Topic 7 - Data Management, Systems & Migration

Overview

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

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

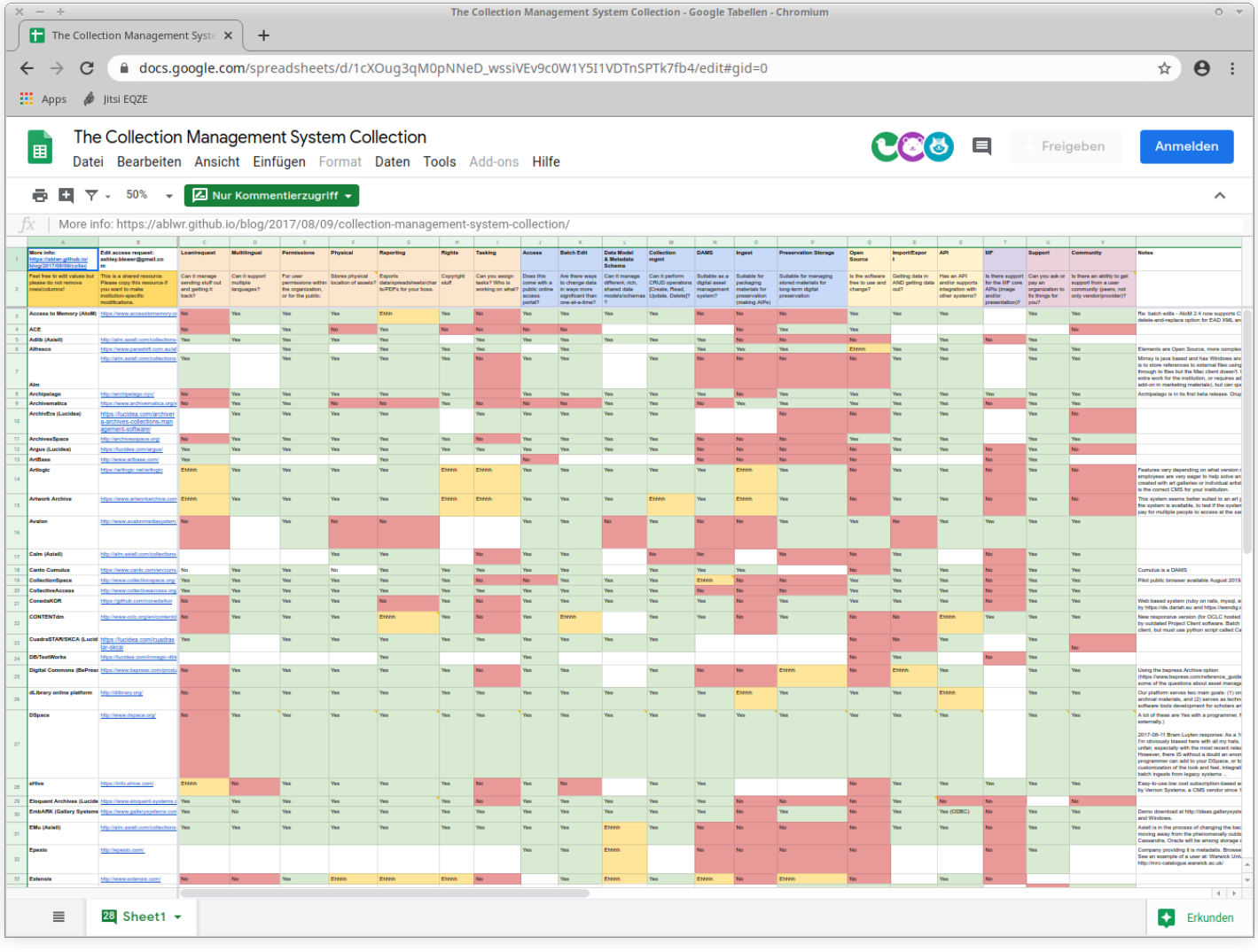
of some popular OpenSource CMS:

- AtoM (Access To Memory)
- Omeka
- ResourceSpace

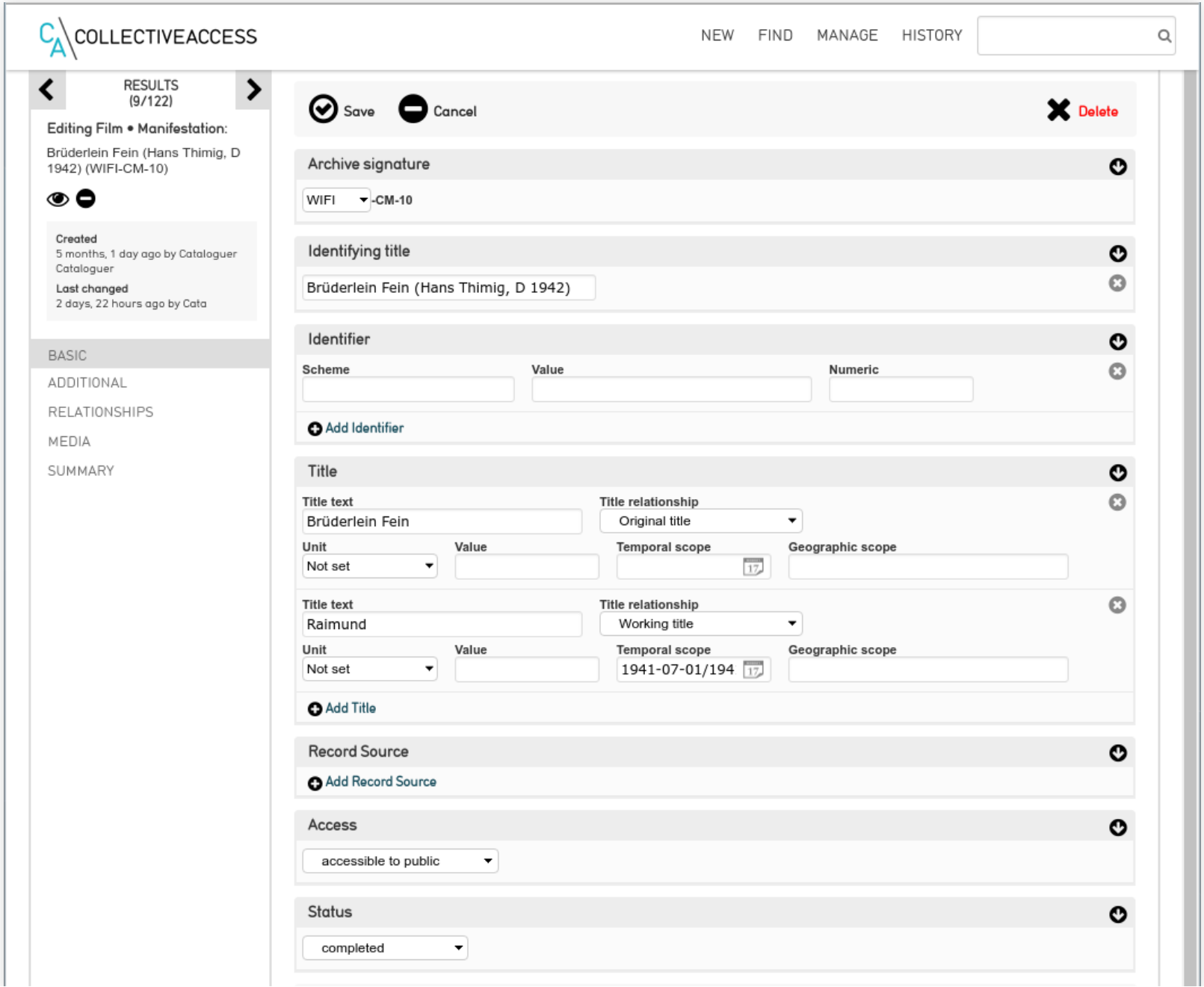
A loooong list...

