



Town of Marblehead

Document Digitization Best Practice

Prepared By: The Office of Municipal & School Technology

EOTSS | Executive Office of Technology Services & Security

Table of Contents

Executive Summary	2
Community Profile	
Efforts to Date	3-5
Best Practice Guidance.....	6-11
Scanning Device.....	
File Format	
Image Resolution.....	
Metadata	
Quality Assurance and Control	
Retention Schedules and Disposal Rules	
Recommendations	11-13
Digitization Protocol	
Acquiring Other Scanning Devices	
Integration of Records Management System	
Maintenance and Training.....	
Potential Opportunities.....	
References	14-17
Appendices.....	18-23



Image: “Full Moon Over Abbot Hall”¹

Executive Summary

In June of 2017, the Town of Marblehead signed a Community Compact agreement with the Baker-Polito administration. The Town adopted the Document Digitization Best Practice to address the lack of physical space for municipal document storage. This report will provide best practice guidance regarding document digitization and retention of municipal documents for the Town of Marblehead.

The Town of Marblehead requested technical and staffing guidance to create operational efficiencies and make municipal documents readily available for staff and the public. The town also identified the following objectives in selecting this best practice:

- Reduce the amount of physical space needed to store municipal documents

¹ Kipp, Rob. “Full Moon Over Abbot Hall”. Town of Marblehead. Accessed February 14, 2019.
<https://www.marblehead.org/about-marblehead/slideshows/photo-gallery>

- Digitize large format documents (past engineering plans and maps)
- Streamline digitization processes and develop guidelines for the future

Community Profile

The Town of Marblehead is located in Essex County, Massachusetts. The Town was first settled in 1629, and incorporated in 1649, becoming independent of Salem.² Being integral to the American Revolution and the Civil War, the Town is known for historic buildings, streets, and beaches. Today the Town has a population of 19,808 and a median household income of \$110,025.³

Efforts to Date

One of the major challenges the Town ran into was the condition of the large format documents to be digitized. As is typical, many of these documents were rolled and stored in an area without any special climate control considerations, resulting in brittle documents that are difficult to digitize while retaining their integrity. To prevent further deterioration, the Town researched the challenge and developed an in-house solution to humidify documents, allowing for scanning and long-term digital preservation as well as flat storage of the original documents.

The document humidifying solution, pictured below, consists of a humidifier connected to a tub with a tray to hold the documents. After being exposed to humidity, documents are ironed with press cloth on top, preventing potential heat damage to the documents. Lastly, documents are laid flat with weight on top in preparation for scanning and storage. Over the past two years, this process has made it possible for the Town to digitize 15,000 rolled large format documents, and more effectively store the original documents.

² Marblehead Carved Out of Salem. MassMoments. Accessed February 14, 2019.
<https://www.massmoments.org/moment-details/marblehead-carved-out-of-salem.html>

³ “Community Facts”. United States Census Bureau. American Fact Finder.
https://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml

