



Integrated Scanning of America

Adding Visual Navigation to Document Capture, QC, Indexing and Search Methodologies

By: Manuel Bulwa March 2021

A successfully digitized document collection requires both effective **search** capabilities in the software sold by document management (**DM**) **vendors** as well as adequate **capture**, often performed by separate **service providers**. These two vendor cultures frequently clash due to conflicting interests, business backgrounds and other reasons. A similar cultural clash occurs when Professional Services (**PS**) organizations, Record Managers (**RM**) and **IT** staff are involved as stakeholders. The end user is quite often at the mercy of the survival of the fittest, not always resulting in their best interest.

Most **DM** products provide **search** capabilities based on a combination of a handful of structured index fields and full text search (**FTS**). Structured indexing is handled by powerful database query technology while full text search is hyper energized by AI technologies such as natural language processing (**NLP**), text analytics, etc.

So far, structured indexing and full text search are two legs supporting the proverbial search stool, so we need to add a third leg: **Visual Navigation**. Users frequently supplement their available search tools by visually navigating entities such as folder tree structures, bookmarks and thumbnails. These visual entities are captured from covers, folders, staples, clips, divider tabs, barcoded target sheets, envelopes, microfilm blips, email attachments, title blocks, headers, footers, logos, etc. Some of these markers are frequently ignored in document capture projects or conflated with indexing data. They convey valuable information that makes the art of organizing and finding more pragmatic. It helps to capture these entities in **context**: Physical context establishes a sequential relationship between document breaks. Logical context responds to semantic and ontological relationships found in the document text.

It is worth mentioning a less frequently used methodology that can work small miracles in certain scenarios: Binary Search (I will not use the acronym for obvious reasons). Binary Search is a combination of automation and visual navigation that relies on responding to Yes/No or Lower/Higher questions to locate sorted records. In less than n clicks a user can locate any of **2n** records with zero indexing cost (some prefer to say that the indexing costs are deferred to the search effort).

Although the industry has produced good technology to semi-automate capture, the human factor is critical all the time to reduce the risk of feeding imperfect data to the DM solution.

Visual Navigation markers are small images that can be created as a combination of:

Rectangles containing the image of divider tabs, title blocks, headings, dividers, etc. These rectangles can be drawn with the mouse and/or computer automation.

Barcoded sheets strategically placed during prepping. Example: to mark clipped or stapled pages.

Rule based extraction of textual headings, logos, image patterns.

As the markers are mostly raster images, they are not necessarily searchable, thus the need for visual navigation of graphic bookmarks.

An example of a combination of structured, FTS and visual navigation follows: A structured index query narrows down potential document candidates to a manageable number, which is then further narrowed down by full text search. This produces a small number of documents, one or more of which may satisfy the original search. As each document may have many pages, to decide which document qualifies as a successful hit we should navigate it visually. In the example below, visual markers

were captured as a combination of mouse rectangles, barcoded sheets on stapled or clipped pages and automated capture of text headings. As you navigate through the multiple pages of the selected document, the corresponding visual marker will be highlighted in blue (left pane). Conversely, if you click on a visual marker, the first page pointed at by that marker will become visible (top right, blurred to protect confidentiality), showing any highlighted keywords used in the full text search (green box, bottom right):

This visual navigation adds value to human interpretation of the structure of a document. For example:

The fact that certain pages were stapled or clipped together means something. This meaning would have been lost (or hard to detect) without a staple marker (notice the paper clip icon).

The fact that a paper document was organized with divider tabs means that certain sections required more attention than others for different search or classification purposes.

Objects in a collection such as CDs, DVDs, envelopes, 3D samples, etc. also carry valuable information. A visual marker will capture at least some of their meaning without interfering with structured indexing schemas.

In architectural or engineering drawings, as well as in historical or geographical maps, the title blocks contain valuable visual qualifiers such as revision numbers, permit numbers, classification categories, names, etc. Although some of it could be captured as structured indexing, a visual (raster) marker is in my opinion more efficient.

