



# 'DIY' DIGITAL PRESERVATION FOR VIDEO

IDENTIFY – SELECTION – ORGANISE - STORAGE

# AGENDA FOR TODAY

**Session 1: Hello + Housekeeping**

**Session 2: Introduction to General Digital Preservation**

**Session 3: Identify + Selection**

**Session 4: Organise + Storage**

## SESSION 1

# WHO ARE WE & WHY ARE WE DOING THIS?

- We are the four London Cohorts of the 'BRIDGING the DIGITAL GAP' traineeship scheme.
- Collaborating since March 2020 on 'DIY' Digital Preservation workshop series, based on a different medium (software, video, audio & web-archiving).
- Sponsored by the UK National Archives and funded by the National Lottery Heritage Fund.
- Supported by: TfL, LMA, UaL, UoW.
- Spread awareness about how easy it is to preserve and manage your own personal digital files.

Hoping to inspire you to take action!



## SESSION 2

# WHAT IS DIGITAL PRESERVATION ?

*" IT IS ABOUT A SET OF ACTIVITIES NEEDED TO ENSURE THAT DIGITAL CONTENT CAN BE ACCESSED FOR AS LONG AS NEEDED. "*

it's an on-going process  
and not a one-off !

it may mean perpetuity  
(e.g. lasting forever, no end).

*taken from DPC*

# WHAT IS A BITSTREAM ?

SWITCH TO MICROSOFT WHITEBOARD

## SESSION 2

# LET'S TALK ABOUT DIGITAL INFORMATION FIRST !

- Digital information:
  - ✓ drives our economy.
  - ✓ helps develop our cultural experiences.
  - ✓ connects our communities together.
- Born-digital material such as text, image & video are items that are created and managed in digital form.

## SESSION 2

# THE MAIN QUESTION TO ASK HERE IS:

*“But what if all or part of this information disappears ?”*

I have learnt during my traineeship so far:

- Our digital assets/collections are just as valuable as traditional information (e.g. maps, books, etc).

By depending on technology, it can put our digital items at risk too!

## SESSION 2

# THE CHALLENGE IS:

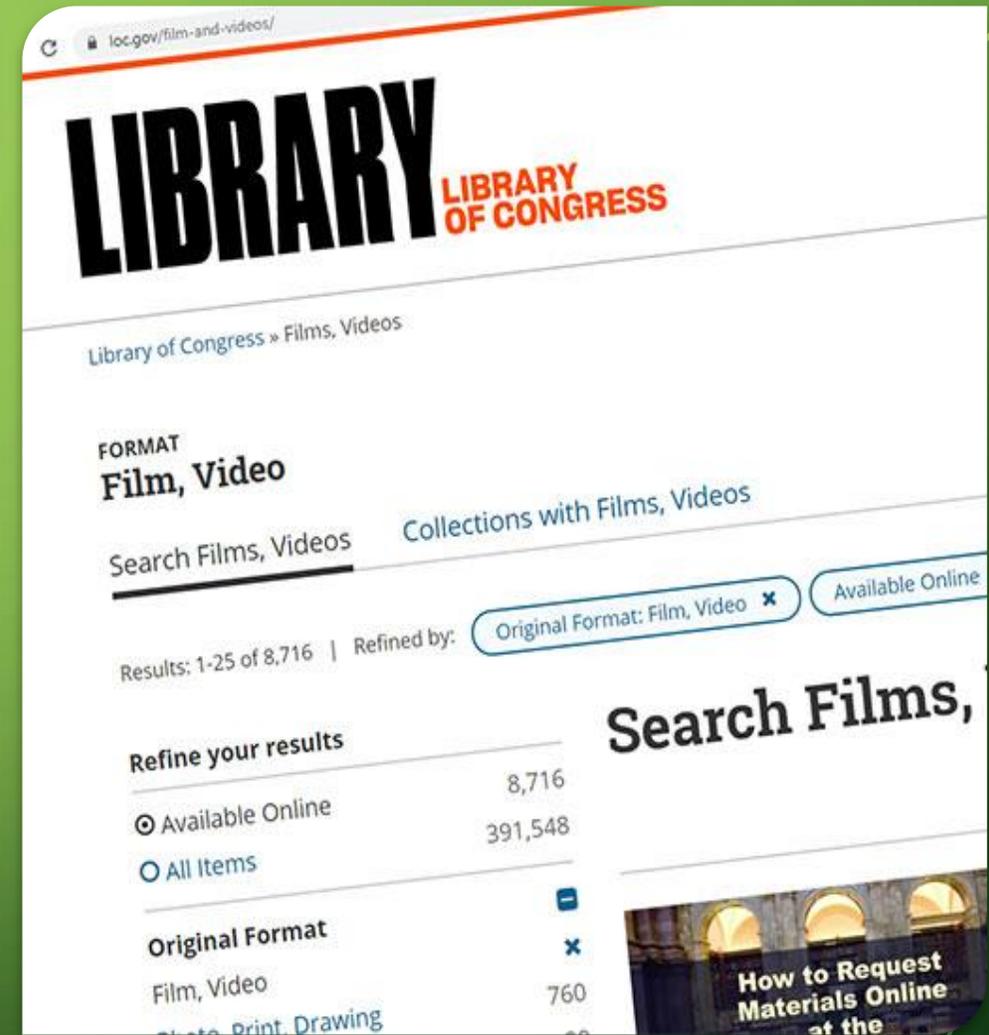
*“ What are the minimal requirements for responsible management of born-digital video ? ”*