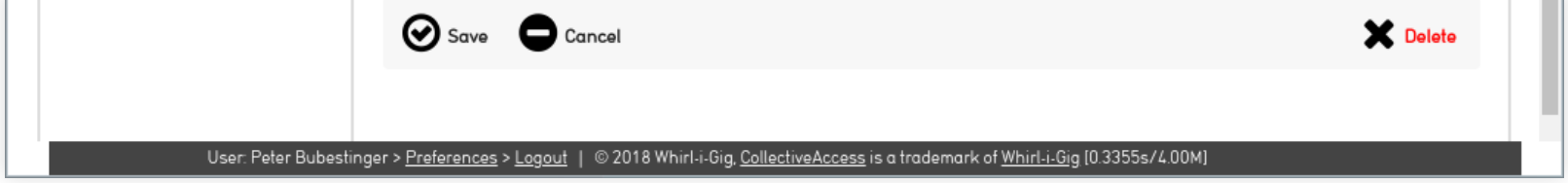
of not only OpenSource CMS:

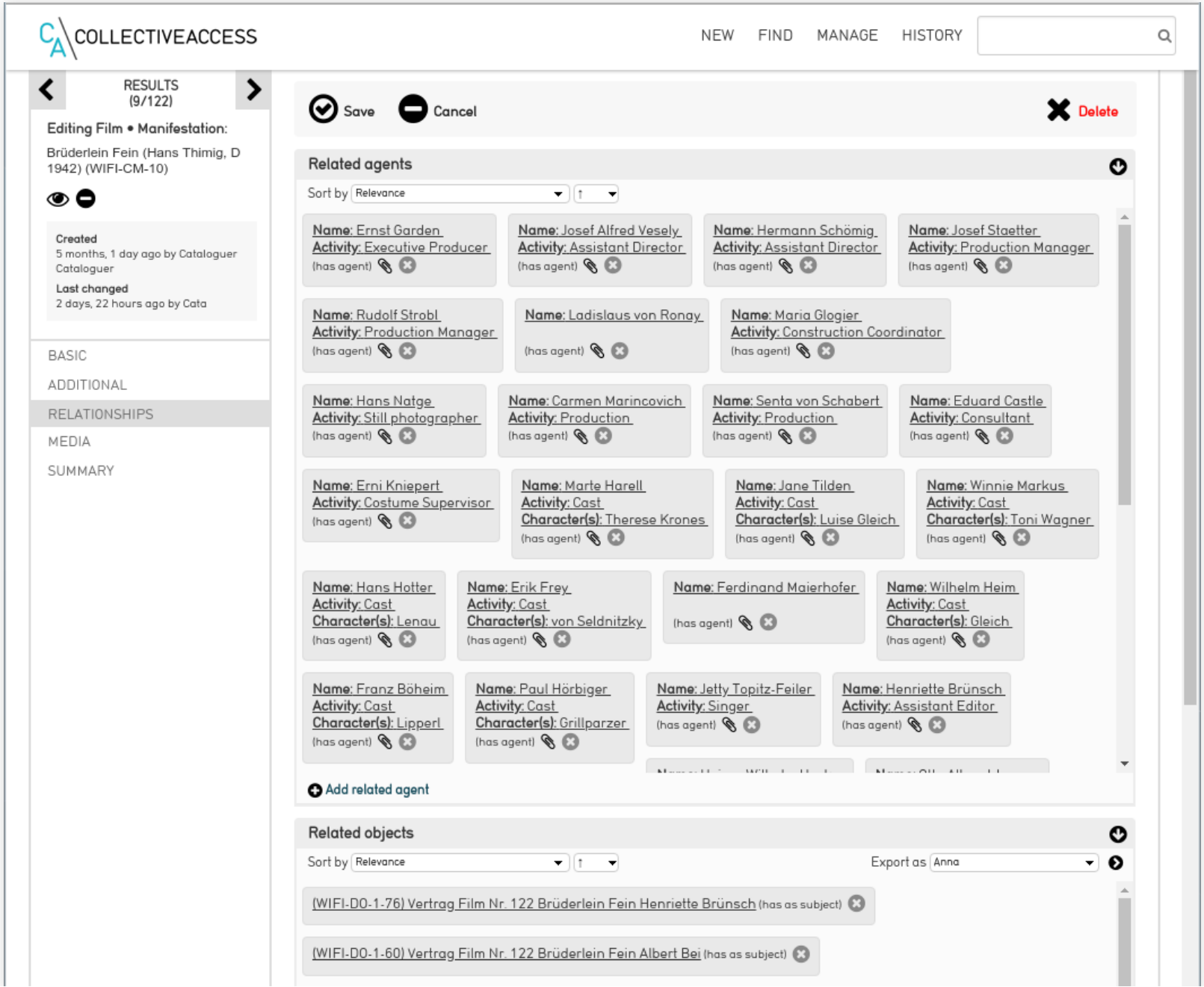
https://bits.ashleyblewer.com/blog/2017/08/09/collection-management-system-collection/

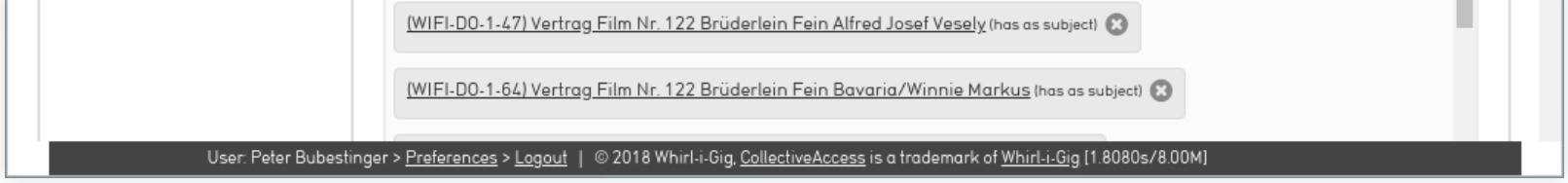

The Collection Management System Collection - Google Tabellen - Chromium
The Collection Management Syste
docs.google.com/spreadsheets/d/1cXOug3qM0pNNeD_wssiVEv9c0W1Y5I1VDTnSPTk7fb4/edit#gid=0
Apps
Jitsi EQZE
The Collection Management System Collection
Datei Bearbeiten Ansicht Einfügen Format Daten Tools Add-ons Hilfe
Freigeben
Anmelden
50%
Nur Kommentierzugriff
More info: https://ablwr.github.io/blog/2017/08/09/collection-management-system-collection/
Sheet1
Erkunden