Multilevel blips in microfilm rolls are examples of visual markers that facilitate locating microfilm records.

Some collections are much harder to classify and organize than others. Over-indexing could be as bad or worse than under-indexing, so visual navigation may come to the rescue and facilitate a practical compromise.

After some time using visual markers, the user may realize that some should be made searchable by adding text to the raster images. In my experience automatic extraction such as OCR or ICR frequently overpromises and under-delivers, while human data entry is more reliable. Fortunately, the visual marker images naturally lend themselves to be used as snippets that can be distributed to remote (even off-shore) data entry operators, volunteers and interns, making human capture more efficient and less costly. This effort could be described as deferring some of the cost of indexing to points in time where users learn through actual use to determine which markers become relevant enough to be promoted to the indexing level.

Besides paper documents and books, visual navigation proves very helpful in other areas. Examples:

Smart video inventorying : used to reduce collection manifest/inventory creation costs and liabilities:

An operator wearing smart glasses captures (hands-free) a video of the contents of a box or drawer. By "contents" I mean the labels or tabs on every folder in the box or drawer. Using his/her hands, the operator simply makes sure that the entire label is exposed for a fraction of a second before exposing the next label or tab. A typical box or drawer can be captured in two to four minutes, depending on how packed the folders are, their orientation and their location.

A legible still frame is extracted by our software off the video for each label. Structured indexing is limited to first and last label of each box. Everything in between can be located in a handful of clicks without costly label-by-label indexing. For example, a collection of over 30,000 documents would require less than 15 clicks to locate any document or determine that it does not exist. Producing a conventional box/drawer catalog that lists each and every label is at least three times more expensive without using this methodology. See also: <https://youtu.be/9TtKt5HGiO>

COM Fiche: When a COM fiche collection is digitized, it can be prepared to use one or more of the following indexing and visual navigation options:

- Fully index each individual record and populate a content management software database.
- Use special software to locate digitized image frames by only indexing index ranges of every sheet of microfiche (from/to).
- Use Microsoft Windows® file system to name each sheet after its title block. This option imitates digitally the mechanism used when the fiche was managed manually through reader-printers.
- Use bookmarks to locate each frame within a multipage PDF per sheet. This option also imitates digitally the mechanism used when the fiche was managed through reader-printers.

Every fiche sheet has a title block that people used to search manually on reader printers. Now they can be visible

as pdf file names that can be navigated visually and/or through Windows search. Double clicking on a named file will get a PDF file to show all frames in that fiche sheet. The last frame of a sheet ("fiche index frame") contains the coordinates of the row and column that will show the starting frame of the document sought. All you need to do now is to click on the bookmark matching that coordinate and visually navigate the document frames:

An improved variation of this method is based on defining "hot spots" in the fiche index frame, so clicking on a coordinate will link straight to the image without having to use the bookmarks.

Large Format Drawings and Maps: An architectural drawing of a property **set** typically includes one or more sheets depicting construction tasks for various sections of the property. After a quick search, the few sheets of the selected set can be visually navigated through a small number of images of the title blocks. This makes possible the finding of records:

- without the cost of keying in certain title block data fields such as sheet number, revision number, section, task, etc.
- without the risk of keying or recognition errors and the cost of subject matter expertise to decide which ones to capture.
- adding the convenience of improvising selection criteria on-the-fly by seeing stamps, signatures, hand written annotations, logos, etc.

Microfilm rolls: If the filming is blipped and that a catalog or manifest exists naming each blip, then the entire roll could be contained in a bookmarked PDF file. The bookmarks can be navigated visually and/or via text search. Producing separate smaller PDF files per bookmark is not recommended, as blips are known to be less than perfect, so document frames may extend beyond a blip.

In summary, visual navigation provides the third leg needed to stabilize the search stool. It can also save money, defer costs, reduce errors and facilitate progressive indexing as dictated by use and practical experience.

Questions and additional levels of detail are available by request emailing the author mailto: mbulwa@isausa.com



2363 Teller Road Unit 102 Newbury Park, CA 91320

Phone: (805) 375-0422