

DIGITIZING MUNICIPAL RECORDS

Elizabeth O'Connell, Digital Archivist
Yolande Bennett, Asst. Digital Archivist

WHAT IS A DIGITAL REPOSITORY?

DIGITAL REPOSITORY

The technical infrastructure, services, and resources for the storage and management of digital information

DIGITIZED VS. BORN DIGITAL |

DIGITIZED VS. BORN DIGITAL

Digitized: Analog information that has been transformed into digital form

□ Example: A physical copy of meeting minutes that has been scanned

Born Digital: An electronic record originating in a computer environment

□ Example: An archived e-mail or social media post

RECORD VS. ACCESS COPIES

RECORD VS. ACCESS COPIES

Record: Data or information in a fixed form that is created or received in the course of individual or institutional activity and set aside (preserved) as evidence of that activity for future reference

Record Copy: The single copy of a document, often the original, which is designated as the official copy for reference and preservation

Access Copy: A reproduction of a document created for ease of access and use by additional users

ACTIVE VS. INACTIVE RECORDS



ACTIVE VS. INACTIVE RECORD

Active Record: A record that is used with sufficient frequency to justify keeping it in the office of creation

Inactive Record: A record that is no longer used in the regular course of business but which may be preserved and occasionally used for legal, historical, or operational purpose

INTELLECTUAL CONTROL



WHAT IS INTELLECTUAL CONTROL?

Intellectual Control over your records means knowing not only what you have, but also how much of it there is, where it's located, and how long you'll need to maintain it for.

Having this control protects you from losing documents, allows you to follow the retention schedule better, and can help you make decisions regarding storage, digitization, and destruction.

THE FIRST STEP

Achieving intellectual control of your documents is the necessary first step before creating a digital repository.

To gain intellectual control:

- Survey
- Organize
- Record
- Capture metadata

TO DIGITIZE OR NOT TO
DIGITIZE |

PROS AND CONS TO DIGITIZATION

PROS

- Increasing ease of access
- Saving physical space
- Having a back-up incase of accidental destruction of Record Copy
- Adding accessibility features such as OCR for screen readers
- Time: Saves time searching for records and when reviewing for destruction

CONS

- Costs: both of digitization and digital storage
- Time: prepping for digitization, installing the infrastructure to house it, and processing the electronic files
- Risk: Digital files have an increased risk of degradation or loss when not properly maintained or stored

WHAT YOU NEED TO KNOW

1. Third party storage vs. Self-storage
2. Never keep a single copy
3. File upkeep
4. The more prep you do prior to digitization, the better
5. Your responsibilities don't change

END GOALS

1. Increase access (Public)
2. Increase access (Your organization)
3. Save space
4. Preserve at-risk records
5. Efficiency

WHAT TO SCAN AND DESTROY

Good Options

High volume collections of printed documents

Records with medium retention periods
(4 – 25 years)

Seldom accessed records

Bad Options

Records with permanent retention

Records with short retention periods
(<5 years)

Records with very long retention periods
(26+ years)

Records of historical importance

Handwritten documents

Records commonly requested by the public

DIGITIZING |

CHOOSING WHAT TO DIGITIZE

Things to consider:

1. Condition
Tears, fasteners (staples, paperclips, etc.), water damage, fading, etc.
2. Uniformity
3. Use
4. Volume
5. Legibility
6. Fragility

PREPARING RECORDS FOR DIGITIZATION

1. Survey documents
2. Determine type of copy to be created
Record Copy, Archival Copy, Access Copy
3. Organize the records
Alphabetical, numerical, by date, etc.
4. Create metadata
5. Prep physical copies
Remove fasteners, unfold, stabilize tears, remove sticky notes
6. Determine file-naming and file folder structure
7. Determine scanner settings

DIGITIZATION GUIDELINES

1. Survey documents designated for scanning
Examine physical characteristics
2. Determine the type of digital record to be created
Record Copy, Archival/Preservation/Master Copy, or Access Copy
3. Establish suitable image capture mode
Black-and-White, Grayscale, or Color
4. Determine resolution settings

