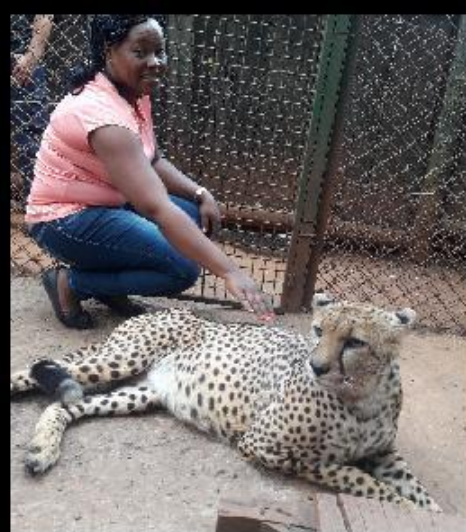




Digitization of Multimedia assets

The Digital Memory project





OUTLINE

- PROCEDURE OF DIGITIZATION
- AUDIO, PHOTO AND VIDEO at CERN
- EXPLOITATION OF THE “OLD” DIGITAL MULTIMEDIA

The Main Steps

- **Estimate** content and cost → get funding
- **Collect** and Organize your Content
 - Identify all the locations of the objects / group everything together
 - Make sure you collect all the existing material of your institute: “more is less”
 - Differentiate material already catalogued and the new material
- **Inventory**
- **Digitize**: Call for Tender/ select a bidder (if not in house)
- **Follow up** the digitization
- **Control** Quality of metadata and data
- **Upload** to your Repo
- **[Enrich** the metadata]

The Physical Supports

- **STORAGE**

- Identify all the locations of the objects
- Collect ALL the existing material:
 - “more (tape) is less (money)”

- **IDENTIFICATION**

- Identify the carrier type, group them by types
 - Still images: negatives, slides, medium and large formats
 - Moving images: listing of [all types](#)
- Priority from best to worst quality: *D3, Digital Betacam, DVCAM, DV/miniDV, DVD, Betacam SP, Betacam, 1 inch C, U-Matic S/SP, U-Matic, VHS, Video 8, Hi 8*
 - E.g if same movie on DVCAM and VHS: digitize DVCAM



The Procedure

- Group the material by types : **Barcode** on card boxes
- Use **Stickers** to make categories:
 - To be trashed
 - Carriers with no labels
 - Carriers duplicates, to put a copy aside
 - Carriers not catalogued
 - Carriers already catalogued
 - Carriers already catalogued with a video file
- Create new records if the material is NOT catalogued
- Optional: Add **Id abels** or **QR-Code** to the carriers, pointing to the record in your catalog <https://www.the-qr-code-generator.com/>



The Inventory: Metadata Update

- Before Digitization
 - Add the barcode value for the card box
 - Add the original physical location of the carrier
 - Each record must have an Id that is also printed on the carrier
- Digitization: File naming convention
 - Using record id - if qr-codes are used
 - Using an extract of your repository, containing the id on the label
- Quality Assurance:
 - Metadata: upload listings of digitized material to review it (OpenRefine)
 - Metadata will contain the URL to the digitized file
 - Data: check quality by random visualisation





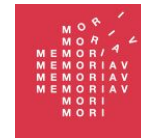
The Audio recordings collection

- Stored in the CERN Archive
 - 7'564 Carriers in 540 boxes of 1'740 Kg
 - 2'733 C60 & C90 cassettes
 - 4'831 open reel tapes
- Speeches of official committees since 1956
- Very restricted access policy
- Bidding
 - In 2016: ~17.7 €/carrier (1'695 carriers at 30'000 euros with $\frac{2}{3}$ of cassettes)
 - In 2018: ~11 €/carrier (6'869 carriers at 75'000 euros with $\frac{2}{3}$ of tapes)
- Options:
 - Formats: .wav uncompressed (48kHz&16 bit) + .ogg and .mp3 (320 kbps)
 - Hum removal, trimming, speed correction, channel config, etc.





Films and Audiovisual collections



- 5'767 rushes, documentaries and seminars recordings
 - public/restricted/copyrighted
- Complex initial situation
 - 120 films & 13 different carrier types
 - from D3 to HDCAMs or 1 inch C tapes
- Systematic Inventory & QR-coding
- Speech to text option tests
 - mitigated conclusions so far
- QA run by Contractor and CERN
 - ~10% of “dead” supports
- Bidding: ~23.7 €/carrier (121'000 €)

