Smart capture of large format drawings

By Manuel Bulwa

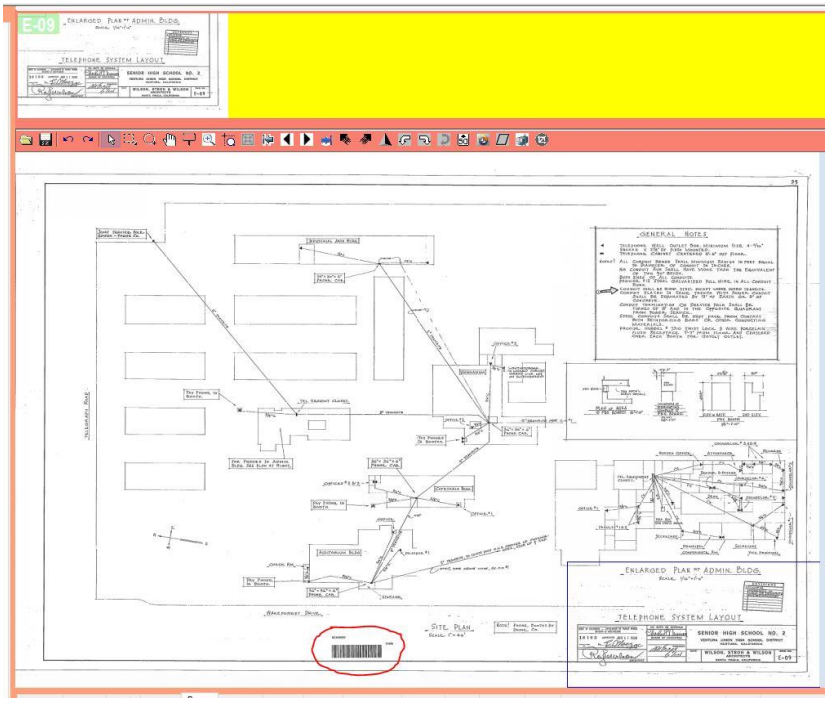
January, 2021

Large format drawings (Engineering or Architectural) are typically digitized in bulk during a backfile conversion phase that may last for weeks or months. During and after that period, some of the sheets in the collection need attention and/or changes. The problem here is that users cannot wait until these particular drawings are fully automated under a Document Management System (DMS), so they physically pull paper originals from their archive location. This action carries some risks and costs, whether they require changes/updates or not.

To reduce or eliminate these risks and costs, it imperative to accelerate the availability of fully validated and indexed digital records, and to reach near-zero tolerance accuracy level in the indexing efforts. In a perfect world, any user at any time will be able to retrieve digitally the latest version of any drawing (or supporting documents), and to always locate the right sheet for changes. In the real world, we face multiple challenges such as:

- Drawings may be needed before their digital counterparts are under DMS control.
- Version numbers are sometimes not fully up to date or visible.
- Index data visible on title blocks, stamps, etc. are usually ambiguous. This makes it difficult to assign a unique digital identity for every sheet both on paper **and** digitally. In other words, it is often hard or impossible to be absolutely certain that the piece of paper in the user's hands matches exactly its digital counterpart. Even if it does match from an indexing perspective, it may not match from an imaging perspective, as small differences may not be apparent.

This article introduces a methodology based on assigning a visible barcode identity to a drawing. It only addresses updates to a drawing made on paper, leaving changes made digitally on images to a separate discussion. This methodology maximizes the certainty that the piece of paper we are working with correlates unequivocally to an image in the digital collection.



The red circle in the figure illustrates the positioning of a barcode sticker in a drawing:

There are two basic methods to produce barcode stickers:

- Stock labels: Just sequential, consecutive numbers with no relation to metadata.
- Barcode labels printed off lookup files. They are still consecutive numbers, but they are univocally tied to your database entries. They will be printed in the same order in which you

currently have your physical drawing sets filed. This makes barcoding very straightforward. Some may be out of sequence, and “widows and orphans”, but we have mature, easy solutions for these cases.

Notes:

- Any changes or annotations made to the paper drawing should be such that they do not obliterate or obstruct the barcode.
- If circumstances are such that a paper copy is needed, then either a new barcode identity must be assigned to the copy, or the original needs to be disposed of or marked as “Do Not Use”.
- Follow a check-in/check-out protocol when removing a sheet from its set, to document the chain of custody, set the “status” flag as “Under Controlled Revision” and start the update workflow.
- Whenever possible, perform the changes digitally using a collaborative editing tool.
- Barcodes should be placed on any empty area within the drawing, preferably on its borders, or even in the backside on extreme cases.

We assume that all research conducing to a decision to update a drawing is made by navigating the digital collection. A well-planned process would identify (directly or via cross-reference) the physical location of the set containing the particular sheet.