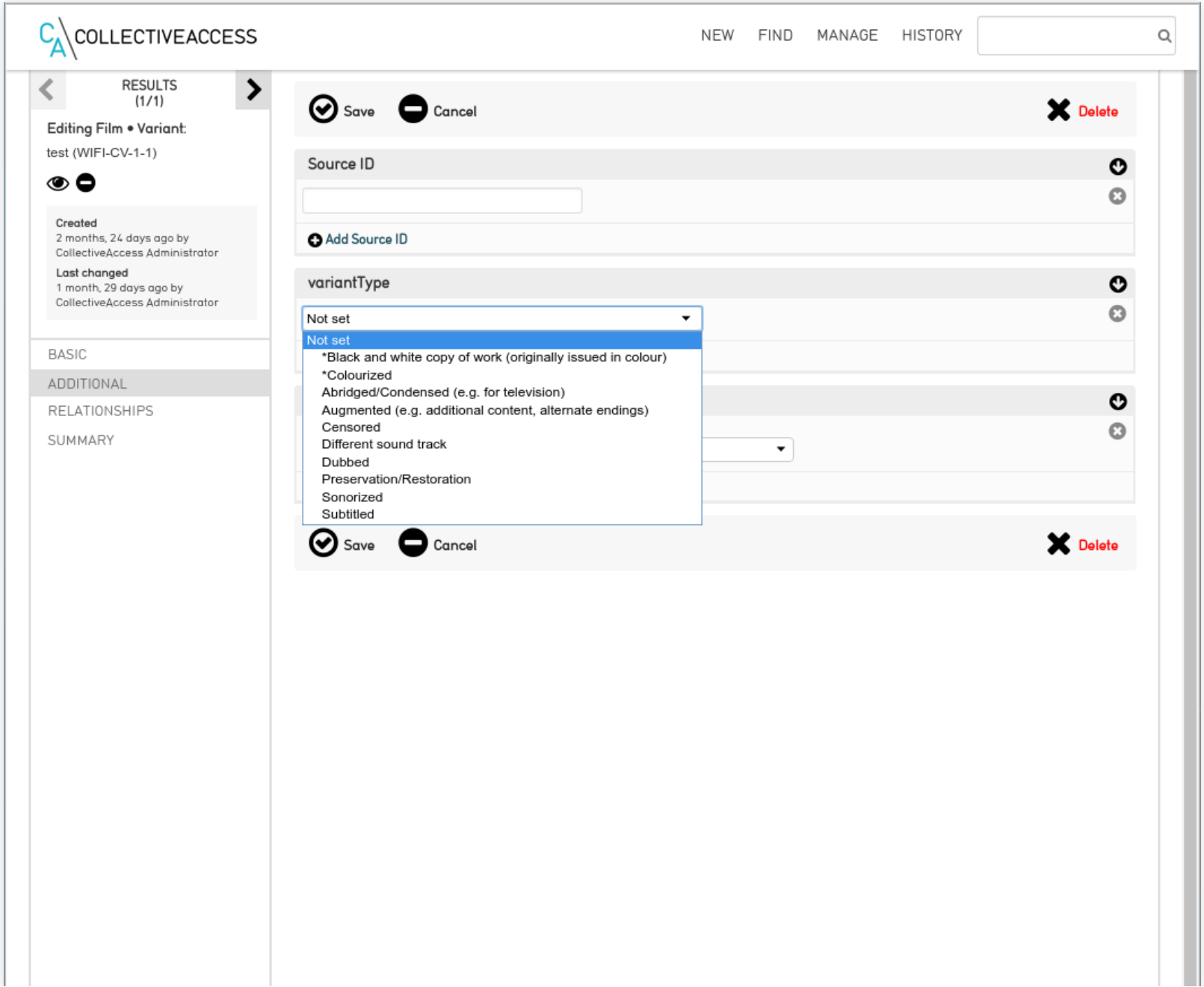
A look into: CollectiveAccess


COLLECTIVEACCESS
NEW
FIND
MANAGE
HISTORY
RESULTS
(9/122)
Editing Film • Manifestation:
Brüderlein Fein (Hans Thimig, D 1942) (WIFI-CM-10)
Created
5 months, 1 day ago by Cataloguer Cataloguer
Last changed
2 days, 22 hours ago by Cata
BASIC
ADDITIONAL
RELATIONSHIPS
MEDIA
SUMMARY
Save
Cancel
Delete
Archive signature
WIFI
-CM-10
Identifying title
Brüderlein Fein (Hans Thimig, D 1942)
Identifier
Scheme
Value
Numeric
Add Identifier
Title
Title text
Brüderlein Fein
Title relationship
Original title
Unit
Not set
Value
Temporal scope
Geographic scope
Title text
Raimund
Title relationship
Working title
Unit
Not set
Value
Temporal scope
1941-07-01/194
Geographic scope
Add Title
Record Source
Add Record Source
Access
accessible to public
Status
completed


Save
Cancel
Delete


COLLECTIVEACCESS
NEW
FIND
MANAGE
HISTORY
RESULTS (9/122)
Editing Film • Manifestation:
Brüderlein Fein (Hans Thimig, D 1942) (WIFI-CM-10)
Created
5 months, 1 day ago by Cataloguer Cataloguer
Last changed
2 days, 22 hours ago by Cata
BASIC
ADDITIONAL
RELATIONSHIPS
MEDIA
SUMMARY
Save
Cancel
Delete
Related agents
Sort by Relevance
Name: Ernst Garden
Activity: Executive Producer
(has agent)
Name: Josef Alfred Vesely
Activity: Assistant Director
(has agent)
Name: Hermann Schömig
Activity: Assistant Director
(has agent)
Name: Josef Staetter
Activity: Production Manager
(has agent)
Name: Rudolf Strobl
Activity: Production Manager
(has agent)
Name: Ladislaus von Ronay
(has agent)
Name: Maria Glogier
Activity: Construction Coordinator
(has agent)
Name: Hans Natge
Activity: Still photographer
(has agent)
Name: Carmen Marincovich
Activity: Production
(has agent)
Name: Senta von Schabert
Activity: Production
(has agent)
Name: Eduard Castle
Activity: Consultant
(has agent)
Name: Erni Kniepert
Activity: Costume Supervisor
(has agent)
Name: Marte Harell
Activity: Cast
Character(s): Therese Krones
(has agent)
Name: Jane Tilden
Activity: Cast
Character(s): Luise Gleich
(has agent)
Name: Winnie Markus
Activity: Cast
Character(s): Toni Wagner
(has agent)
Name: Hans Hotter
Activity: Cast
Character(s): Lenau
(has agent)
Name: Erik Frey
Activity: Cast
Character(s): von Seldnitzky
(has agent)
Name: Ferdinand Maierhofer
(has agent)
Name: Wilhelm Heim
Activity: Cast
Character(s): Gleich
(has agent)
Name: Franz Böheim
Activity: Cast
Character(s): Lipperl
(has agent)
Name: Paul Hörbiger
Activity: Cast
Character(s): Grillparzer
(has agent)
Name: Jetty Topitz-Feiler
Activity: Singer
(has agent)
Name: Henriette Brünsch
Activity: Assistant Editor
(has agent)
Add related agent
Related objects
Sort by Relevance
Export as Anna
(WIFI-DO-1-76) Vertrag Film Nr. 122 Brüderlein Fein Henriette Brünsch (has as subject)
(WIFI-DO-1-60) Vertrag Film Nr. 122 Brüderlein Fein Albert Bei (has as subject)