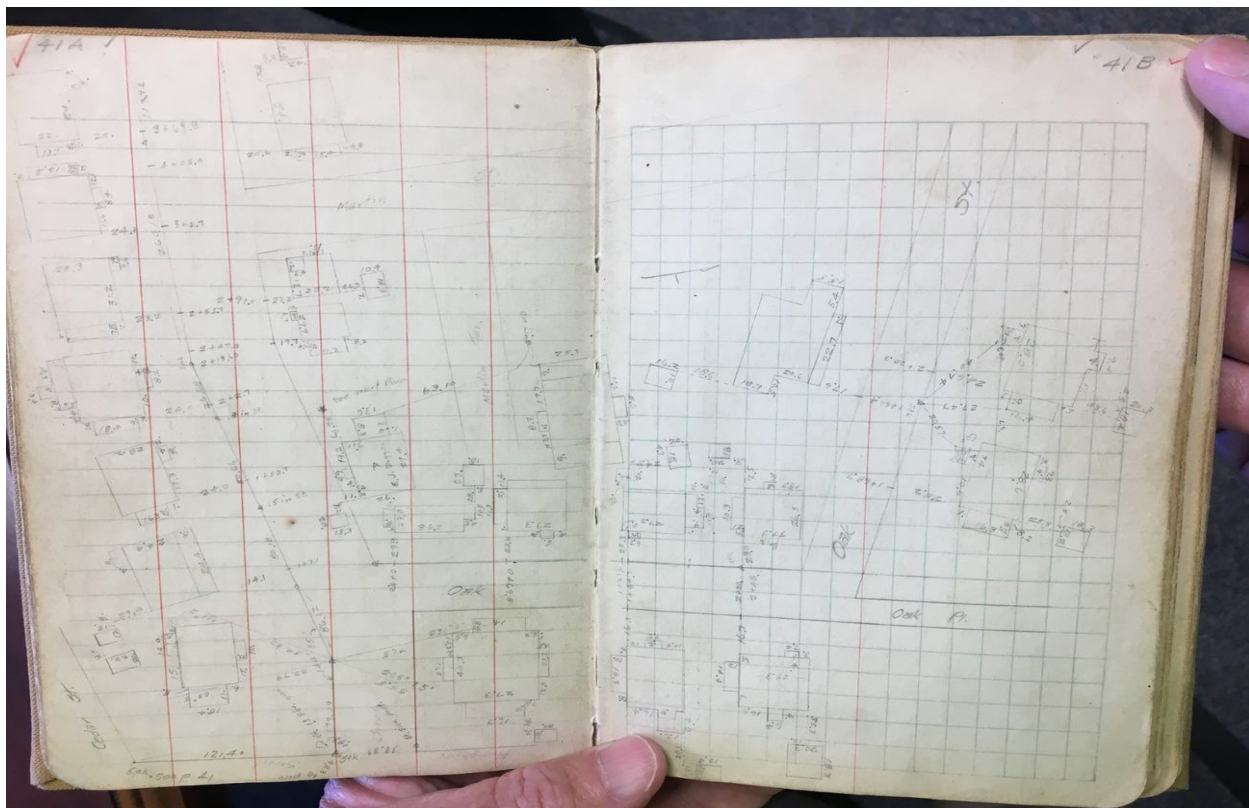
In addition to the challenge of scanning rolled documents, the Town is interested in digitizing bound books. The planning book pictured below is an example of a fragile, historical document, with intricate detail that the Town would like to preserve in a digital format.



Marblehead's custom document humidifying solution



Marblehead's document ironing station



Hand-drawn Marblehead Planning on a bound book

Best Practice Guidance

While the Town of Marblehead has shown great initiative and has made incredible progress in digitizing historic documents, they recognized a need to implement standards as the operation grows. The Federal Agencies Guidelines Initiative developed a Digitization Activities document that may be useful in understanding the process and phases involved in digitization initiatives.⁴

The initiative defines the components of the digitization process as:

- Selection
- Assessment
- Prioritization
- Project management and tracking
- Preparation of originals for digitization
- Metadata collection and creation
- Digitizing
- Quality management
- Data collection and management
- Submission of digital resources to delivery systems and into a repository environment
- Assessment and evaluation of the digitization effort.

These components are part of a workflow with four primary phases, including:

1. Project Planning
2. Pre-Digitization
3. Digital Conversion
4. Post-Digitization work

⁴ Federal Agencies Digitization Guidelines Initiative. Digitization Activities. Page 14. 2009.
<http://www.digitizationguidelines.gov/guidelines/DigActivities-FADGI-v1-20091104.pdf>

In the project planning phase, the organization will want to select, assess and prioritize the documents to be digitized. This includes assessment of physical characteristics, audience, and whether the whole collections or a partial collection will be digitized. The documents considered should also be prioritized, important factors might include the value of the document or frequency of requests for the type of document. During this phase the organization should also determine who will perform the scanning, including whether it will be done in-house or outsourced and how the project will be communicated to the various stakeholders. Finally, there should be a plan in place for the tracking of digitization activities and documentation around procedures and processes to be followed.

In the pre-digitization phase, in addition to the items identified above, the organization will want to identify whether metadata or other descriptive information for documents to be digitized exists. If this information already exists, it can be very helpful in the digitization process and record organization. The metadata and naming convention requirements for the various documents should be defined at this point. Documents should also be evaluated for scanning readiness and any steps necessary to prepare the documents for scanning should be performed. Finally, the approach to scanning the documents, such as format and quality should be determined, quality control and assurance processes should be established, and file storage availability should be confirmed.

During the digitization phase, the process above and specifications identified should first be validated. Once validated, images can be scanned, metadata entered, optical character recognition (OCR) performed (as appropriate) and image and metadata quality evaluated. The scanned documents should also be indexed to ensure that documents can be effectively searched and retrieved. It is also important to confirm access to the documents have been granted to appropriate users.

