

# **'DIY' Digital Preservation for Audio**

Identify · Appraise · Organise · Preserve · Access

# Plan + housekeeping

## **Part 1** (10:00am-10:50am)

10:00am-10:10am **Intros and housekeeping**

10:10am-10:30am **Digital Preservation for audio**

*What material do you care about and hope to keep?*

10:30am-10:50am **Locate, Appraise, Identify**

*What kind of files are you working with?*

## **Part 2** (11:00am-11:45am)

10:50am-11:00am **[10-minute comfort break]**

11:00am-11:15am **Organise (Migrate)**

11:15am-11:45am **Store, Maintain, Access**

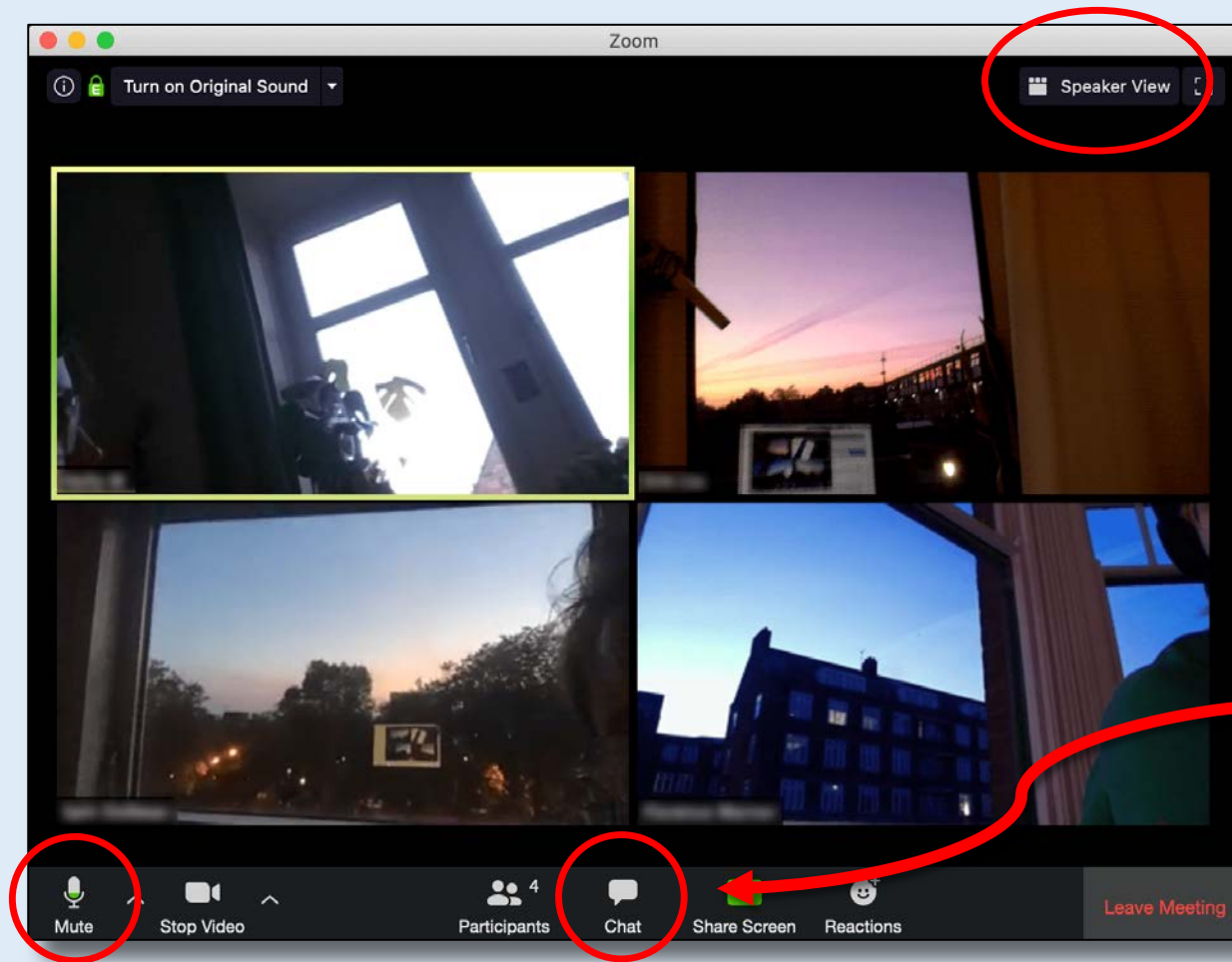
*Discussion & Questions*

# Using Zoom

Feel free to swap to 'Grid View' when slides are not in use, to gage 'the room'.