Originals		Preservation Master	Access Master	Access Copy
35mm films	Wrapper:	.mkv	.mov	.mp4
	Video codec:	FFV1 - 10 bits RGB	Apple ProRes 422 HQ	H.264 @ 5Mbps
	Audio codec:	24 bits PCM, 48kHz	24 bits PCM, 48kHz	16 bits AAC, 44.1kHz, 256kbps
	Definition / Aspect ratio:	4096x? / Original	4096x? / Original	1920x1080 / Pillar-letterbox
16mm films	Wrapper:	.mkv	.mov	.mp4
	Video codec:	FFV1 - 10 bits RGB	Apple ProRes 422 HQ	H.264 @ 5Mbps
	Audio codec:	24 bits PCM, 48kHz	24 bits PCM, 48kHz	16 bits AAC, 44.1kHz, 256kbps
	Definition / Aspect ratio:	2048x? / Original	2048x? / Original	1920x1080 / Pillar-letterbox
Analogue and digital SD video	Wrapper:	.mkv	.mov	.mp4
	Video codec:	FFV1 - 10 bits YCbCr	Apple ProRes 422 LT - SD profile	H.264 @ 1Mbps
	Audio codec:	24 bits PCM, 48kHz	24 bits PCM, 48kHz	16 bits AAC, 44.1kHz, 256kbps
	Definition / Aspect ratio:	7x576 / Original	7x576 / Original	640x360 / Pillar-letterbox
All		Maintain the original recording standard, interlacing, number of audio channels and auxiliary information such as original timecode	Maintain the original recording standard, interlacing, number of audio channels and auxiliary information such as original timecode	Progressive scan at 25 frames per second. Maintain the original number of audio channels and auxiliary information such as original timecode



Photos: Slides & Negatives

- Black & White : 120'298 images organized in 13'612 albums (scanned in 2014)
- Colour: 299'265 images (on-going)
 - 178'618 24x36 negatives
 - 38'844 24x36 slides
 - 81'803 Medium & Large formats
- Keys:
 - TIFF 48 bits (RGB) & 4800 ppi @ 24x36 size
 - File naming pattern to enable album identification
- Bidding: ~0.35 €/image (62'000 € for slides/ML & 47'000 € for negatives)

PhotoLab Archives

Search 19,094 records for:

[Search Tips](#)[Advanced Search](#)

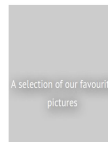
☐ search also PhotoLab Archives Images

Photolab Archives images

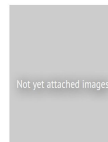
Please note: the Photolab Archive images are currently being scanned, uploaded and automatically matched with old album records. Inspected manually to check for errors. You can use the links below to see various sets of images:



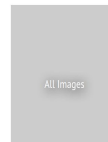
Inspecting Albums



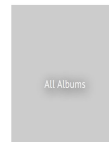
A selection of our favourite pictures



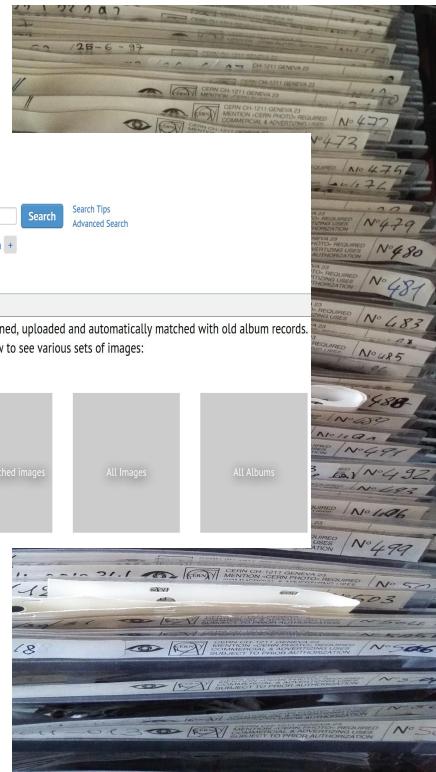
Not yet attached images



All Images



All Albums

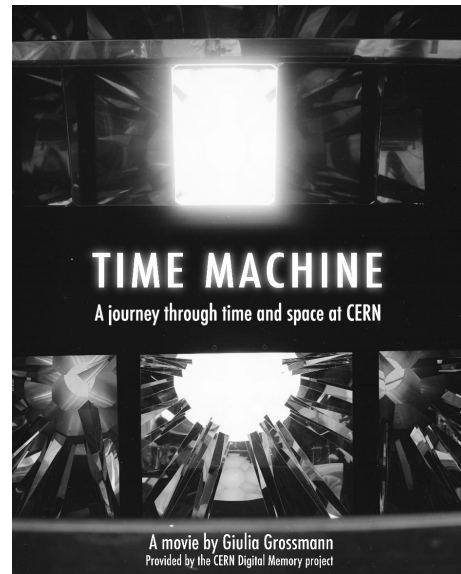


[Surprise!](#)



The ending point

- All multimedia files to be available by the end of 2018:
 - ~350'000 access files → up to ~ 3'150'000 files with all the formats, replicas and md5
 - > 500 Tb : less than 3% of CERN Data Cloud
- Entering CERN Document Server Multimedia Repo
- Gaining new visibility
 - Film-maker to exploit new “old” content
 - Interest from Museums
 - Exhibitions - collaboration with galleries
- And what about the good old tapes ?



[Trailer !](#)