**THIS IS WHERE DIGITAL PRESERVATION  
ENTERS !!!**

## SESSION 3

**STRATEGY:** (TAKEN FROM LIBRARY OF CONGRESS  
[HTTPS://WWW.LOC.GOV/](https://www.loc.gov/))

1. **Identify** what video types you want to save (e.g. avi, mp4, mkv, flv, xvid format).
  - Where are they saved? On your computer PC/Mac?
  - Is it stored on your Cloud platform online?
  - Is it scattered around in several folders or external hard drive / flash drives?
  - It may not be practical to save every single file.
2. **Decide or select** on what is most important to you to save.
  - This will help you reduce the amount of files that you have accumulated.



website screenshot of Loc.gov

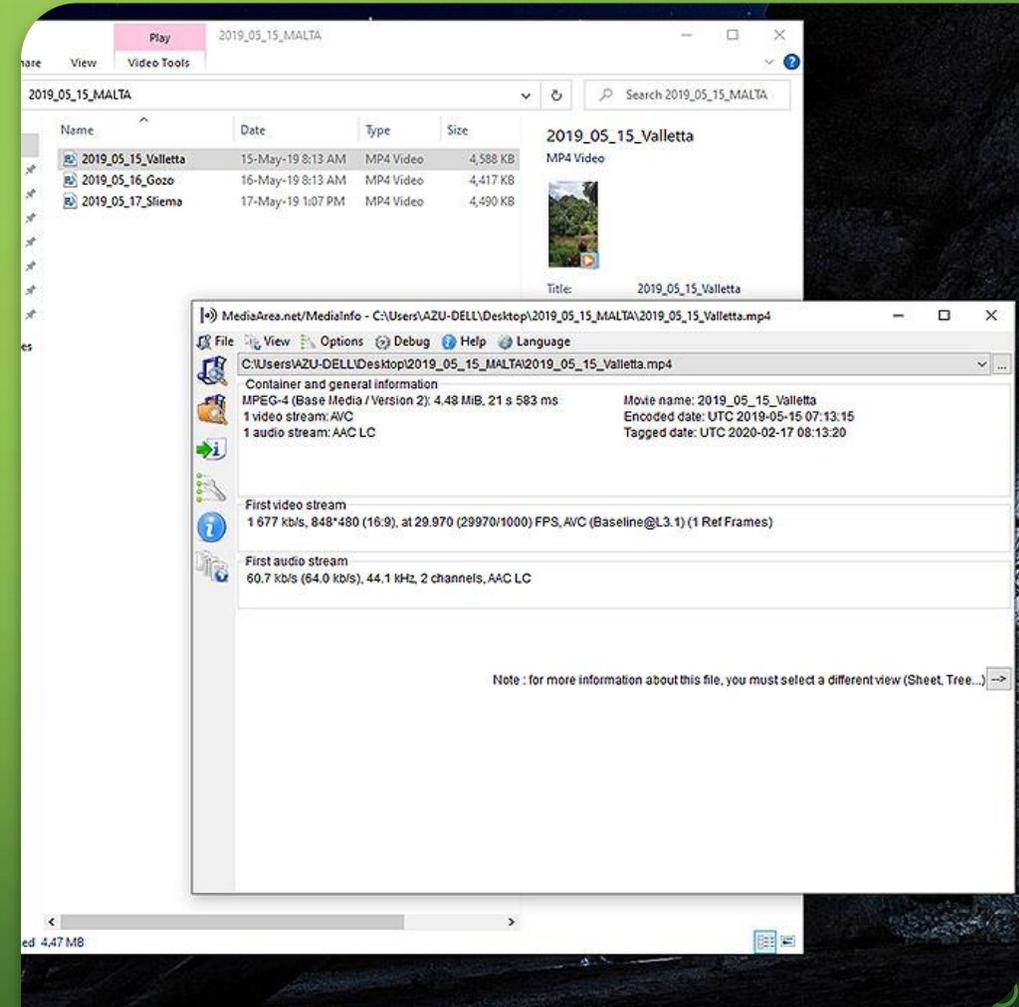
## SESSION 3

# GET THE METADATA

- **Metadata** is, essentially, 'data about data'. It is the information stored to describe aspects of a file, and it's content.
  - E.g. technical specifications, resolutions, bit-rates.