In the post-digitization phase, the organization should finalize metadata, ensure that the digitized documents meet quality assurance and control standards, and make the records available to relevant stakeholders. As appropriate, the digital assets can then be exported to other systems in applicable formats. Finally, an assessment should be performed to review the overall project,

including cost-benefit, review impact on other activities, identify any ongoing issues to be addressed, document lessons learned and identify opportunities for process improvement.

Scanning Devices

Determining an appropriate scanning device depends on the types, sizes, and conditions of documents. Generally, there are three types of scanners that can be leveraged: flatbed, overhead, and sheet-feed. Flatbed and overhead scanners are generally versatile, not limited by the overall condition and the binding of the documents, but they may be limited with the sizes of the documents. Oversized materials must be scanned using equipment that features a scanning bed that is as large or larger than the object to be scanned.**Error! Bookmark not defined.** With sheet-feed scanners, they can expediently digitize documents with multiple workloads, however, they may not be appropriate for deteriorating documents. According to the Preservation Review of Scanning Equipment from the Library of Congress, form feed [or sheet-feed] equipment is not acceptable for scanning fragile, high-value, fine art, or special or archival collection materials.⁵ See Appendix I for additional details on the types of scanners.

For images that are larger than scanning devices, stitching also may be an option to consider. Stitching is an imaging processing method combining multiple overlapping images to create a single image. This can be used in scanning where a single scan of a large object is not able to produce sufficiently high resolution.⁶ The Town may be able to utilize this method with various software dedicated to image stitching after digitizing multiple parts of documents. However, this method may have a potential for distortion during stitching process.

File Formats

Another important consideration in any document digitization initiative is file format and scanning resolution. Recognizing this, the Town requested guidance around considerations to be evaluated. The Town's high-speed, sheet-feed, large format scanner includes SmartWorks Touch software, which supports scanning in a variety of formats, including PDF and TIFF, and optical

⁵ Collections Care. Library of Congress. Accessed April 26, 2019. <http://www.loc.gov/preservation/care/scan.html>

⁶ Stitching. Federal Agencies Digital Guidelines Initiative. Accessed May 14, 2019. <http://www.digitizationguidelines.gov/term.php?term=stitching>

image resolution of up to 1200 dpi, or dots per inch, which represents the resolution of the printed material.

TIFF, or Tagged Image File Format, is widely considered appropriate for image files, particularly in photography and records management. Being a “lossless” format, TIFF keeps every pixel of an image intact, enabling high-resolution image rendering. However, the TIFF format can yield large file sizes, so digital storage space constraints must be considered.

PDF, or Portable Document Format, is also widely known and utilized today and may come with features such as searchable text, embedded fonts, lossless compression, and high-resolution images.⁷ The PDF file format also has a subtype that can be used for preservation purpose: PDF/A, or PDF/Archive. PDF/A is an ISO-standardized (ISO 19005) subtype of the PDF format intended for long-term preservation.⁸ It should be noted that the PDF/A format disables some features, such as video, audio, JavaScript, and any executable file launches.⁸

While there are a variety of factors that can influence the quality of a scanned PDF, generally an image-embedded PDF, such as a map or plan, will include some form of compression, meaning that image quality might be slightly lower than a TIFF. While resource constraints need to be considered, it is generally a good practice to store a master image in the TIFF format, with a secondary service image in PDF format, which serves as a backup, is more easily accessible and smaller.

Image Resolution

Scanning documents with higher optical resolution is generally considered a good practice for cultivating a high-quality digital collection. Higher optical resolution leads to less reliance on software enhancement, or interpolation, which is to improve the quality of the scanned image by filling in extra pixels based on the scanning software algorithm. While interpolation is intended

⁷ PDF Family. Library of Congress. Accessed May 2, 2019.
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000030.shtml>

⁸ PDF/A Family, PDF for Long-term Preservation. Library of Congress. Accessed April 16, 2019.
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000318.shtml>

to improve the quality of scanned image, it may compromise the integrity of the scanned image by artificially filling-in pixels.

The Association for Library Collections and Technical Services, which is a division of the American Library Association (ALCTS), has published a variety of recommendations on digitization. Their recommendations range from 300 dpi to 600 dpi depending on the type of document.⁹ They provide image comparisons ranging from 200 ppi to 700 ppi. While historically significant or especially detailed documents may require higher resolution, 300 dpi is generally adequate for typical documents.