RESOLUTION SETTINGS

Archival Record	Spatial Resolution	Bit-Depth	Optimal Format	Acceptable Format
B & W printed text documents	300-400 PPI	B & W	PDF/A	PDF
Handwritten text documents	300-400 PPI	Grayscale	PDF/A	PDF
Damaged text documents	300-600 PPI	Grayscale	PDF/A	PDF
B & W photographs	300-600 PPI	Grayscale	TIFF	JPEG
Color photographs	300-600 PPI	Color	TIFF	JPEG
Oversized materials	300-600 PPI	Grayscale or Color	TIFF	JPEG

DIGITIZATION GUIDELINES

1. Survey Documents Designated for Scanning
Examine physical characteristics
2. Determine the Type of Digital Record to be Created
Record Copy, Archival/Preservation/Master Copy, or Access Copy
3. Establish Suitable Image Capture Mode
Black-and-White, Grayscale, or Color
4. Determine Resolution Settings
5. Choose a Final Format
PDF vs. PDF/A, TIFF vs. JPEG
6. Quality Check
A minimum of 10 images or 10 percent should be inspected, whichever is higher

AFTER DIGITIZATION

- Do a quality control check
- Make sure file names adhere to best practices
- Organize digital records
- Upload to Digital Repository and back-up locations
- Schedule maintenance checks for records
- Schedule destruction for records with finite retention schedules
- Optional: Run OCR on documents to make them searchable/usable by screen readers

FILE ORGANIZATION BEST PRACTICES

FILE NAMING

- Create a consistent and descriptive naming practice and document it through policy
- Avoid the characters: ~! @ # \$ % ^ & * () ' ; , < > ? “
 - *Replace spaces with underscores, camel case, or hyphens*
- Avoid using periods in file names
- Keep file names between 25-30 characters
- Use leading zeros to sort files in sequential order
 - *Example: 001, 002 ... 010, 011, etc.*

FILE NAMING

- Format dates by year, month, and day for chronological sorting

 - Use YYYYMMDD, or YYYY-MM-DD, or YYYY_MM_DD

- Document version control via file names

 - Example: *CollectionPolicy_firstdraft_2022_06_03.pdf*, *CollectionPolicy_finaldraft_2022_06_07.pdf*

- Examples of well-formed file names:

 - *2009_01_15_planning_board_meeting_minutes.docx*

 - *2009_02_10_planning_board_meeting_minutes.docx*

 - *2009_CommitteeEventImage_001.jpeg*

 - *2009_CommitteeEventImage_002.jpeg*

FILE ORGANIZATION

Folder structures can help organize and provide contextual information for files

□ Example: Keep track of a file's retention period by including this information in the folder name

Determine what critical information should be kept in the file name, such as date information, and what other information may be helpful to record in the folder or directory name

When implementing policy and procedure, consider file organization strategies together with file naming preferences and review these policies on a regular basis.

MAINTAINING DIGITAL RECORDS



PROPER STORAGE

The 3-2-1 Rule

3 – Make 3 copies

Your central computer can count as one copy, but you should make two additional copies in case more than one backup fails.

2 – Use 2 different types of storage media

Rather than rely on the endurance of one type of storage media, keep at least one of your copies in a separate storage format (I.E. Cloud, hard drive, server).

1 – Store 1 back up in a different location

In the event of a physical disaster, such as a fire, keep one of your backups in a different geographic location.

PROPER STORAGE

All digital data is encoded onto physical media which in itself is susceptible to damage and degradation. For best results, digital storage media should be kept in the same kind of environmental conditions that you find comfortable:

1. Avoid extreme temperatures or changes in humidity
Maintain conditions around 70 degrees Fahrenheit and 40% humidity
2. Avoid direct light, moisture, dust, dirt, mold, and pests
3. Keep digital storage media away from powerful magnets

CHECKSUMS

A checksum is a small-sized block of data derived from another block of digital data for the purpose of detecting errors that may have been introduced during its transmission or storage.

Checksums should be run right after a digital file is created, after a digital file is transferred to a new location, and on a regular schedule to ensure file integrity.

Make sure to keep a record of each checksum string.

If the checksums don't match, that means that the file has degraded in some way, and should be replaced in that location with a file from another location.