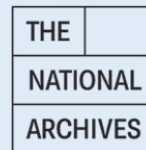
Please use the chat function to ask questions while slides are in use.

Please keep your mic on 'mute' when not speaking.



# Who are we?

- **Bridging the Digital Gap**  
traineeship scheme
- UK National Archives  
(National Lottery Heritage Fund)
- bringing 'digital' skills into the  
archives sector



ual:

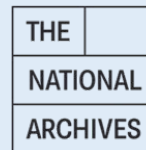


UNIVERSITY OF  
WESTMINSTER



# Why are we doing these workshops?

- agitate the cultural record to reflect lived experience
- embrace tools that support historical self-determination among non-specialists
- raise awareness, share skills, share knowledge



ual:

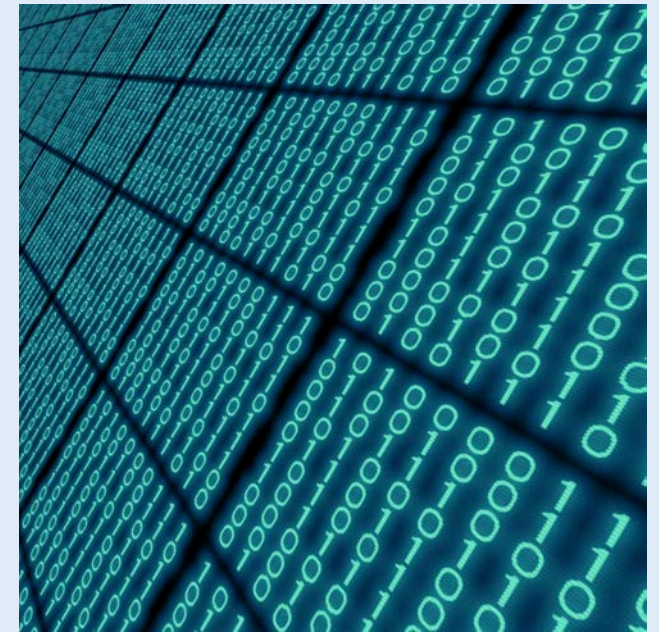


UNIVERSITY OF  
WESTMINSTER



# What is digital preservation?

- digital material is vulnerable in different ways to analog material
- **digital preservation =**  
“a series of managed activities undertaken to ensure continued access to digital materials for as long as necessary.”