Digitization Methods

There are three methods to consider for conversion: backfile, day-forward, and on-demand conversion. Backfile conversion refers to the conversion process “...of replacing large volumes of documents with digital images.”¹⁰ In other words, this method goes backwards, putting emphasis on past documents over present documents. Day-forward conversion is “...the process of scanning, indexing and storing documents to an imaging system as they are produced or received in the normal course of business”.¹¹ Finally, on-demand conversion refers to digitizing municipal documents as they are requested. Determining an appropriate and realistic conversion timeline depends on how documents are prioritized for digitization.

Metadata

Establishing or collecting metadata or any other relevant archiving information is a vital process prior to or during the digitization. Metadata is defined as data that provides information about other data.¹² Metadata can be identified into three categories: descriptive, structural, and administrative. See Appendix II for definitions of the different types of metadata.

⁹ Minimum Digitization Capture Recommendations. Accessed April 19, 2019. <http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations#maps>

¹⁰ Backfile conversion definition. Webopedia. Accessed March 4, 2019. https://www.webopedia.com/TERM/B/backfile_conversion.html

¹¹ Day forward definition. Webopedia. Accessed March 4, 2019. https://www.webopedia.com/TERM/D/day_forward.html

¹² Metadata. Merriam-Webster Dictionary. Accessed April 17, 2019. <https://www.merriam-webster.com/dictionary/metadata>

Utilizing proper metadata fields are necessary for consistency and regularity of information organization. The Dublin Core Metadata Initiative, or DCMI, provides metadata best practices,¹³ endorsed by National Information Standards Organization and International Standards Organization. The Dublin Core Metadata Element Set, describes fifteen metadata vocabularies, which are broad and generic, and can be used to describe a wide range of resources.¹⁴

Quality Assurance and Control

Quality assurance and control are a vital part of the digitization process and ensure the quality of digitized documents. As such, each step in the process of digitization should include evaluation of whether technical requirements are being met for capture, image processing and delivery.¹⁵

Retention Schedules and Disposal Rules

The Town must adhere to all regulatory requirements for document retention and disposal, this applies to both physical and digital documents. The Secretary of the Commonwealth's Office has published several resources that may be useful, including an updated retention schedule and guidance on electronic records. Please refer to Appendix III for links to these resources.

Recommendations

The Town of Marblehead has shown much progress on the document digitization project. The Office of Municipal and School Technology has prepared recommendations for efficient and sustainable document digitization practice.

Digitization Protocols

The Town may want to document their digitization protocols and processes, such as scanning instruction, project progress, and factors related to quality control. As the Town is finishing the

¹³ About DCMI. Dublin Core Metadata Initiative. Accessed April 23, 2019. <http://www.dublincore.org/about/>

¹⁴ Dublin Core Metadata Element Set, Version 1.1: Reference Description. Dublin Core Metadata Initiative. Accessed April 23, 2019. <http://www.dublincore.org/specifications/dublin-core/dces/>

¹⁵ Session 7: Reformatting and Digitization. Northeast Document Conservation Center. Accessed February 27, 2019. <https://www.nedcc.org/preservation101/session-7/7digitization>

digitization of engineering plans and maps, documentation of digitization protocol may become a useful reference for other parts of the organization.

The Town has been digitizing in PDF format at 300 dpi to balance the file size and quality of digitized documents. According to the State Electronic Records Preservation (SERP) Framework from the Council of State Archivists, both TIFF and PDF/A formats are actively utilized for scanned images and texts.¹⁶ As long as the necessary details of documents are properly captured in scanning resolution ranging from 300 dpi/ppi to 600dpi/ppi, scanning in PDF format may be acceptable. As a part of future best practice, the Town may also want to consider incorporating the TIFF as a master copy format and the PDF as a service copy format. This will ensure that a high-quality, lossless master copy exists, while also retaining a backup that would be easier to consume for most interested parties. Furthermore, incorporating PDF/A format will ensure long-term digital preservation of the scanned documents. There are a variety of software solutions that can generate or covert PDF/A-compliant documents, including the SmartWorks Touch software the Town currently uses.

Acquiring Other Scanning Devices

Considering multiple factors for digitization, the Town may want to consider acquiring additional scanning devices. The Town has been unable to digitize some of the target documents, particularly bound books and documents in fragile condition. As such, either acquiring a flatbed or overhead scanner may be appropriate for the Town to continue with digitization. For further reference, Appendix I may be helpful in considering the options for the Town.

