



Book Scanning Considerations

Scanning:

Book scanning is not an easy task and requires an operator to be present at all times, even when using a robotic book scanner to flip the pages.

As we flip through the pages of a bound book, the center of the cutter will effectively move from left to right in relation to the cameras by the dimension of the thickness of the book. The book cradle allows us to move the book to keep the gutter centered, however it is a manual process and not an exact science. If the book binding is tighter at the top than at the bottom it will cause either the left or right facing pages to skew slightly. We can adjust the yaw of the cameras so that the page is captured square, however it can still be noticed in the final scans.

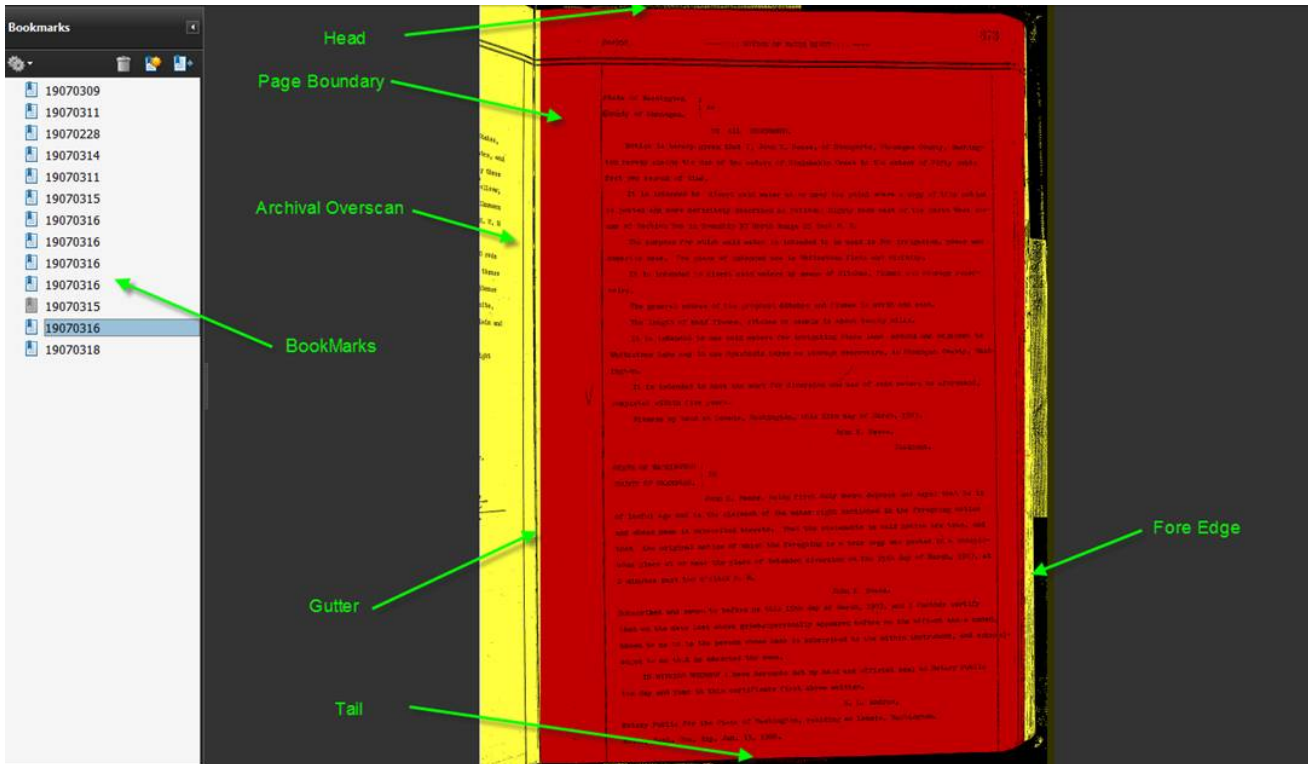
Generally pages are overscanned when captured to accommodate the movement of the book in the cradle, the movement of the gutter position and any skew caused by inconsistent or loose bindings. This overscan is indicated by the yellow area in the graphic below. (Typically the overscan at the gutter is not as extreme as in the graphic)

After scanning the operator is presented the scanned left and right pages showing where the page will be cropped. Each scan is then cropped to within 1/8" of the page border as the operator flips to the next page.