If the paper drawing just located has already been assigned the barcode identity, then it can be submitted to the update workflow. If instead the paper drawing just located has never been assigned a barcode identity, then one of the following options is followed:

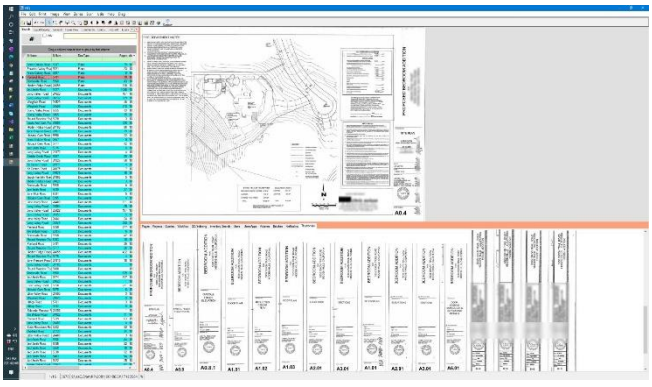
- A request is made to print a barcode sticker, which must be affixed to a blank area near any border.
- A request is made to print a “Smart Barcode Sheet”, which must be clipped or stapled to the drawing, usually in the back.
- A request is made to print an identical copy that does have the barcode identity embedded. The software looks for the whitest area on any border to embed the barcode. The original must be marked as “Substituted” or disposed after that.
- If the user decides not to use identity barcodes, then the user, at its own risk, should use its best judgment to locate its digital counterpart and flag it as “Under Uncontrolled Revision”.

Aside from managing changes and updates, this technology also applies to “near 100%” QC processes, which frequently rely on counting sheets and comparing against file counts. Counting sheets to control scanning integrity cannot be fully trusted, as it is likely that plus/minus (excess/defect) miscounts may yield false positives.
Using identity barcodes introduce several supplemental conveniences, such as the ability to handle sequence “holes” in the barcode numbers, ability to individually check-in/check-out sheets of a set, ability to reconstruct the original order within the set, ability to control future revisions or duplicates and more.
The barcode technique, especially if sets are tied to metadata, worked exceptionally well for us numerous times due to the following reasons:

- Index values are not always conspicuous in the drawing; some may not even be present in some sheets. Hunting for index data in a drawing is inefficient. Assigning a subject matter expert (SME) to index is expensive. Assigning a non-SME to index with insufficient expertise in your document data is risky. The barcode technique ties the physical paper world with the digital world of existing databases, facilitating accurate population of index data.
- In some cases, we need to separate paper to be scanned on separate production lines, such as overhead scanners, large format scanners and standard size scanners. In this case, the barcode sheets are critical to assure a smooth reconciliation of scanned image files and paper originals.
- The preparation and placement of barcoded sheets serves as a valuable "manifest/inventory" of records to be sent for scanning.
- The accounting of recognized barcodes and the continuity of their sequencing serve as valuable checks and balances methods against the inventory.
- The ability to deal with "widows and orphans", i.e. sheets with barcodes extracted off your files but with no matching drawings to scan, and conversely orphan drawings for which no barcodes were produced. This offers extremely valuable information about possible inaccuracies in your electronic records and proof of missing physical records.

Finally, following is a summary of other techniques we at ISA use to address the many challenges of large format drawings:

- **Raw Capture**: We always scan in color using the healthiest edge first to feed, regardless of the requested deliverable formats and rotations. The color originals are a valuable archive set that may eventually be needed to avoid rescans. For original such as translucid Vellum sheets and difficult blueprints we deliver a table of contents spreadsheet with all index data items and 3 hyperlinks: color, grayscale and bitonal. We also deliver software that can read the color originals and perform various conversions, cleanup and extractions to fine tune particular needs.
- **Title Block Thumbnails**: 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 brute force located by visually navigating through a small number of images of the title blocks. This makes possible the finding of records: a. without the cost of keying in certain title block data fields such as sheet number, revision number, section, task, etc. b. without the risk of keying or recognition errors and the cost of subject matter expertise to decide which ones to capture. c. adding the convenience of improvising selection criteria on-the-fly by seeing stamps, signatures, handwritten annotations, logos, etc.



- **Twin Barcode Sheets**: We pre-print pairs of identical twin barcodes. They are used to physically split portions of originals to diverse production lines (business size, overhead, large format scanners). One of the twins stays with the main set of pages, signaling the point at which sheets were detached and sent to a separate scanner preceded by the identical twin. The scanned images are perfectly and accurately reconciled by barcode

recognition and automation. The physical presence of the barcoded sheets makes the manual process of reconciling originals a breeze. This technique also applies to 3D objects such as CDs, DVS, samples, etc.

A separate white paper will elaborate on other techniques, including physical prepping of old, damaged, brittle, dusty originals.