# Audio practices and technological dependency



Image credit: Don Shall



Image credit: Tarje Sælen Lavik



Image credit: JuneAugust



Image credit: Mk2010

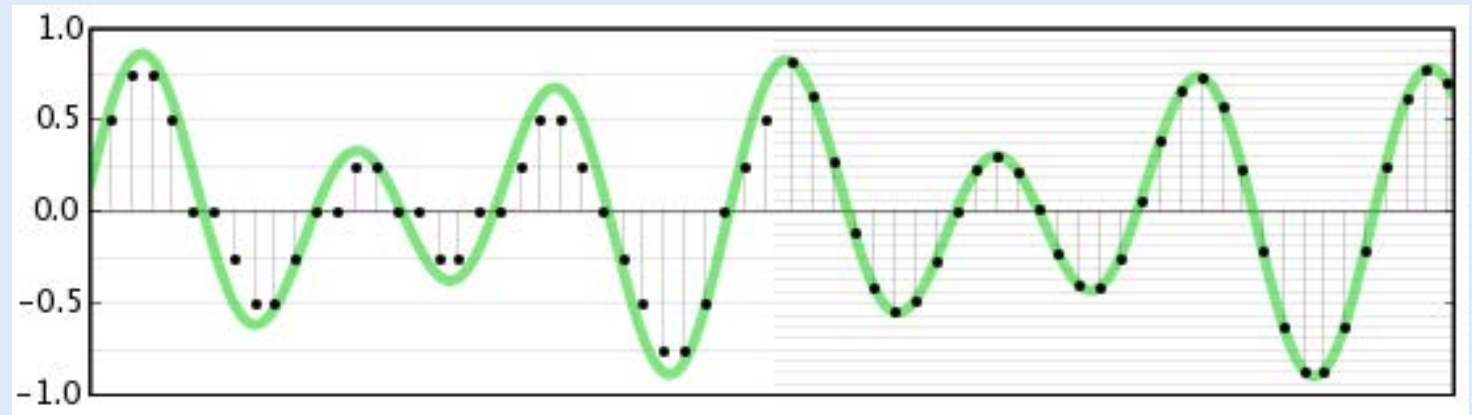


Image credit: Stahlkocher

# Digital preservation and 'born-digital' audio



- codecs  
(e.g. LPCM)
- formats and  
containers  
(e.g. WAV, Ogg)



*Image credit: audacityteam.org*



# Challenges for personal digital preservation

## Technical

- hardware failure
- bit rot
- file corruption
- virus/malware
- media obsolescence (file format, hardware, software)

## Non-technical

- unclear ownership/responsibility
- lack of documentation or long-term planning (neglect)
- overdependence on third-party solutions (e.g. Google Drive)
- loss/theft, natural disaster



**Peter Van Garderen** @pjvangarderen · Jul 28  
Not your spreadsheet, not your data. [#digipres](#)



Dear Google User,

We are writing to inform you of a technical issue affecting Google Sheets between December 11, 2013 and November 18, 2019. During this time period, in very rare circumstances, a cell's data ostensibly appeared empty in the current version of a sheet.



Dear Google User,

We are writing to inform you of a technical issue affecting Google Sheets between December 11, 2013 and November 18, 2019. During this time period, in very rare circumstances, a cell's data ostensibly appeared empty in the current version of a sheet.

The root cause that led to this issue has been identified and resolved. However, in some instances, we were, due to technical limitations, unable to restore the impacted data completely, and the cells may continue to appear empty.

In case a cell's data appears empty from your Google Sheets, Google recommends that to locate the data by viewing the [cell's edit history](#) and replacing the cell's contents with the data present in the historical version.

Attached is information on your potentially impacted documents. The .csv attachment includes the following information:

- Spreadsheet URL
- Cell Number
- Issue happened after
- Issue happened before

We appreciate your business and sincerely apologize for any inconvenience this may have caused.

# **Personal digital preservation is necessary and possible**

## **From an amateur filmmaker:**

lost iPhone and all film/audio recordings within - didn't back-up project material - only low-quality test export shared over Facebook Messenger retrievable

## **From a CSM Graphics student:**

programmed a piece of software for thesis project - file corrupted days before degree show - had to take time off work to rapidly recreate