(WIFI-D0-1-47) Vertrag Film Nr. 122 Brüderlein Fein Alfred Josef Vesely (has as subject)
(WIFI-D0-1-64) Vertrag Film Nr. 122 Brüderlein Fein Bavaria/Winnie Markus (has as subject)
User: Peter Bubestinger > Preferences > Logout | © 2018 Whirl-i-Gig. CollectiveAccess is a trademark of Whirl-i-Gig [1.8080s/8.00M]


COLLECTIVEACCESS
NEW
FIND
MANAGE
HISTORY
RESULTS
(1/1)
Editing Film • Variant:
test (WIFI-CV-1-1)
Created
2 months, 24 days ago by
CollectiveAccess Administrator
Last changed
1 month, 29 days ago by
CollectiveAccess Administrator
BASIC
ADDITIONAL
RELATIONSHIPS
SUMMARY
Save
Cancel
Delete
Source ID
Add Source ID
variantType
Not set
Not set
*Black and white copy of work (originally issued in colour)
*Colourized
Abridged/Condensed (e.g. for television)
Augmented (e.g. additional content, alternate endings)
Censored
Different sound track
Dubbed
Preservation/Restoration
Sonorized
Subtitled
Save
Cancel
Delete


User: Peter Bubestinger > Preferences > Logout | © 2018 Whirl-i-Gig. CollectiveAccess is a trademark of Whirl-i-Gig [0.2711s/8.00M]

RESULTS
(1/2)

Editing Manifestation:

Day of the subgenius (M7272)

Created
1 month, 3 days ago by Peter Bubestinger

Last changed
1 second ago by Peter Bubestinger

BASIC
ADDITIONAL
RELATIONSHIPS
MEDIA
SUMMARY
LOG

Saved changes to Manifestation

Save Cancel Delete

Media representations

Scene at 1h25m15s PRIMARY

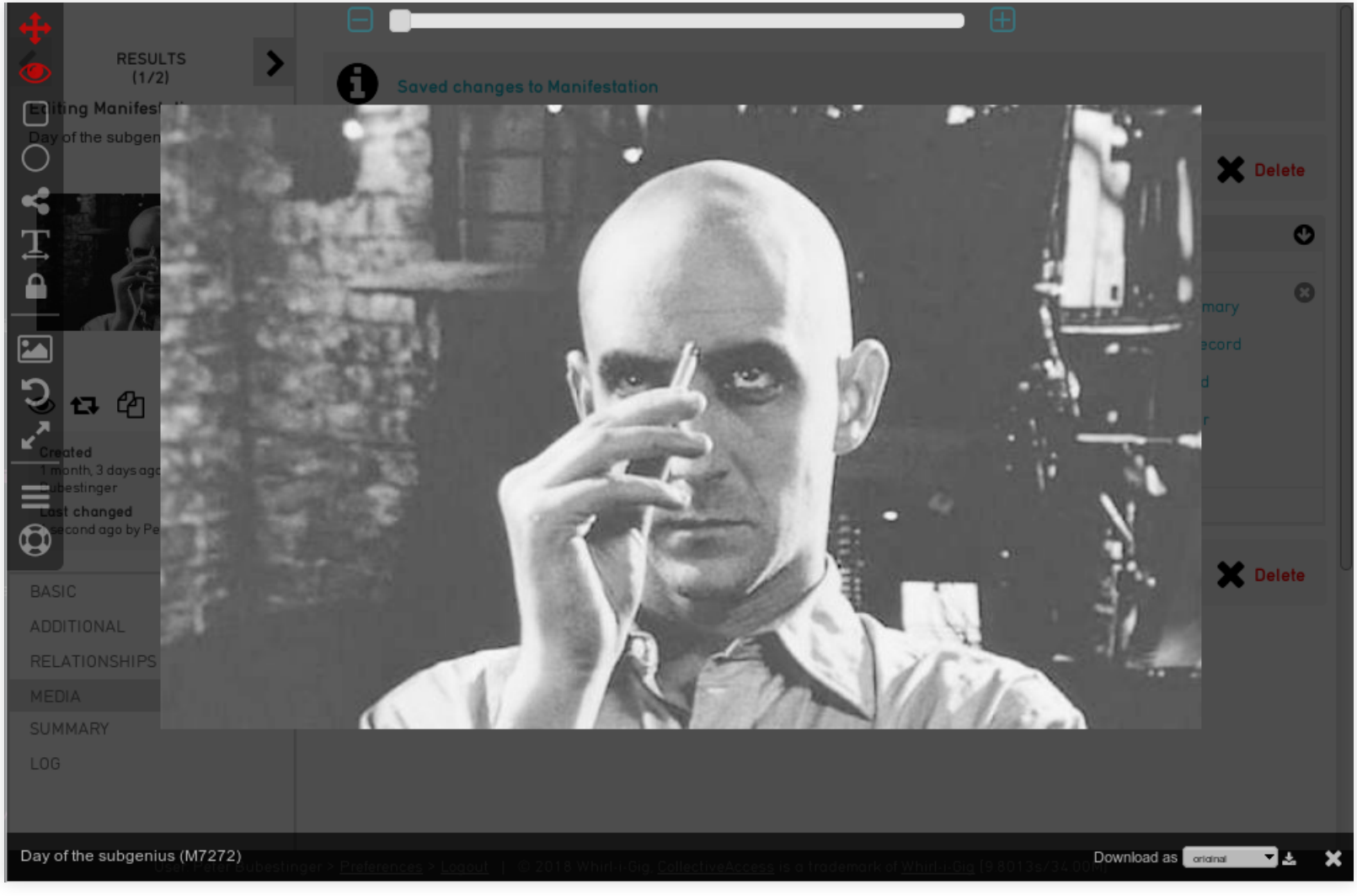
File name pi-3.jpg
Format JPEG; **Dimensions** 640p x 384p; 8 bpp; GRAY; 72ppi; 29.49kb;
Type generalView; **Access** accessible to public; **Status** new