## USEFUL TOOLS!

- ❖ Try **Tubekit** ([tubekit.org/](http://tubekit.org/)): It's a YouTube crawler that allows you to extract YouTube video data (e.g. author, keywords, genre, no. of views, ratings, comments, etc).
- ❖ Try **MediaInfo** ([mediaarea.net/en/MediaInfo](http://mediaarea.net/en/MediaInfo)): displays metadata embedded in video & audio files.
- ❖ Try **VLC player** ([videolan.org/](http://videolan.org/)): Free & open source multimedia player.



MediaInfo

# QUICK POLL

WHICH OF THE FOLLOWING VIDEO FORMATS DO YOU WORK WITH  
MOSTLY ?

## SESSION 4

# STRATEGY: (TAKEN FROM LIBRARY OF CONGRESS)

- **3. ORGANISE** the content.
  - Gather everything into one place.
  - Create one main archive folder with an understandable name.
  - Then create sub-folders inside the master folder and name them simple descriptive titles. It may help to include name and date of subject.
- **RENAMING FILES:** follow file-naming conventions.

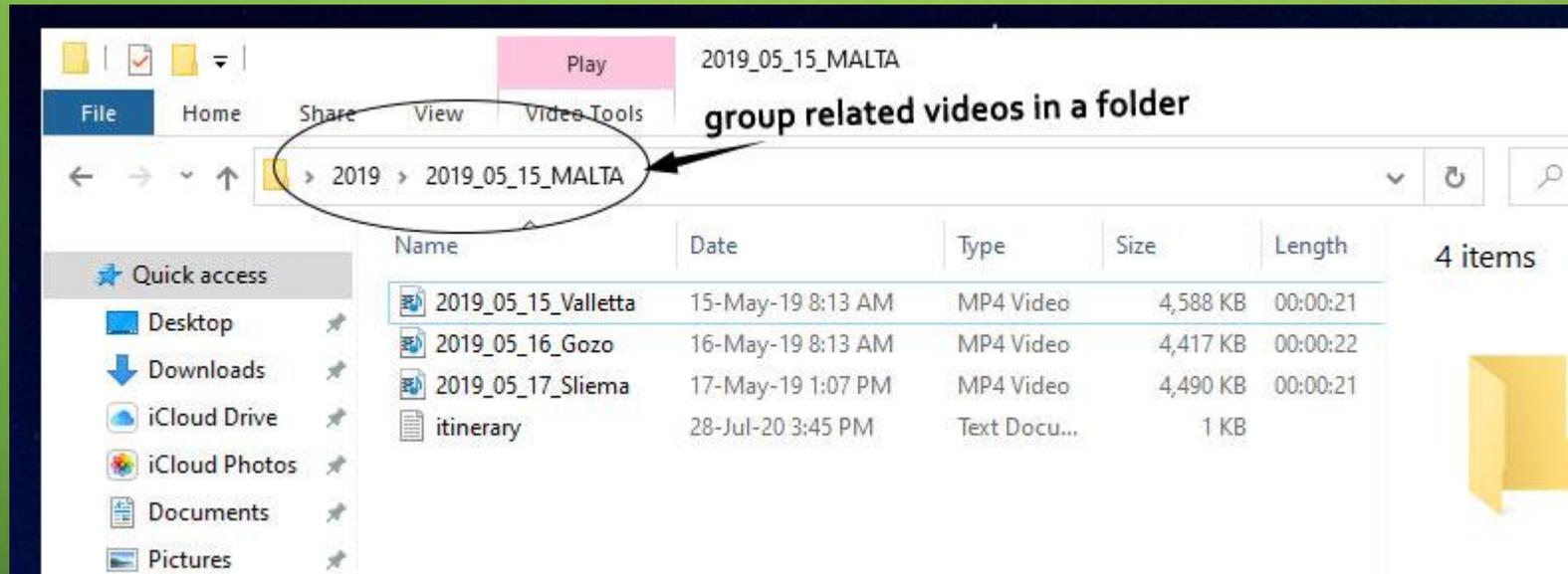
azu001\_2019\_05\_15\_valletta.mp4



**N.B:** Avoid using “illegal” characters @#\$%&\*<>?/\~

## SESSION 4

- Create an “INFORMATION PACKAGE”
  - Basically, group related videos or even text files in a folder.
  - This helps maintain the original order & context.



