitor technology news
- Don’t wait too long…

Speaker notes

No notes on this slide.

Questions?

Comments?

Speaker notes

No notes on this slide.