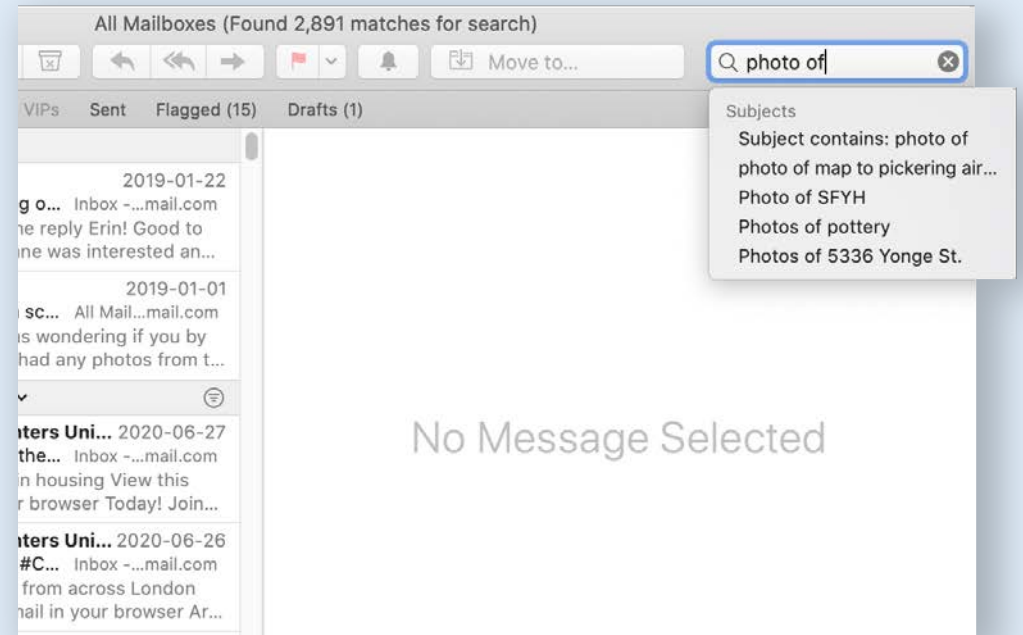
## **From my dad:**

taped over home video footage with daily news

**What material do you care about and hope to keep?**

# (1) Locate your material

- **hardware:** floppy disks, CDs, USB/Flash drives, tascam, camera, mobile device
- **shared drives:** Google, DropBox, your institution/workplace
- **other places to look:** email attachments, chat history, social media



Determining when and in which contexts specific files were used is easier for more recent projects, but can become increasingly difficult the longer you leave them.

## (2) Appraise your digital material (content)

- use **VLC** to playback most AV formats/containers
- consider how you want to access your files
  - high quality exports only?
  - project files + linked files?
  - process documentation?



Screenshot of VLC being used to inspect a movie file.



## (2) Appraise your digital material (technical)

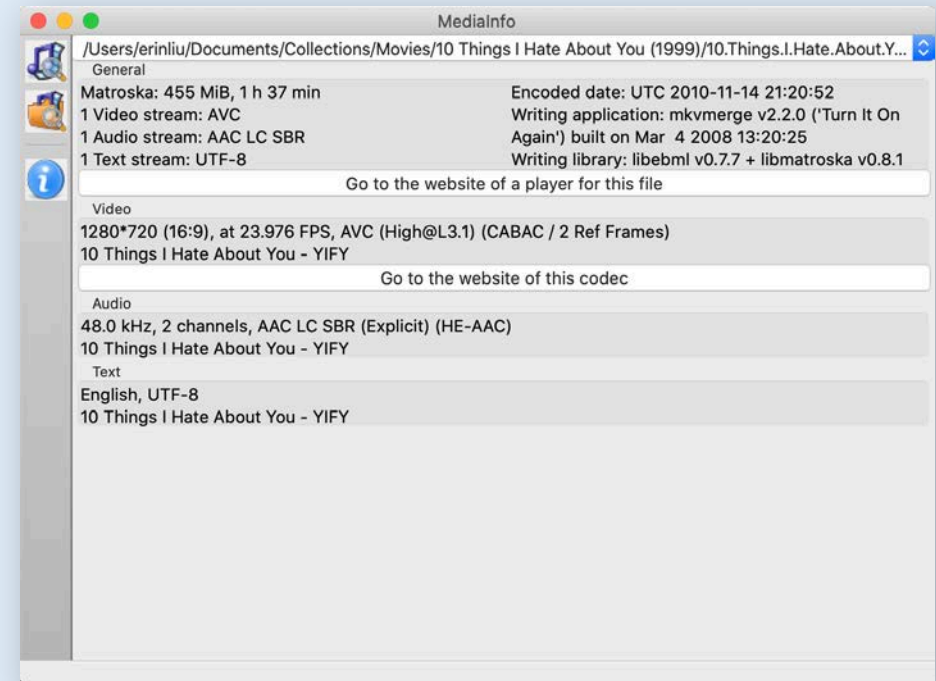
- consider deleting low-res duplicates
- fine to delete Database (.db), Peak (.pkf) files
- **TreeSize Free** (Windows) or **GrandPerspective** (MacOS) can help visualise your files according to size and maintain any existing folders



Screenshot of a GrandPerspective scan for a 'Documents' folder.