The resulting image has 1/8" over scan at the gutter, head, tail and fore edge.

If a customer needs this overscan area cropped off so that only the page content is displayed it requires meticulous adjustment on every page or manual cropping in a post scan process. Typically this can add 20% or more to the scan cost or more if the book bindings are very loose.

The fore edge of one side of the book will display similar to a fanned stack of cards and will be most noticeable at the first and last pages of the book.





Indexing or Bookmarking:

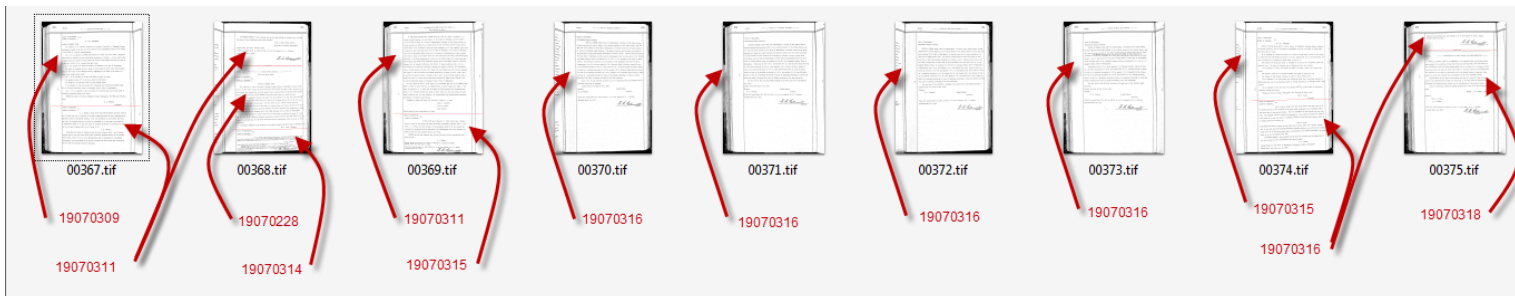
When indexing a book the service bureau and the customer need to keep in mind that the format of a book is designed to flow from one page to another, so a topic often will start in the middle of a page and flow to the next. This can present problems if the customer desires to split the book up into topics and publish each topic as a separate document. The best example is older County Recorders books where meeting minutes were taken in free form continuous fashion. In order to save space in the book the recorder did not start a new page for each meeting. Instead they continued mid page for each new meeting. The example below illustrates this and provides some solutions on how to deal with this format.

In the examples below I have drawn a red line where there is a split in a recording. You can see that in the examples provided we have two and even three recordings on a page.

This book was very challenging to index and publish since there were multiple recordings and the recording date was at the end of the recording not at the beginning. This meant that our operators had to page down to the end of the recording, key the date, then move it back to the beginning of the recording one or more pages previous.

In addition the recordings are not in date order and they have multiple recordings on the same date.

Below is a graphic that attempts to display the challenge:



There are two basic strategies with options on how to publish a book in this format:

1. Split the book for each recording or meeting date, then copy the image when it applies to more than one recording and redact as desired.
2. Publish the entire book as a single pdf which is linearized for quick loading and:
 - a. Bookmark the meeting dates (which are searchable)
 - b. Set a pdf destination (which is searchable and can be web linked)

The split book strategy is more common because it is easier for most vendors as they do not have the tools to publish and link to specific content inside of a pdf. With a little forethought and discussions upfront between a knowledgeable vendor and the customer, it can be quite easy to deploy books which are searchable and for which content can be linked.

An example of using a destination in adobe to open a book to a specific meeting is provided here >> [1907 March 11](#)