Make primary
Edit full record
Download
Set center

Media metadata

Add representation ›

Save Cancel Delete


RESULTS
(1/2)
Saved changes to Manifestation
Delete
Delete
BASIC
ADDITIONAL
RELATIONSHIPS
MEDIA
SUMMARY
LOG
Created
Last changed
Day of the subgenius (M7272)
Download as
original

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:

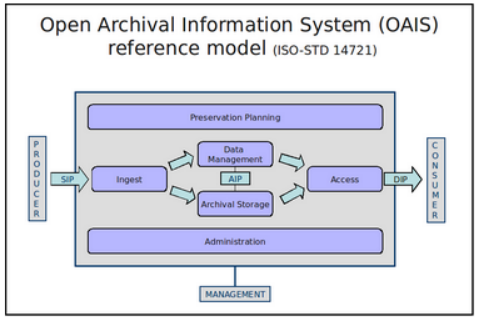
- Fixity creation
- Filename documentation & detox
- Virus checking
- Adding metadata
- Create preview images
- ...
- Adjust structure to in-house rules
- ... A popular example of such a system is "Archivematica"

Archivematica 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

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)

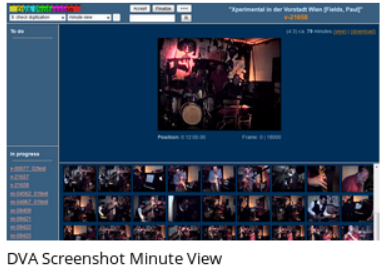
DVA-Profession

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:

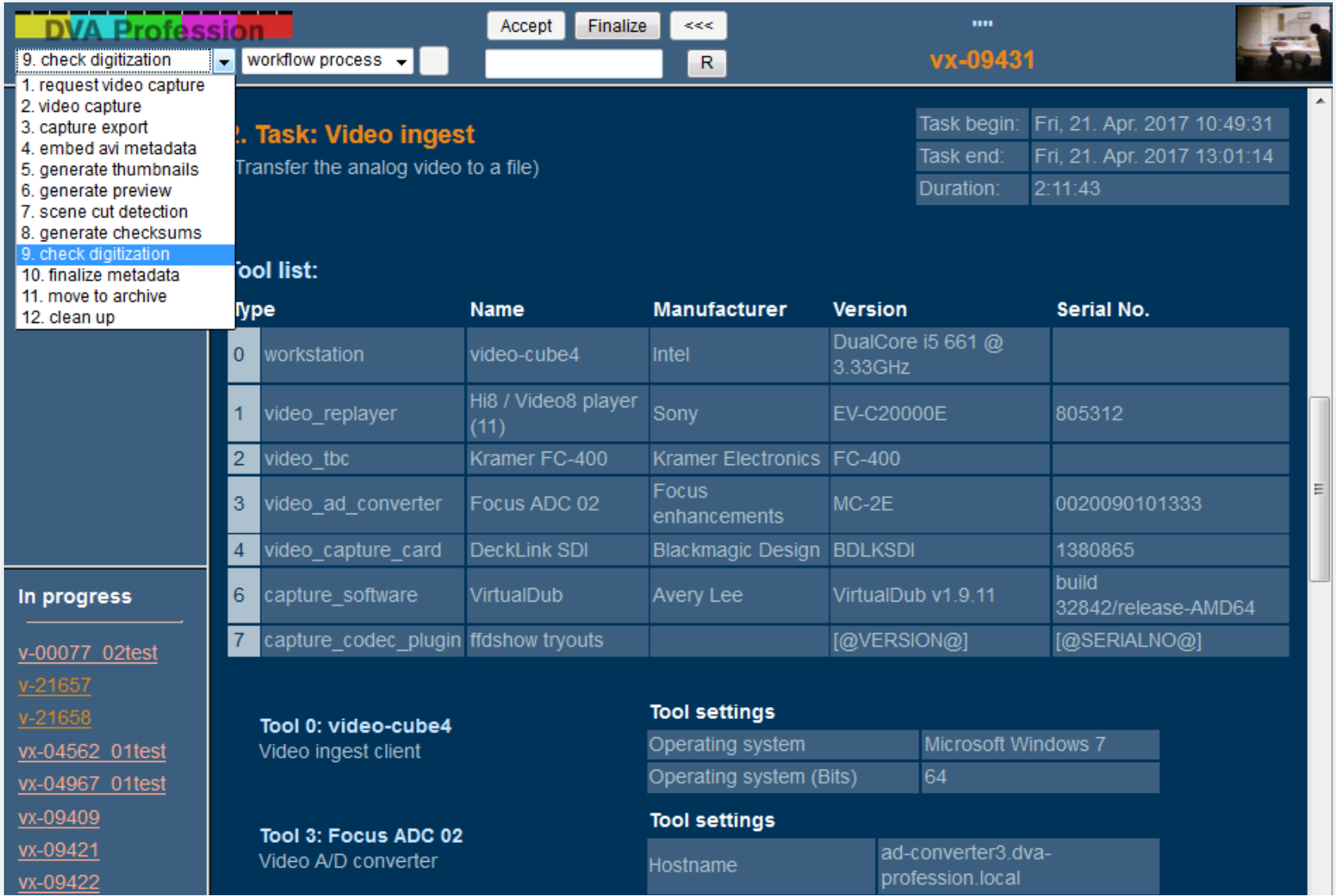
Summary

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 can be reconstructed years after the conversion of the original tape: software, tape replayers, A/D converters