### (3) Identify what you have

- file extensions identification alone may not always be accurate
- knowing what formats you have will support in determining how to care for them
- **DROID, MediaInfo**
  - MediaInfo designed specifically for AV material



**Link to download MediaInfo:**  
[mediaarea.net/en/MediaInfo](http://mediaarea.net/en/MediaInfo)

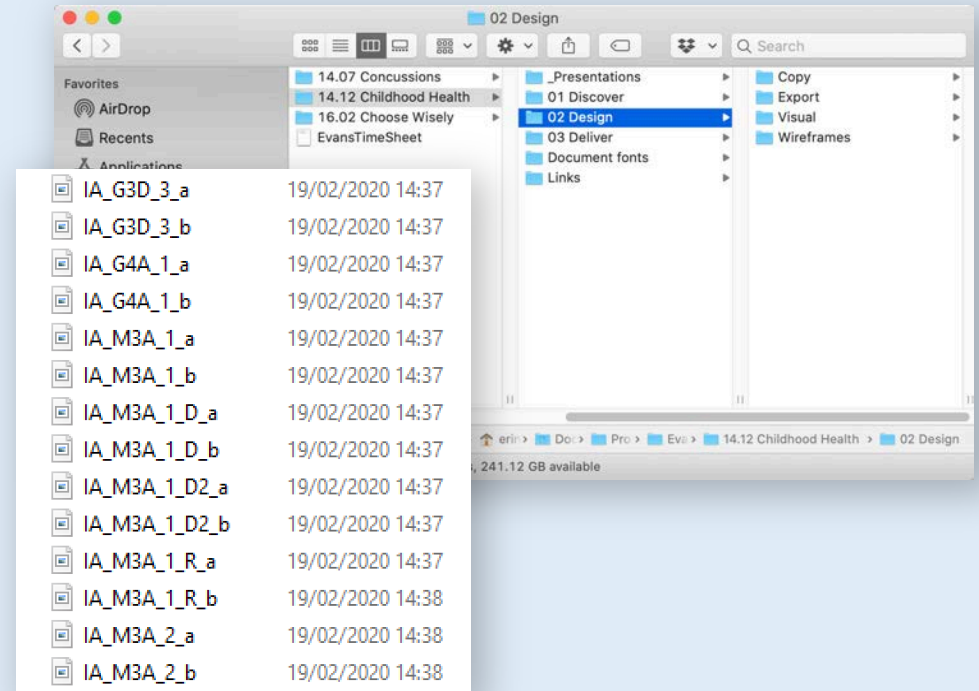


What material do you care about and hope to keep?

**What kind of files are you working with?**

## (4) Organise your files

- use concise, consistent, meaningful filenames and folders
- avoid special characters
- e.g. "Img\_3081" vs.  
"2018\_1087\_1187" vs.  
"20181118\_COL\_nisha-voiceover\_01"
- many file-renaming tools on the web,  
e.g. **Bulk Rename Utility** (Windows)



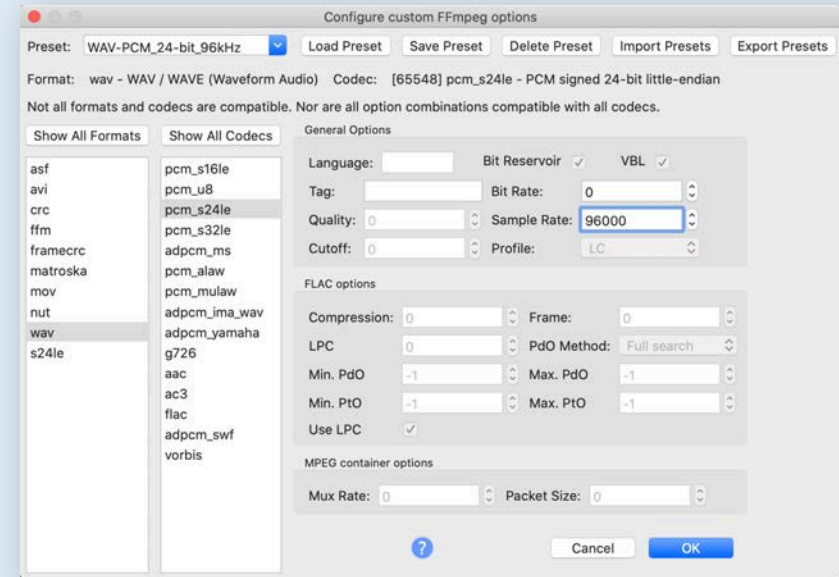
You may already have a file-naming and folder-organising convention. The key is to think about how you'd like to access your files long-term, bearing in mind what relationships need to be maintained among your files to enable access/use.