Metadata and Integration of Records Management System

To date, the Town has been capturing metadata based on file naming convention. While this is a meaningful first step, the Town might want to consider going a step further in establishing metadata, which could prove useful in the future for document retrieval by Town staff and the

¹⁶ SERP Framework. Council of State Archivists. Accessed May 6, 2019. <https://www.statearchivists.org/electronic-records/serp-framework/open-standards-neutral-formats/>

public. Leveraging a template, like the example shown in Appendix IV, might assist the Town in consistently and effectively establishing document metadata.

Integration of a records management system is often part of a digitization initiative. A records management system supports the creation and maintenance of a robust collection of municipal documents readily available to staff and constituents. Further, collection and storage of metadata for all digitized documents will support efforts to organize and access documents in the future. Appendix V provides questions to consider in assessing the viability and obstacles to consider in implementing a records management system.

Maintenance and Training

Document digitization is not just about digital conversion and records management; it is also about maintaining the established system and processes for Town staff and the public. While the Town Planner and Town Engineer are leading this practice, training for other Town staff members may support efforts to expand into other departments.

Potential Opportunities

The Town might want to consider pursuing other opportunities to support this initiative, given the interest in expanding digitization activities. One program the Town might consider is the [Community Compact IT Grant program](#) that supports IT capital investments in communities. It is a competitive program open to communities that have signed on to the Community Compact Best Practice program. Given interest in collaborating with other communities, the Town might also consider the competitive Community Compact [Efficiency and Regionalization program](#) that support regionalization and other efficiency initiatives aimed for long-term sustainability.

References

Care, Handling, and Storage of Works on Paper. Library of Congress. Accessed February 8, 2019. <https://www.loc.gov/preservation/care/paper.html>

Christensson, Per. "Lossless Definition." TechTerms. (2006). Accessed Apr 18, 2019. <https://techterms.com/definition/lossless>

Collections Care. Library of Congress. Accessed April 26, 2019. <http://www.loc.gov/preservation/care/scan.html>

Digitization Activities. Federal Agencies Digitization Guidelines Initiatives. November 2009. Accessed March 12, 2019. <http://www.digitizationguidelines.gov/guidelines/DigActivities-FADGI-v1-20091104.pdf>

Digitization Guidelines. Wisconsin Historical Society. Accessed April 1, 2019. https://www.wisconsinhistory.org/pdfs/la/Digitization-State/9_Digitization-Format-Guidelines.pdf

Dublin Core Metadata Elements Set, Version 1.1: Reference Description. DCMI Usage Board. June 2012. Accessed April 23, 2019. <http://www.dublincore.org/specifications/dublin-core/dces/>
Electronic Records Management Guidelines. The Office of Secretary of the Commonwealth of Massachusetts. Accessed March 30, 2019. https://www.sec.state.ma.us/arc/arcpdf/Electronic_Records_Guidelines.pdf

Flatbed Scanner vs Overhead Camera Digitization. Sustainable Heritage Network. September 2015. Accessed February 23, 2019. <https://www.sustainableheritagenetwork.org/digital-heritage/flatbed-scanner-vs-overhead-camera-digitization>

“Guidelines for Digitization”, UMass Amherst Libraries, May 2011. <https://www.library.umass.edu/assets/Digital-Strategies-Group/Guidelines-Policies/UMass->

[Amherst-Libraries-Best-Practice-Guidelines-for-Digitization-20110523-templated.pdf](#) (Accessed March 15, 2019)

Introduction to Records and Information Management Presentation. Council of State Archivists. July 2009. Accessed March 21, 2019. <https://www.statearchivists.org/programs/introduction-records-and-information-management-presentation/>

Make PDFs ADA Compliant. Cuesta College. April 2017. Accessed April 18, 2019. https://www.cuesta.edu/departments/documents/training_docs/Make_PDFs_ADA_Compliant.pdf

“Minimum Digitization Capture Recommendations”, American Library Association, June 1, 2013. <http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations> (Accessed April 2, 2019)

Number 1: Digitization Projects. Council of State Archivists. August 2017. Accessed March 4, 2019. https://www.statearchivists.org/files/6015/0272/2035/COSA_DigitizationProjects_final.pdf

DocuFi. PDF vs. TIFF, An Evaluation of Document Scanning File Formats. LinkedIn SlideShare. January 2014. Accessed February 11, 2019. <https://www.slideshare.net/DocuFi/pdf-vs-tiff-for-document-scanning>