DVA Screenshot Minute View

Workflow management

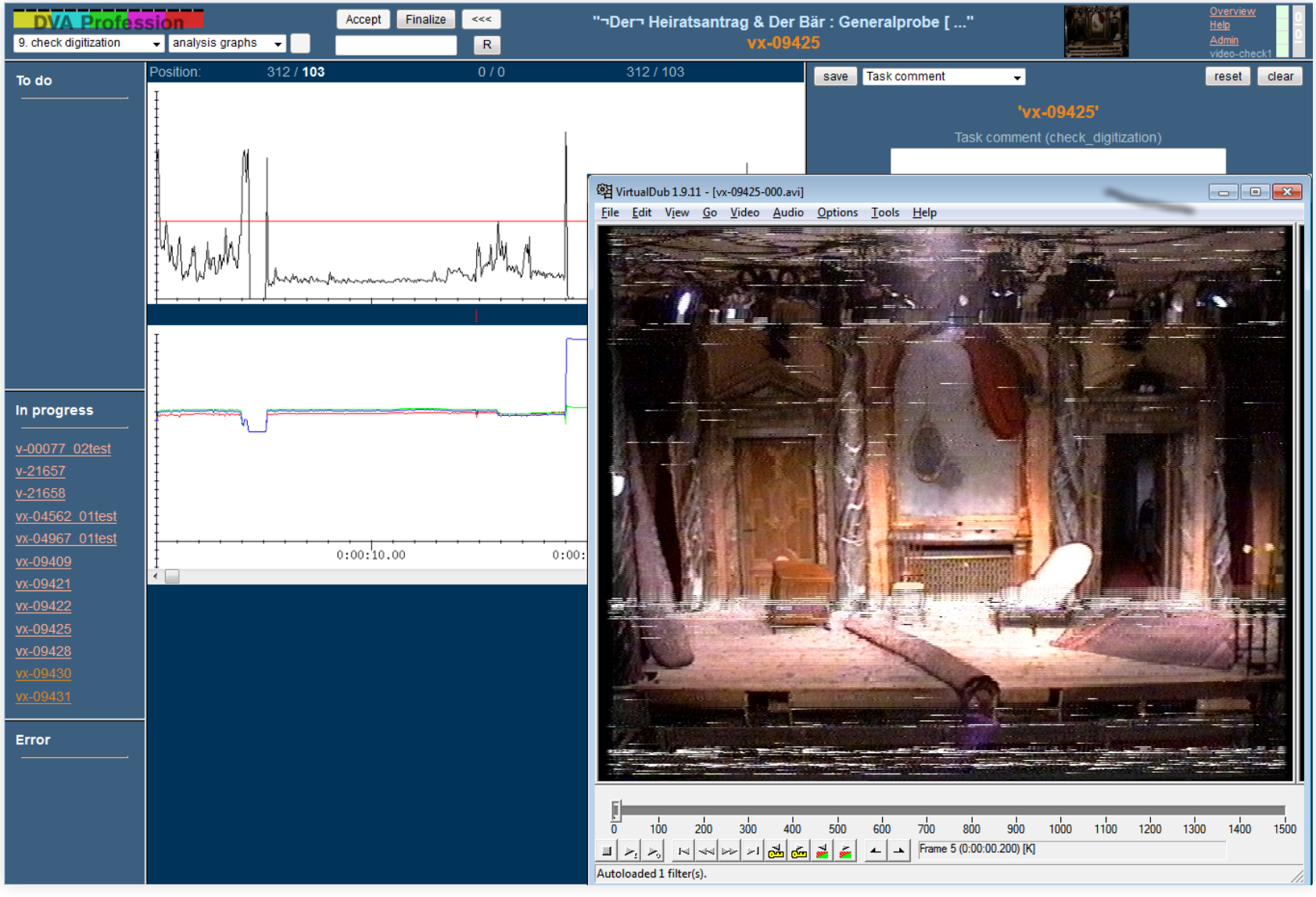


vx-09425
vx-09428
vx-09430
vx-09431

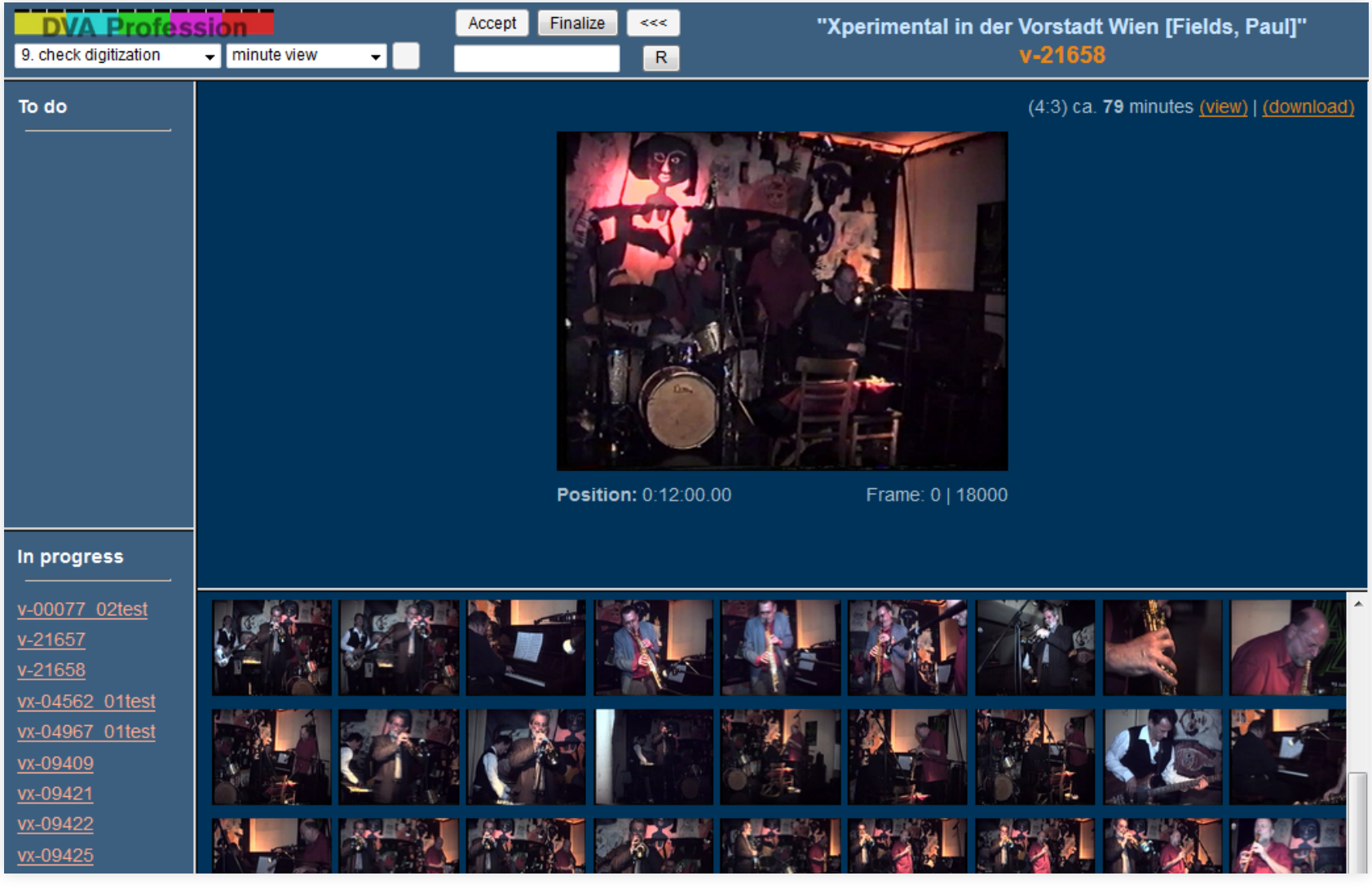
Error

Video input	Y/C
Video system	PAL
Genlock	Frame Sync
YUV level in	SMPTE
Genlock 75 ohm (terminated)	on
YUV level out	SMPTE
On screen display	off
VBI settings	Transparent
Contrast	127
Saturation	64
Hue (degrees)	0

Quality control


DVA Profession
9. check digitization
analysis graphs
Accept
Finalize
<<<
R
"¬Der¬ Heiratsantrag & Der Bär : Generalprobe [..."
vx-09425
Overview
Help
Admin
video-check1
To do
In progress
v-00077_02test
v-21657
v-21658
vx-04562_01test
vx-04967_01test
vx-09409
vx-09421
vx-09422
vx-09425
vx-09428
vx-09430
vx-09431
Error
Position:
312 / 103
0 / 0
312 / 103
0:00:10.00
0:00:
save
Task comment
reset
clear
'vx-09425'
Task comment (check_digitization)
VirtualDub 1.9.11 - [vx-09425-000.avi]
File Edit View Go Video Audio Options Tools Help
0 100 200 300 400 500 600 700 800 900 1000 1100 1200 1300 1400 1500
Frame 5 (0:00:00.200) [K]
Autoloaded 1 filter(s).

