

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

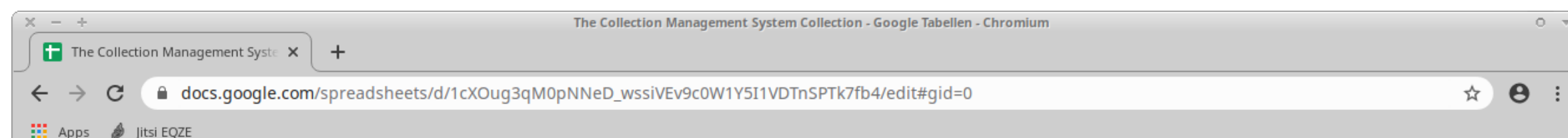
of some popular OpenSource CMS:

- [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/>



A look into: CollectiveAccess

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

User: Peter Bubestinger > [Preferences](#) > [Logout](#) | © 2018 Whirl-i-Gig. [CollectiveAccess](#) is a trademark of [Whirl-i-Gig](#) [9.8013s/34.00M]

11

12

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"

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 maintaining these tasks.

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 OAI

Open Archival Information System (OAIS) reference model (ISO-STD 14721)

```
graph LR
    PRODUCER -- SIP --> Ingest
    Ingest --> DM[Data Management]
    subgraph DM
        AIP[AIP]
    end
    DM --> AS[Archival Storage]
    AS --> Access
    Access -- DIP --> CONSUMER
    PP[Preservation Planning] --- Ingest
    PP --- DM
    PP --- AS
    PP --- Access
    Admin[Administration] --- Ingest
    Admin --- DM
    Admin --- AS
    Admin --- Access
    Mgt[MANAGEMENT] --- Ingest
    Mgt --- DM
    Mgt --- AS
    Mgt --- Access
```

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

[What is Archivematica?](#)

- [Open source OAI](#)
- [Lowering the barriers to best-practice digital preservation](#)

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)

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.

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.

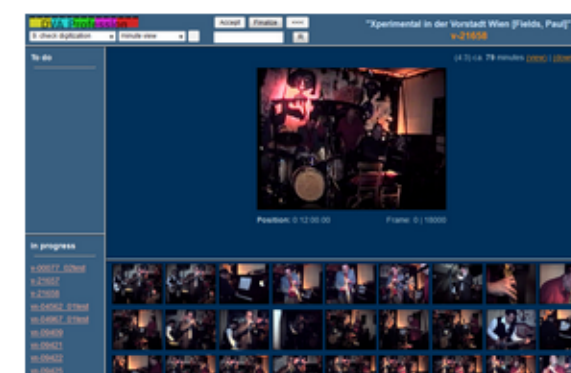
Respecting this massive intervention, accurate documentation of every step in the process of conversion is necessary.

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 tapes: software, tape players, A/D converters

Auf dieser Seite:



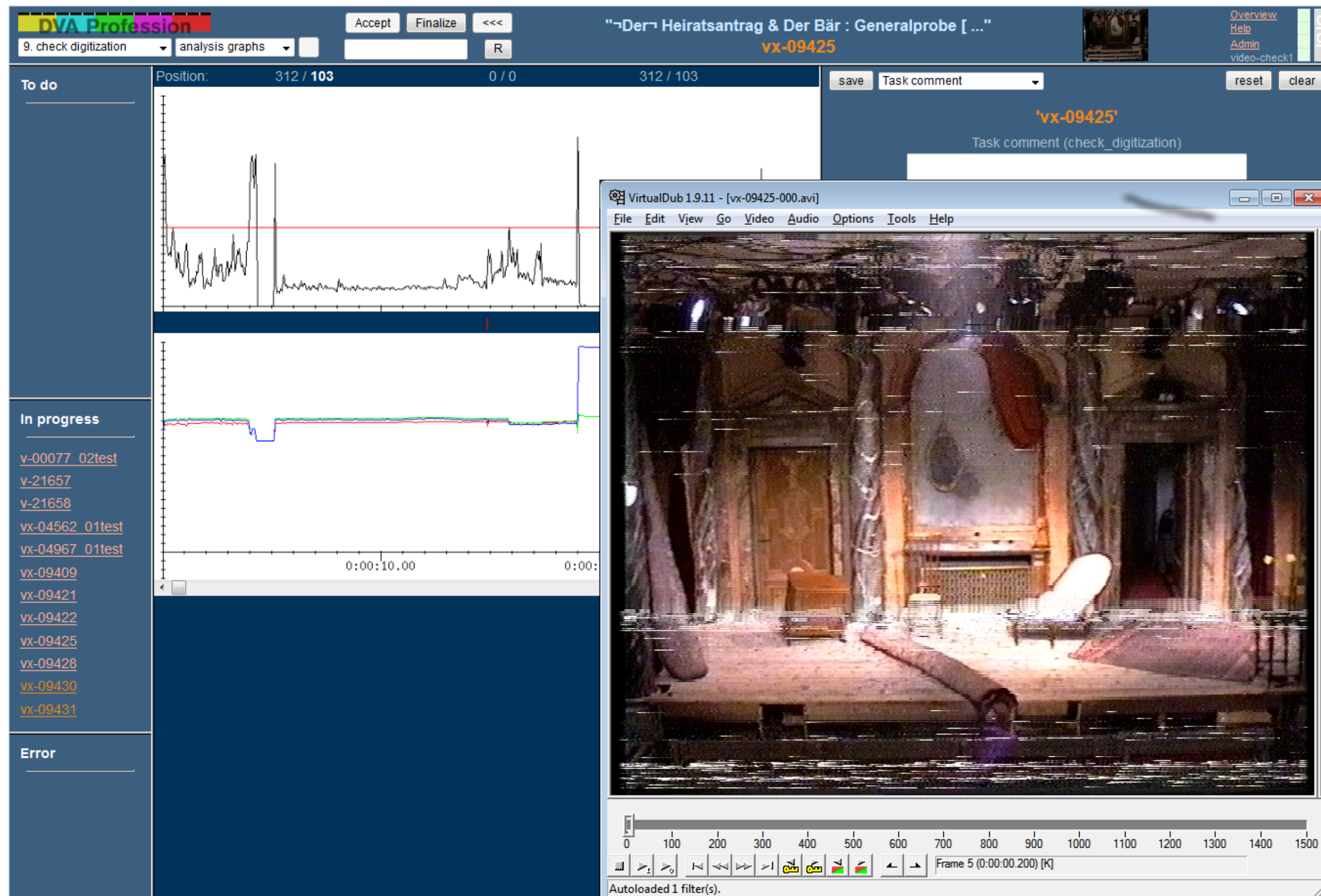
DVA Screenshot Minute View

Workflow management

Quality control

Speaker notes

No notes on this slide.



Other stuff

DVA Profession

9. check digitization

minute view

Accept

Finalize

<<<

R

"Xperimental in der Vorstadt Wien [Fields, Paul]"
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](#)