- Also prevents videos from becoming disassociated with each other.
- REMEMBER: There is no one correct way to structure the organisation of YOUR video files.
- Use a structure that works for YOU !

## SESSION 4

# STRATEGY

## 4. STORAGE

- Storing isn't just about finding a device or service to hold your videos.
  - It requires a SET OF ACTIONS to ensure your media stays intact, secure, and accessible.
- 
- SIMPLE STRATEGIES for keeping your videos safe:
    1. Making copies.
    2. Checking files.
    3. Controlling access.
    4. Refreshing your devices.

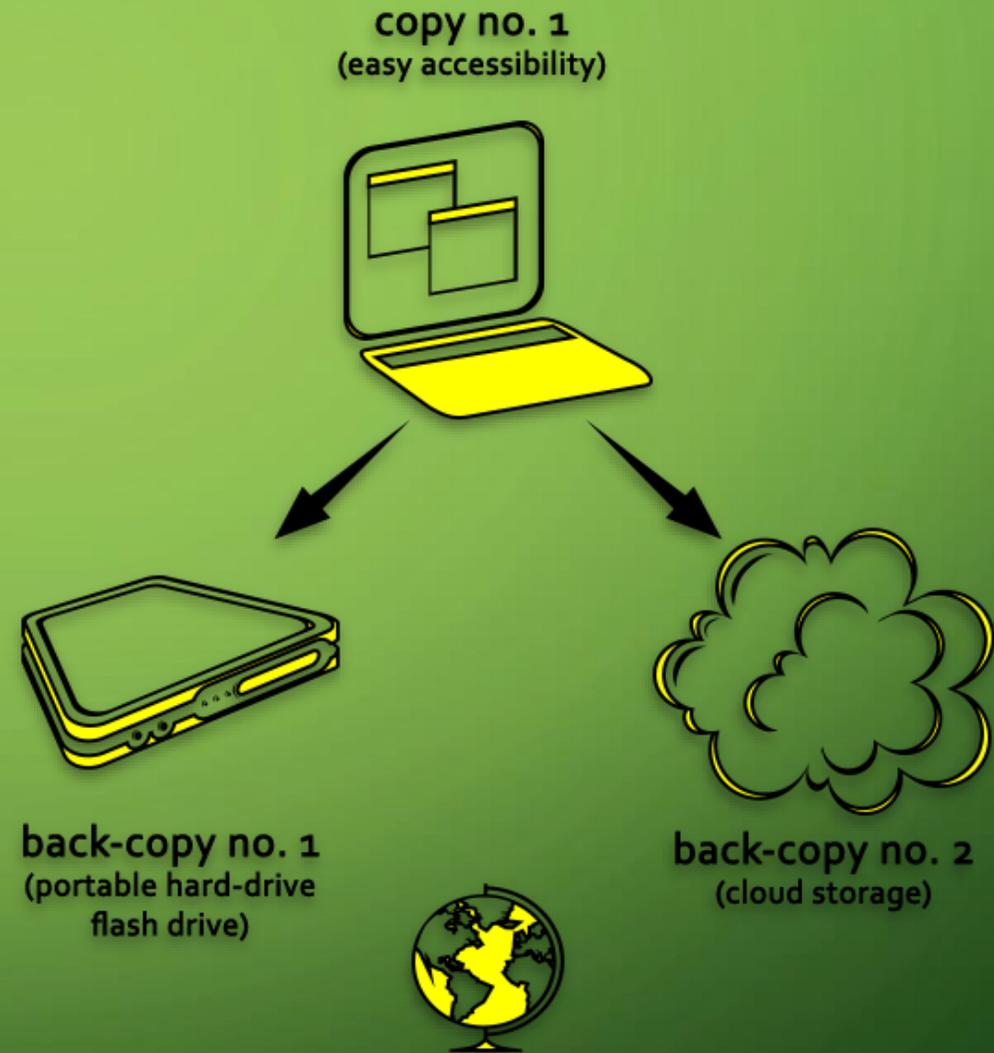
## SESSION 4

### 1. MAKING COPIES:

- LOCKSS (Lots of Copies Keep Stuff Safe).
- 3-2-1 back-up rule used by pros!
  - It simply states that
    - ✓ 3 copies of your data.
    - ✓ 2 different media.
    - ✓ 1 copy off-site.

