



OHIO PUBLIC LIBRARY INFORMATION NETWORK

2323 West 5th Avenue, Suite 130 • Columbus, Ohio 43204 • 614-728-5252 (phone) • 614-728-5256 (fax)

So you're thinking of having a vendor digitize your yearbooks.

Many libraries have had success using an outside vendor to digitize their yearbooks. This can be an easy and cost-effective solution. However, it's important to do your homework beforehand. This short guide is intended to help you ask the right questions about the digitization services you will receive from the vendor. While much of your decision will be governed by budget limitations, asking these questions will save you headaches in the long-term. Following the vendor questions, we have included a few points for you to consider as well.

Questions to ask the vendor:

What is the DPI of your images?

DPI, or dots per inch, is a standard measurement of digital images. PPI or pixels per inch is also often used. In an ideal world, all vendors would offer 600 DPI, but 400 DPI is more common. 300 DPI is commonly considered acceptable, but not ideal. The higher the DPI, the less likely it is that you will need to have the yearbooks rescanned in the future.

Can you send me a sample image of a yearbook in both color and black-and-white?

Having an image in front of you will help you assess its quality. It will also help you determine if the DPI offered by the vendor is adequate to suit your needs.

What file type will my images arrive in?

There are many different types of image files that are available. If the vendor will only offer you one file type, make sure you ask for TIFF files. FYI - Image files can be quite large and may take up a lot of storage space. The TIFF file format is a lossless image storage format that uses no compression at all—making sure you get the best image possible. Once you have the TIFF images, you can use various software programs to convert the TIFF files to other file formats. However, it's best to ask the vendor for JPEG and PDF as well, if possible, to save yourself the time of converting the files. (JPEG or PDF files are commonly used for web display.)

Will my yearbooks be returned to me in the condition I sent them?

Some vendors will debind, or take apart a yearbook, to get a better image. If your yearbooks are your library's only set, or if some of them are on loan, you want to be sure that the

yearbooks will not be harmed. On the other hand, if your yearbooks have a narrow gutter, or space in the inner margin, you will not get a good scan unless the yearbook is debound. Follow up this question with “If my yearbooks are damaged in shipment or processing, do you have insurance to cover that damage?”

What is the turnaround time on returning these materials?

If your library only has one set of yearbooks, it is helpful to know how long you will go without access to this information. This way you can manage your patron’s expectations.

How are files named? How will I know what order the pages go in?

Asking this innocent-seeming question beforehand will save you and your staff hours of work figuring this out afterward. One easy solution to is to request that the files be named in a title_year_page format. In this format, the first page of 1975 edition of “The Dart,” a yearbook from Ashtabula, would be label “thedart_1975_0001.tiff” for example.

What kind of quality control is performed?

Small mistakes are common in digitization projects: pages not lining up correctly, bent pages, or glare on glossy pages that makes the text or image unreadable. Making sure that the vendor has a mechanism in place to catch these types of mistakes will ensure that you don’t have to have the yearbooks rescanned later. Follow up this question with: “Will you pay to fix errors missed in quality control?”

Do you provide OCR?

OCR, or optical character recognition, is the process by which text is recognized as text by the computer. OCR is essential for making the document searchable and understandable by a computer. If the vendor does not offer OCR, there are solutions you can try, such as converting the image to a PDF. Once in PDF format, Adobe Acrobat can be used for OCR. Some special fonts, such as calligraphy, will not always be understood by OCR. Having the vendor supply the OCR will save you doing these steps.

How will two-page spreads be handled? Do I need to mark pages that should be digitized together?

Some yearbooks have text and images that span the gutter, or area between the facing pages of a book. If you have the left and right pages digitized as separate images, it may not make as much sense when viewed in your online collection.

Do you offer hosting as well as digitization?

Hosting, also referred to as web hosting, is a service that allows content to be placed on the web. Some vendors may offer online hosting of your yearbooks as well as scanning and other services. Before you ask this question, be sure to research how the vendor displays the online content to make sure you are satisfied with the look and feel of the online collection. Is it searchable? Is the interface intuitive? Also, be sure to ask what will happen to the content if the vendor ceases to exist in the future. Is there a long-term solution?

Before you send off your yearbooks:

- Look through all pages of your yearbook collection. Often pages will be missing or pictures will be cut out. Set any damaged yearbooks aside. If you do find damage in your yearbooks, you can try:
 - Asking your school library to borrow a copy.
 - Putting a call out in your community for donations through your website and/or social media outlets.
- For post-1923 materials, make sure you have permission from the school or copyright holder to digitize the materials.
- Consider where the images will be hosted for viewing. ContentDM, an OCLC product, is popular hosting service used by many archival collections and libraries. Another popular option is issuu, which is a hosting service used by digital magazines. Internet Archive may be another option, as they have free hosting. First consider how you would like your yearbooks to be displayed and used. Then investigate hosting options.
- If, in the future, you would like your yearbooks to show up in searches of the Digital Public Library of America (DPLA) you will need to create metadata, or data about your yearbooks, in a specific format or profile. One easy way to create metadata is to type the required fields in an Excel spreadsheet. From there, you can export the data as a .CSV spreadsheet. Consult OHC and DPLA metadata recommendations and take some time to think through how you want your metadata set up, before you upload images into your online collection. Often, this will be tricky to change afterwards.
- Put together a policy for when and how your library will take down images. Copyright of images in the yearbook can often become complex quickly. For helpful tips, take a look at OCLC's Well-intentioned Practice Guidelines:
(<http://www.oclc.org/content/dam/research/activities/rights/practice.pdf>)
- In addition, consider how your library will handle an individual's concerns about having their yearbook image displayed even if that individual is not the copyright holder. For example, some libraries have removed metadata about the individual, keeping the image in the collection, but making the individual's name unsearchable.