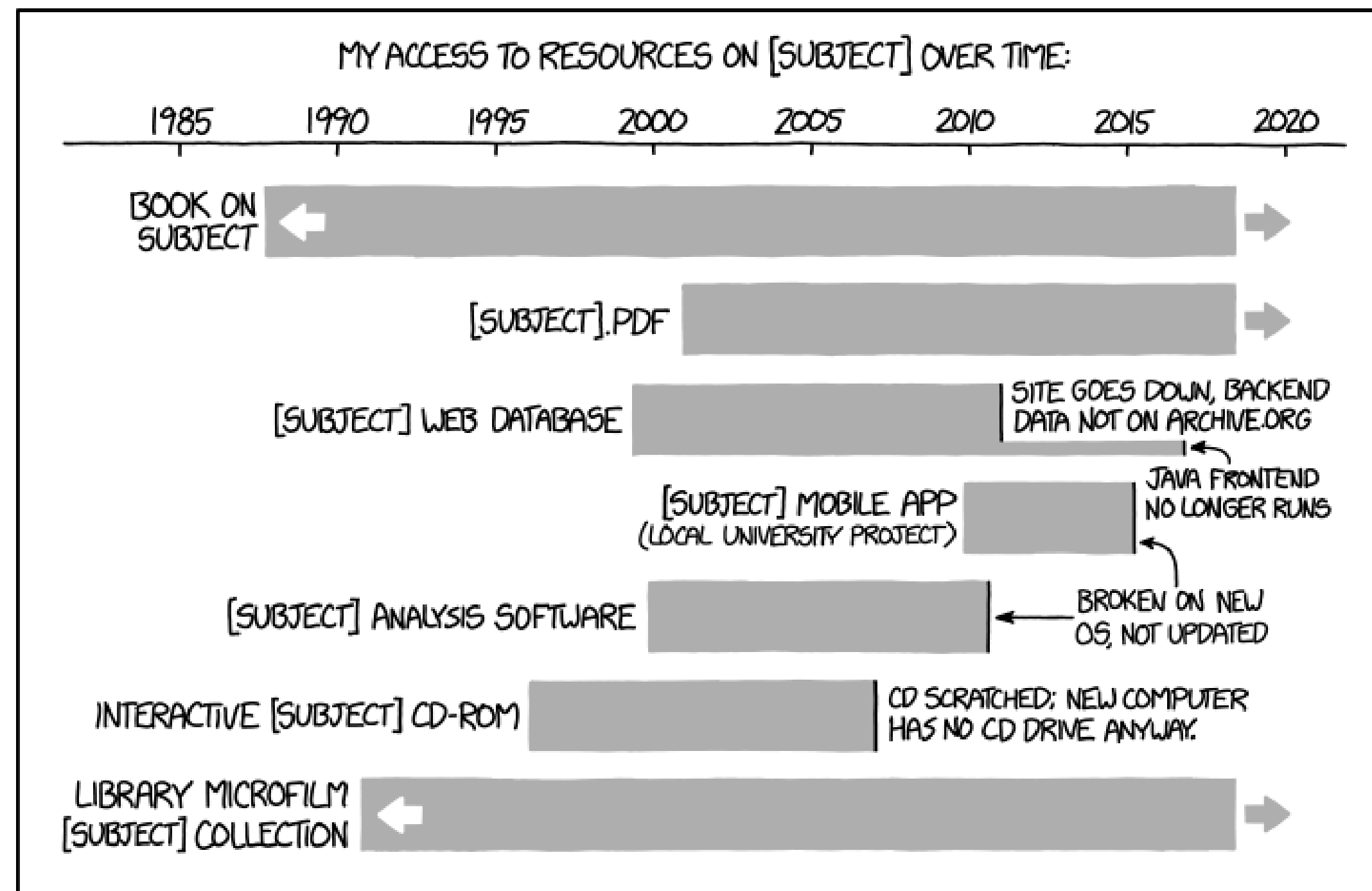
[vx-09425](#)



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?

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.

Migration Types

- Storage
- Format
- Software / platform / environment

Migration Types

Or more generically speaking:

- Hardware
- Software
- File

Device Media

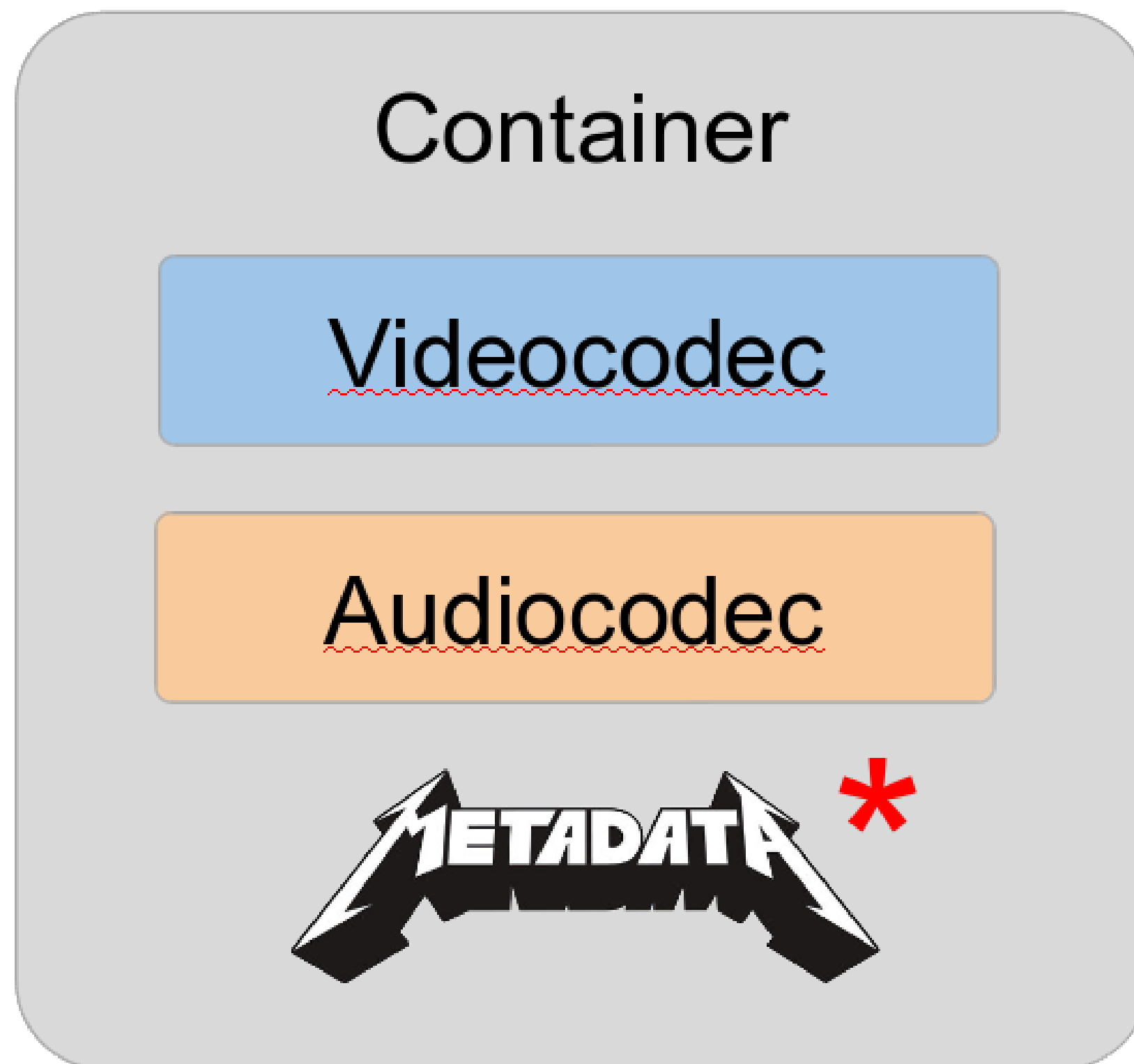


...or other media types (holograms, DNA, etc)

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?

Data Format(s)



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 ;)

Device Format Migration

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

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.