## (5) Migrate any files formats at-risk of obsolescence

- archival format for digital audio files:

### WAV - PCM; 'native' resolution

- bear in mind lossless versus lossy encoding for long-term preservation
- use **Audacity with FFmpeg** for customisable file conversion, keep originals in case of issues with conversion



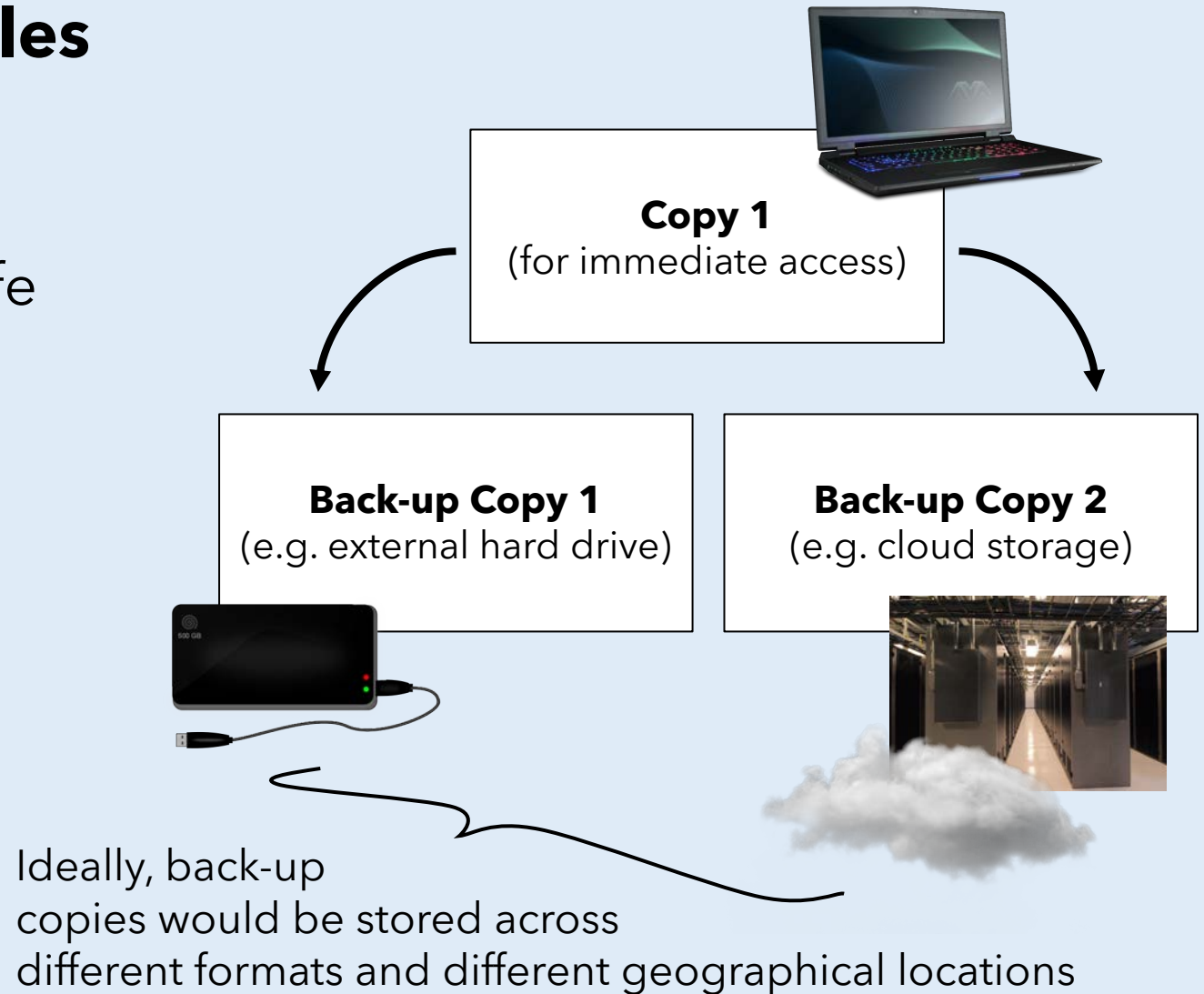
**Link to download Audacity:**  
[audacityteam.org/download](http://audacityteam.org/download)