### USEFUL BACKUP TOOLS!

- Time Machine (Mac)
- Backup & Restore (Win-PC)
- File History (Win-PC)
- Teracopy (<http://codesector.com/teracopy>)
- Bacula (<https://www.bacula.org/>)



with 1 copy off-site for disaster recovery.

## SESSION 4

### 2. CHECKING FILES

- ensure File Fixity – perform FIXITY checks.
- means computing & comparing a file's HASH VALUE (aka checksum) with a previously computed hash value.
- If the file is altered in any way, it's hash value will be different.

#### USEFUL TOOL

- <https://www.weareavp.com/products/fixity>
- **DROID**

### 3. CONTROL ACCESS

- Try to control who has physical & electronic access to your storage devices.
- Only give access to people who needs access.
- Store video files on a volume separate from your other files.
  - Robert McMilen  
<https://youtu.be/Gv3Jkib51i0>

### 4. REFRESHING YOUR STORAGE MEDIA

- Many different types of storage media (use them in different combination).
- But, how do you decide?
  - I. What is the level of support you require?
  - II. What is the size of your video collection?
  - III. Is it easy to refresh your video collection?

## SESSION 4

# TYPICAL HARDWARE / STORAGE MEDIA

### Portable hard drive

Ideal for:

- Collections no larger than 2-3 TB.
- Collections that only need to be accessed by one computer/user at a time.

#### PROS

- Relatively low cost (usually £70 - £120).
- Portable.

#### CONS

- Drives (especially Firewire) fail often.
- Platform-dependent.

### Cloud storage

Ideal for:

- Small collections.
- Collections that need to be accessed by people in different locations.

#### PROS

- Collections can be shared worldwide.
- Storage is maintained by 3<sup>rd</sup> party. Most have very good infrastructure. (e.g. amazon aws, vimeo, dropbox, google drive)

#### CONS

- Plenty of bandwidth is needed to upload & download files.
- On-going subscription fees.
- Service can terminate at any time, sometimes without notice!

## SESSION 4

# WHICH FORMAT IS BEST ?

- Unfortunately, a single format does not exist that will meet every need related to digital video.
- Instead, let's ask ourselves a few questions that can be used as a guideline (taken from Digital Preservation Office at Uni. Of Michigan):
  1. Is this format currently in common use and well supported by playback software?
  2. Is this format a proprietary format? Does it have open documentation?
  3. What are the current preservation standards?
    - Library of Congress's Sustainability of Digital Formats.  
[https://www.loc.gov/preservation/digital/formats/fdd/video\\_fdd.shtml](https://www.loc.gov/preservation/digital/formats/fdd/video_fdd.shtml)



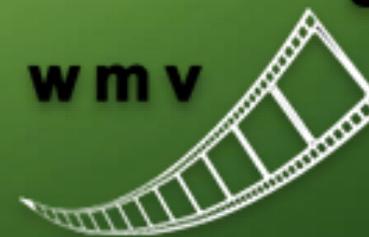
## SESSION 4

# COMMON FILE FORMATS

- mpeg-4 (.mp4) and mpeg-2
- avi (Microsoft)
- mov (Quicktime)
- wmv (Microsoft)
- mkv (Matroska) *future?*

At this time, **MPEG-2** and **MPEG-4** stand out as the best formats for long-term preservation.

- They are open format.
- Both used for commercial broadcast.
- Both used for government entities (Library of Congress).



# “PERSONAL DIGITAL PRESERVATION PLAN” TEMPLATE

SWITCH TO TEMPLATE

## IN SUMMARY:

### Appraise

1. **Identify** what video types you want to save.
2. Decide or **select** on what is most important to you to save.

### Preserve & Access

3. **Organise** your content.
4. **Save** copies in different places.

# USEFUL RESOURCES

- The National Archives (TNA)

<https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/guidance/>

- Digital Preservation Coalition (DPC)

<https://www.dpconline.org/handbook>

<http://digitalpreservation.gov/personalarchiving/video.html>

- Library of Congress (LOC)

<https://www.loc.gov/preservation/digital/>

- WITNESS

<https://archiving.witness.org/archive-guide/>

- Borthwick Institute

<http://borthwickinstitute.blogspot.com/2017/11/save-your-digital-stuff.html>

# USEFUL ONLINE COURSES

- Novice to Know-How

<https://www.dpconline.org/digipres/train-your-staff/n2kh-online-training>

- An introduction to identification of archive film & video

<https://www.eventbrite.co.uk/e/an-introduction-to-identification-of-archive-film-video-online-training-tickets-115180548120>