Other stuff

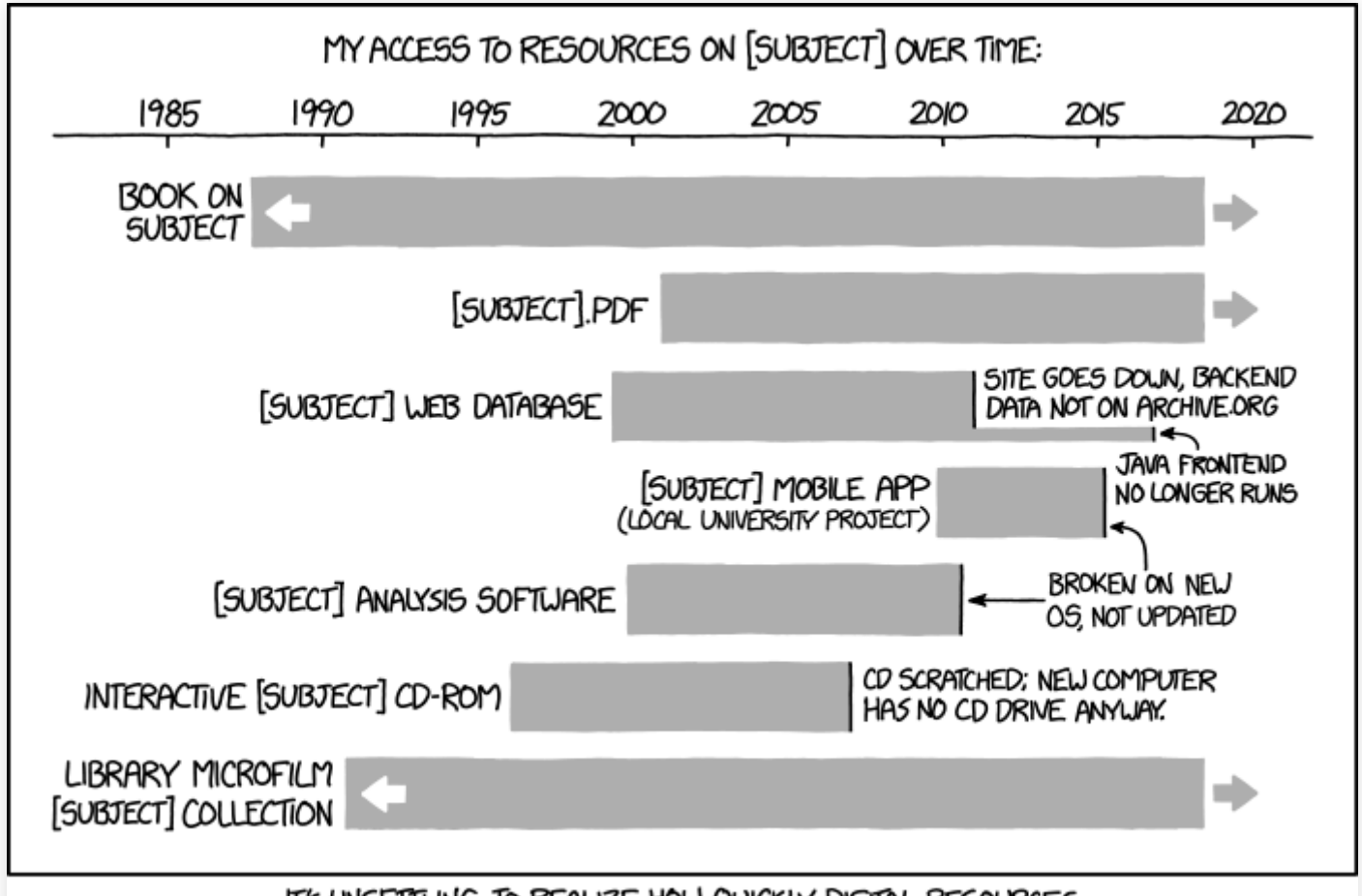

DVA Profession
Accept
Finalize
<<<
"Xperimental in der Vorstadt Wien [Fields, Paul]"
9. check digitization
minute view
R
v-21658
To do
(4:3) ca. 79 minutes (view) | (download)
Position: 0:12:00.00
Frame: 0 | 18000
In progress
v-00077_02test
v-21657
v-21658
vx-04562_01test
vx-04967_01test
vx-09409
vx-09421
vx-09422
vx-09425

Comments?

Questions?

Now for something completely different...

Migration



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

How long?

Eternal Migration

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

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

Migration Types

- Storage
- Format
- Software / platform / environment

Migration Types

Or more generically speaking:

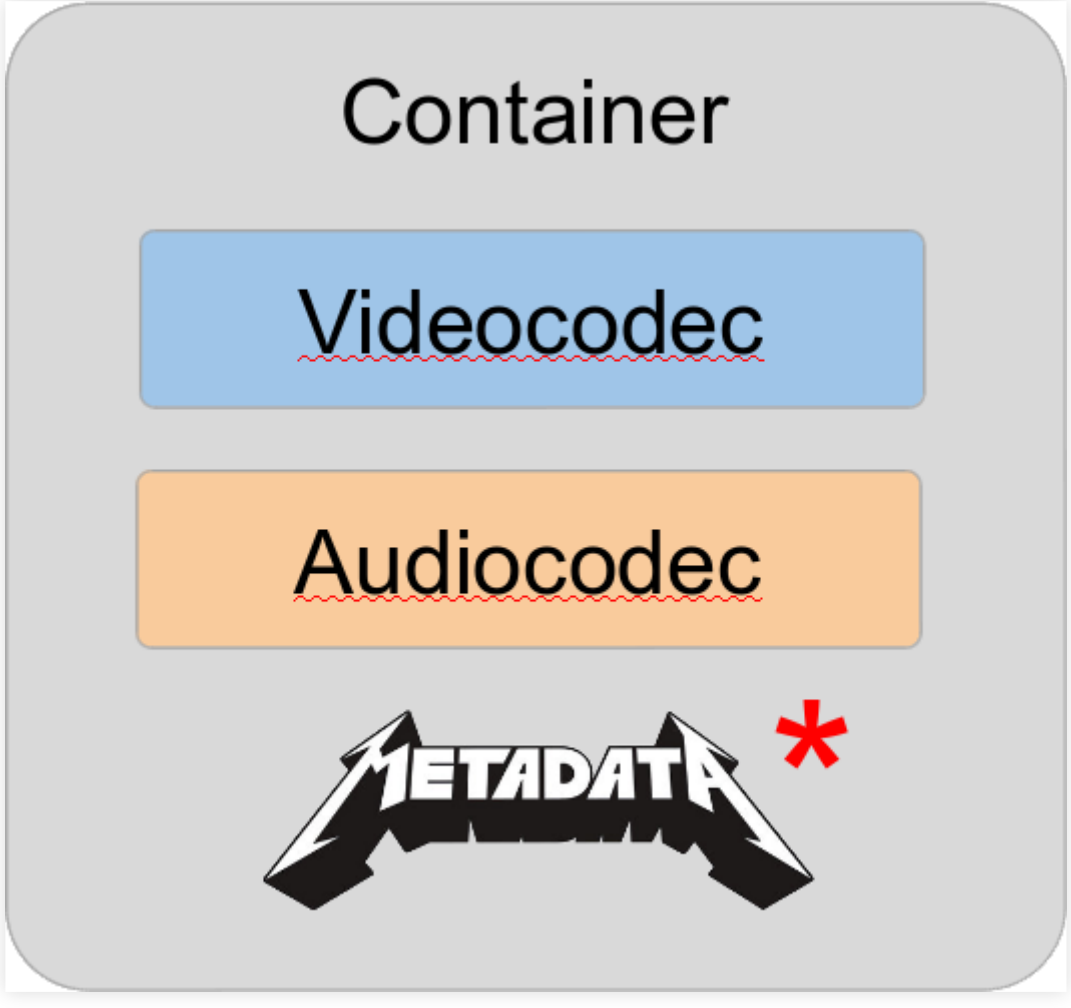
- Hardware
- Software
- File

Device Media




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

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?


MISSION:
MIGRATION

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?

Software / platform / environment

Might require:

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

Things can happen...

- Silent regression
- Unseen (meta)data changes
- Domino-effect: Forced updates of other things

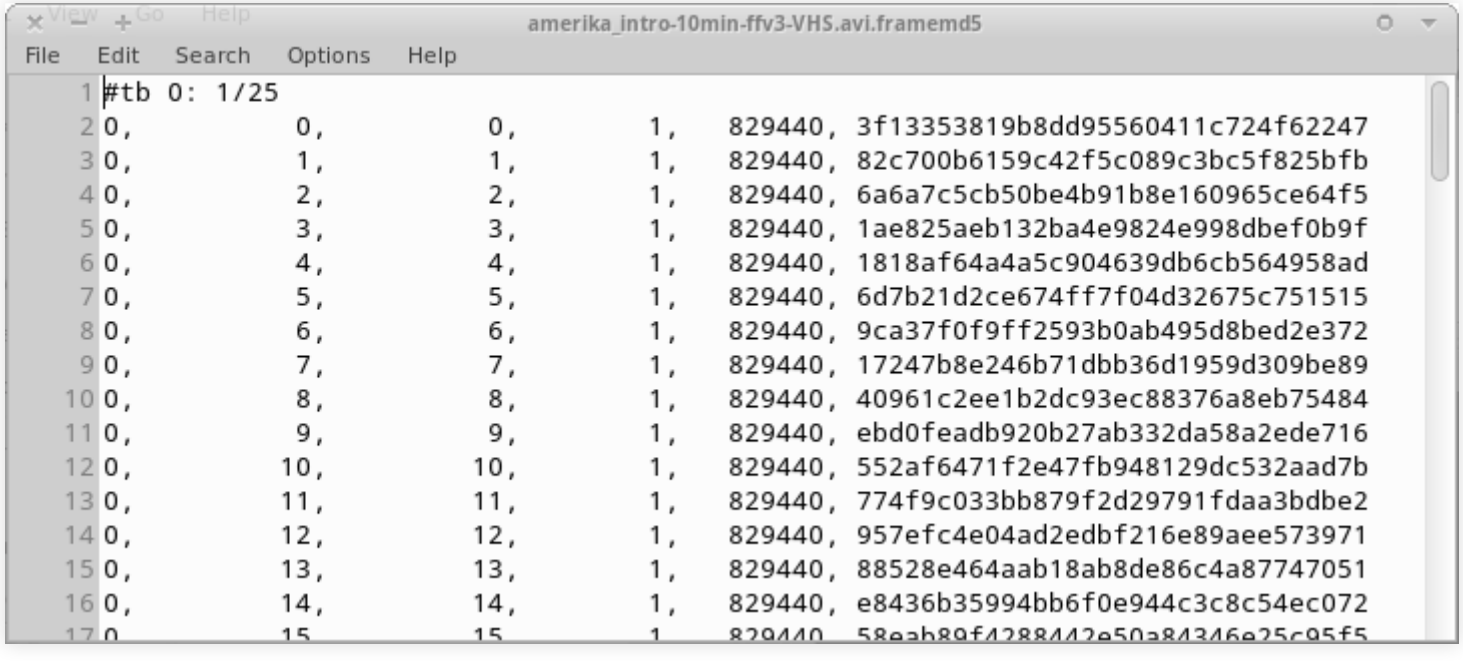
  I'm loving it! ®

- etc.

Media Format Migration

How to check integrity of format/codec migration?

FrameMD5


amerika_intro-10min-ffv3-VHS.avi.framemd5
File Edit Search Options Help
#tb 0: 1/25
0, 0, 0, 1, 829440, 3f13353819b8dd95560411c724f62247
0, 1, 1, 1, 829440, 82c700b6159c42f5c089c3bc5f825bfb
0, 2, 2, 1, 829440, 6a6a7c5cb50be4b91b8e160965ce64f5
0, 3, 3, 1, 829440, 1ae825aeb132ba4e9824e998dbef0b9f
0, 4, 4, 1, 829440, 1818af64a4a5c904639db6cb564958ad
0, 5, 5, 1, 829440, 6d7b21d2ce674ff7f04d32675c751515
0, 6, 6, 1, 829440, 9ca37f0f9ff2593b0ab495d8bed2e372
0, 7, 7, 1, 829440, 17247b8e246b71dbb36d1959d309be89
0, 8, 8, 1, 829440, 40961c2ee1b2dc93ec88376a8eb75484
0, 9, 9, 1, 829440, ebd0feadb920b27ab332da58a2ede716
0, 10, 10, 1, 829440, 552af6471f2e47fb948129dc532aad7b
0, 11, 11, 1, 829440, 774f9c033bb879f2d29791fdaa3bdbe2
0, 12, 12, 1, 829440, 957efc4e04ad2edbf216e89aee573971
0, 13, 13, 1, 829440, 88528e464aab18ab8de86c4a87747051
0, 14, 14, 1, 829440, e8436b35994bb6f0e944c3c8c54ec072

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...

Questions?

Comments?