PDF/A Family, PDF for Long-term Preservation. Library of Congress. Accessed March 29, 2019. <https://www.loc.gov/preservation/digital/formats/fdd/fdd000318.shtml>

PDF/A Frequently Asked Questions. PDF Association. PDF Association Staff. June 2015. Accessed March 7, 2019. <https://www.pdfa.org/pdfa-faq/?highlight=pdf%2Fa>

Recommended Formats Statement. 2018-2019. Library of Congress. Accessed April 21, 2019. <https://www.loc.gov/preservation/resources/rfs/RFS%202018-2019.pdf>

Recommended formats. UK Data Service. Accessed March 9, 2019.

<https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats.aspx>

Records Management for Municipal Governments. Municipal Technical Advisory Service. September, 2017. Accessed February 20, 2019.

<https://www.mtas.tennessee.edu/reference/records-management-municipal-governments>

Regulatory factors pushing ECM from TIFF to PDF. PDF Association Staff. PDF Association. March 2015. Accessed April 11, 2019. <https://www.pdfa.org/regulatory-factors-pushing-ecm-from-tiff-to-pdf/>

SERP Framework. Council of State Archivists. Accessed May 6, 2019.

<https://www.statearchivists.org/electronic-records/serp-framework/open-standards-neutral-formats/>

Session 7: Reformatting and Digitization. Northeast Document Conservation Center. Accessed March 12, 2019. <https://www.nedcc.org/preservation101/session-7/7digitization>

Sustainability Factors. Library of Congress. Accessed March 29, 2019.

<https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml>

TIFF, Revision 6.0. Library of Congress. Accessed March 29, 2019.

<https://www.loc.gov/preservation/digital/formats/fdd/fdd000022.shtml#identification>

TIFF-PDF/A White Paper. Property Records Industry Association. Published November 2003. Adopted March 2017. Accessed April 10, 2019.

https://www.pria.us/files/resource_library_files/Business_Processes_and_Procedures/2017-04-07_TIFF%20and%20PDF_FINAL_CLEAN.pdf

Understanding Metadata. Riley, Jenn. National Information Standards Organization. 2017. Accessed April 1, 2019.

https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf

Web Accessibility Checklist. Perkins School for the Blind. Accessed April 9, 2019.

<https://data.library.arizona.edu/data-management-tips/data-documentation-and-metadata>

You Say You Want a Resolution: How Much DPI/PPI is Too Much? Ashenfelder, Mike. Library of Congress. July 2013. Accessed March 30, 2019. <https://blogs.loc.gov/thesignal/2013/07/you-say-you-want-a-resolution-how-many-dpippi-is-too-much/>

Appendix I: Matrices for Choosing Scanners

	Overhead Scanner (i.e. Crowley ODS, Zeuschel OS 14000)	Flatbed Scanner (i.e. Context IQ Flex flatbed scanner)	Sheet-feed Scanner (i.e. Colortrac SmartLF SC, Fujitsu SP-1120 or SP-1130)	V-shaped Scanner (i.e. Qidenus)
Summary	Scanner with overhead lens and lighting	Flatbed scanner (either large format material or normal-sized material)	Feeding different sizes of documents for scanning	V-shaped scanner for books
Limitations	Can be impacted by ambient lighting conditions and typically cannot scan above 600dpi.	Scanning light, document sizes	Not suitable for fragile documents and cannot handle stapled documents	Only suited for scanning bound books up to the size of the scanner
Notes	Relatively versatile in scanning different documents	Relatively versatile in scanning different documents	Expedient processing for multiple workloads	May not be cost-effective
Best Suited For	Bound books, fragile documents and documents that may be damaged by intense direct light.	Bound books, letter and legal and fragile documents	Large format documents (depending on scale of scanner), letter and legal, small documents, high-volume documents	Bound books

Appendix II: Important Terms in Document Digitization

Records Management System: an application used for identifying, classifying, storing, securing, retrieving, tracking, and destroying or permanently preserving records (Wikipedia)

Bitonal: black and white

Metadata: information describing the history, tracking, or management of an electronic document, which includes file designation, create and edit dates, comments, authorship, and edit history (Electronic Records Guidelines)

OCR (optical character recognition): mechanical or electronic conversion of typed, handwritten, or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo, or from subtitle text superimposed on an image (Wikipedia)