**Using FFmpeg in Audacity:**  
Audacity>Preferences>Libraries



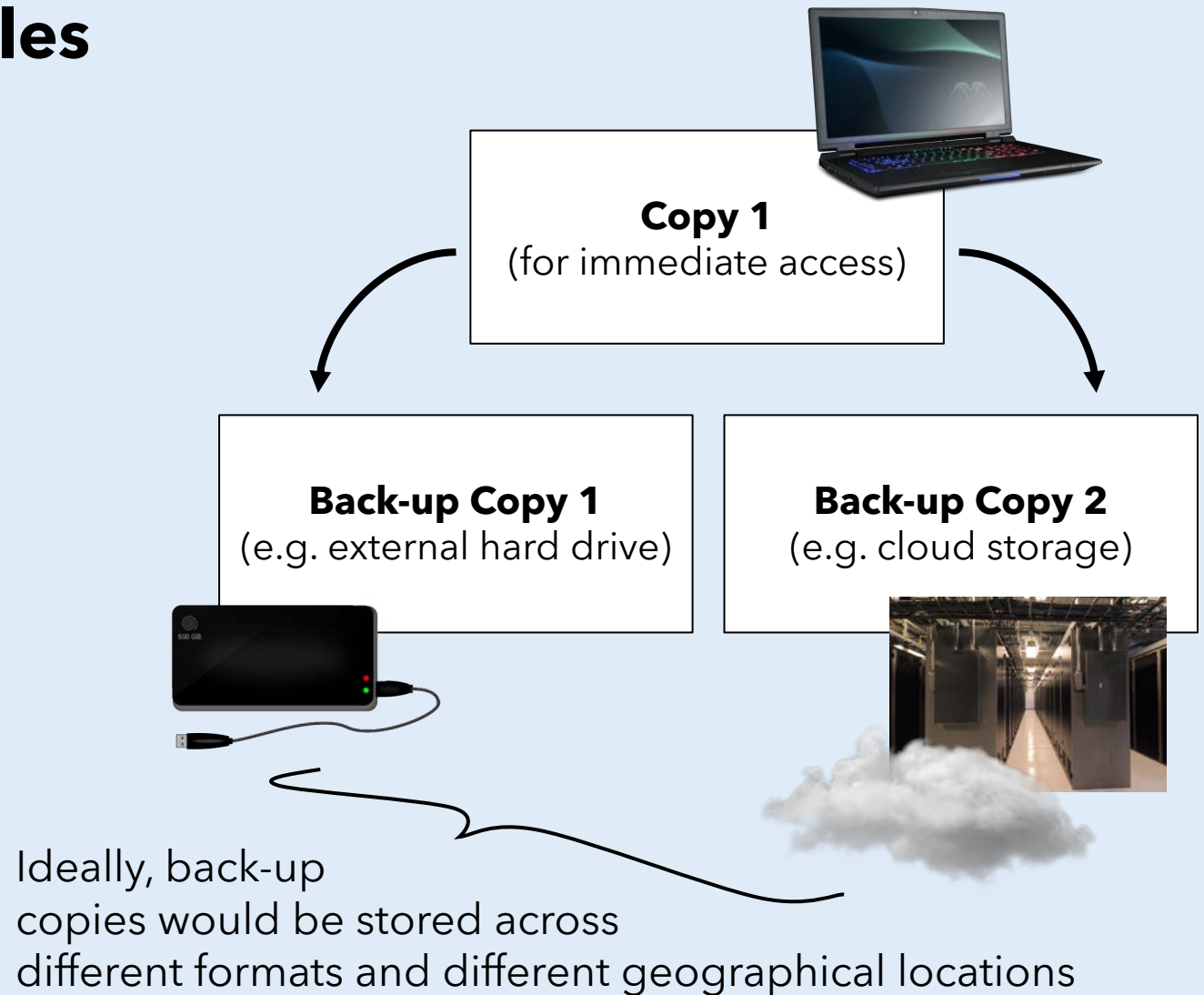
## (6) Storing your digital files

- LOCKSS =  
Lots of Copies Keep Stuff Safe
- 3-2-1 back-up rule used by professionals
- save copies (ideally, three) in multiple locations, ideally in different formats



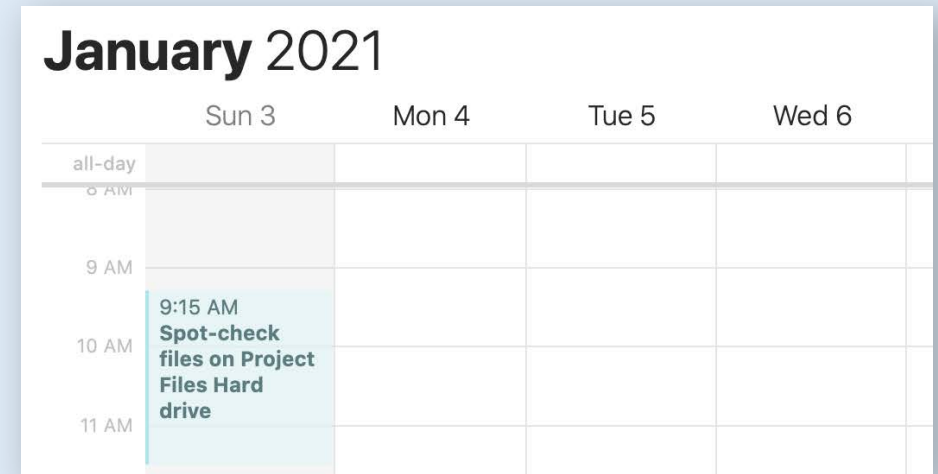
## (6) Storing your digital files

- external hard drives
- cloud storage (be aware of any legal or privacy obligations)
- CDs are not an archival storage medium



## (7) Maintaining your files over time

- refresh storage media, ideally every five years
- **test access to files**, spot-check once a year
- **bear in mind digital preservation when creating files, particularly when using subscription software or emerging software/formats**



A calendar for January 2021. The header shows the month and year. The days of the week are listed at the top: Sun 3, Mon 4, Tue 5, Wed 6. The time slots are listed on the left: all-day, 8 AM, 9 AM, 10 AM, 11 AM. A task is scheduled for 9:15 AM on Sunday, 3rd: "Spot-check files on Project Files Hard drive".

|         | Sun 3 | Mon 4 | Tue 5 | Wed 6 |
|---------|-------|-------|-------|-------|
| all-day |       |       |       |       |
| 8 AM    |       |       |       |       |
| 9 AM    |       |       |       |       |
| 10 AM   |       |       |       |       |
| 11 AM   |       |       |       |       |

Taking care of your digital material can be low-demand when scheduled at regular intervals, like other forms of housekeeping.

# Further Resources:

## Digital Preservation Coalition

- Technology Watch Report on Personal Digital Archiving:  
[www.dpconline.org/news/personaldigitalarchiving-twr](http://www.dpconline.org/news/personaldigitalarchiving-twr)
- BitList:  
[www.dpconline.org/our-work/bit-list](http://www.dpconline.org/our-work/bit-list)

## Library of Congress

- Recommended File Formats Statement:  
[www.loc.gov/preservation/resources/rfs/audio.html](http://www.loc.gov/preservation/resources/rfs/audio.html)
- Sustainability of Digital Formats, descriptions for sound:  
[www.loc.gov/preservation/digital/formats/fdd/sound\\_fdd.shtml](http://www.loc.gov/preservation/digital/formats/fdd/sound_fdd.shtml)

What material do you care about and hope to keep?

What kind of files are you working with?

**Questions / Discussion**