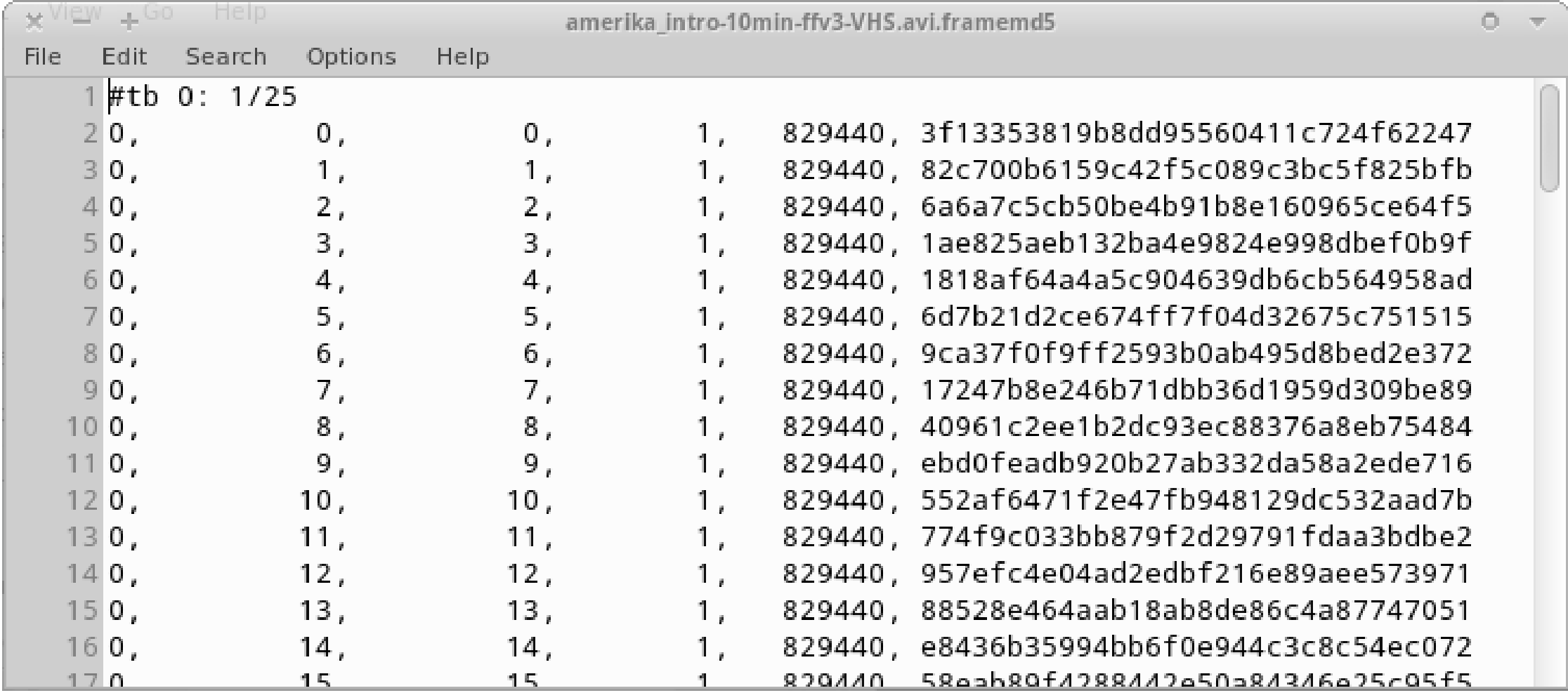
Things can happen...

- 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/codecs migration?

FrameMD5



A screenshot of a text editor window titled 'amerika_intro-10min-ffv3-VHS.avi.frameemd5'. The window has a menu bar with 'File', 'Edit', 'Search', 'Options', and 'Help'. The text content shows a list of 17 frames, each with a frame number, a CRC value, and a MD5 hash. The first frame is highlighted with a blue background.

Frame	CRC	MD5
1	0	3f13353819b8dd95560411c724f62247
2	0	82c700b6159c42f5c089c3bc5f825bfb
3	0	6a6a7c5cb50be4b91b8e160965ce64f5
4	0	1ae825aeb132ba4e9824e998dbef0b9f
5	0	1818af64a4a5c904639db6cb564958ad
6	0	6d7b21d2ce674ff7f04d32675c751515
7	0	9ca37f0f9ff2593b0ab495d8bed2e372
8	0	17247b8e246b71dbb36d1959d309be89
9	0	40961c2ee1b2dc93ec88376a8eb75484
10	0	ebd0feadb920b27ab332da58a2ede716
11	0	552af6471f2e47fb948129dc532aad7b
12	0	774f9c033bb879f2d29791fdaa3bdbe2
13	0	957efc4e04ad2edbf216e89aee573971
14	0	88528e464aab18ab8de86c4a87747051
15	0	e8436b35994bb6f0e944c3c8c54ec072
16	0	58e3b80f4288442e50a84346e25c95f5
17	0	

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

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)

Obsolescence monitoring

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

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?