Backfile Conversion: process of scanning past documents into digital format

Day-forward Conversion: process of scanning documents produced or received in the normal course of business

On-demand Conversion: process of scanning documents based on the document requests

TIFF (Tagged Image File Format): a file format developed by Aldus and Microsoft (later acquired by Adobe) for storing and interchanging raster images

PDF (Portable Document Format): a file format developed by Adobe as a general document representation language

PDF/A (Portable Document Format/Archive): electronic document file format for long-term preservation

PPI (Pixels per Inch): description of the resolution capabilities of an imaging device such as a scanner or the resolution of digital image (Federal Agencies Digital Guidelines Initiative)

DPI (Dots per inch): a representation of number of *dots* in an inch on prints; used with PPI interchangeably

Lossless compression: a method of reducing file size with no loss of quality

Descriptive metadata: information about the content of a resource that aids in finding or understanding

Administrative metadata: an umbrella term referring to the information needed to manage a resource or that relates to its creation

Structural metadata: metadata that describes the relationships of parts of resources to one another (NISO: Understanding Metadata)

Appendix III: Resources and Guidelines from the Office of the Secretary of the Commonwealth of Massachusetts

Municipal Records Retention Schedule – Quick Guide

https://www.sec.state.ma.us/arc/arcpdf/Municipal_Retention_Schedule_20190321.pdf

Massachusetts Records Retention Schedules

<http://retweb.sec.state.ma.us/retweb/retention/schedules.asp>

Electronic Records Management Guidelines

https://www.sec.state.ma.us/arc/arcpdf/Electronic_Records_Guidelines.pdf

Frequently Asked Questions - Born-Digital/Electronic and Digitized Records

https://www.sec.state.ma.us/arc/arcpdf/Frequently_Asked_Questions_Digital.pdf

Appendix IV: Example Template for Compiling Document Information based on Dublin Core Metadata Initiative

Title	Creator	Subject	Date Created	Last Date Modified	Department Responsible (Contributor)	Format	Type	Notes

- Headers may change depending on the actual criteria of municipal documents
- Prioritize which metadata is looked up the most for access (reflecting on the business process)
- Refer to Appendix VI for other metadata fields
- As a part of metadata, file directory location may be helpful for potential integration of records management system (e.g. C:\Users\John_Doe\Marblehead\ScannedImages)

Appendix V: Document Digitization and Management Checklist

Project Preparation & Planning

- ☐ Yes ☐ No Has a list of existing municipal documents been compiled?
- ☐ Yes ☐ No Have the priorities of digitization been set based on the information derived from the list?
- ☐ Yes ☐ No Have you established quality assurance and control rules and procedures for digitization processes?
- ☐ Yes ☐ No Are metadata being captured during the conversion process for future access and organization?
- ☐ Yes ☐ No Is the Town planning on integrating the records management system with existing business platforms?
- ☐ Yes ☐ No Is there IT staff to oversee and maintain the integration of the records management system with other existing business platforms?
- ☐ Yes ☐ No Are current business platforms used by Town built for integration to the records management system?
- ☐ Yes ☐ No Does the Town have backup plans in place for disaster recovery (e.g. data center, servers, and cloud backup)?
- ☐ Yes ☐ No Does the Town have access to appropriate equipment for proper conversion?

Maintenance & Training

- ☐ Yes ☐ No Has the Town established a protocol for future document digitization and records management process?
- ☐ Yes ☐ No Will Town leadership (Town Administrator, Town Clerk, Town Planner, Board of Selectmen, etc.) be briefed on regular basis on digitization process and protocol?
- ☐ Yes ☐ No Have staff members received proper training, and do they understand the need for the process?
- ☐ Yes ☐ No Will there be a point person assigned to handle the digitization?

Appendix VI: Dublin Core Metadata Element Set, Version 1.1 Terms and Definitions

Contributor: an entity responsible for making contributions to the resource

Coverage: the spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant

Creator: an entity primarily responsible for making the source

Date: a point or period of time associated with an event in the lifecycle of the resource

Description: an account of the resource

Format: the file format, physical medium, or dimensions of the resource

Identifier: an unambiguous reference to the resource within a given context

Language: a language of the resource

Publisher: an entity responsible for making the resource available

Relation: a related source

Rights: information about rights held in and over the resource

Source: a related resource from which the described resource is derived

Subject: the topic of the resource

Title: a name given to the resource

Type: the nature or genre of the resource