Managing Legacy Paper Files in the Digital Era

Addressing your organization's legacy paper files and capturing them in your digital information ecosystem may feel like a daunting task. As discussed in our recent white paper, *capture* is the first step in the information lifecycle and is essential for achieving digital transformation and enabling the strategic alignment of information activities envisioned by information governance (IG). Paper files are effectively "dark," inaccessible to your organization's digital information ecosystem *until* they are captured through scanning, either imaging or digitization.[1]

Like many organizations, you are probably facing a mountain of file cabinets and boxes, years' worth of records and documents, with a level of uncertainty of what information lies within. This begs the question *Where do I even start?*

Here are a few things to keep in mind as you plan a scanning project.

Break Down a Scanning Project by Prioritizing Which Paper Files Should Be Captured First

If you have a specific driver for your legacy paper file scanning project, like e-discovery in a legal case, it's pretty clear which files you should target—those relevant to the matter. But, if you are scanning legacy paper as part of a more general effort (e.g., "going paperless"), a starting point may be less clear. Most organizations have enough legacy paper that a "scan everything" approach, without prioritization, can be hard to get off the ground and can quickly overwhelm

---

[1] Imaging is scanning to a picture-based format. Digitization is scanning with some type of additional text or metadata extraction. See ARMA Capture Framework.

budget limits and other resources. Where possible, break down a massive backfile scanning project into smaller, more manageable (and fundable) steps.

Consider starting with paper files that are accessed most frequently and tiering subsequent phases by frequency of access. Files stored on-site are likely candidates for a starting point. They may be on-site specifically *because* they are used often, but confirm this assumption with their owners before you scan.

For files stored off-site, review the access logs to determine which boxes are recalled most frequently. Prioritizing the scanning of paper files by their frequency of access will target those files that are in the more active phases of the information lifecycle, and it can result in immediate and measurable gains as improved ease of access, more effective collaboration, and reduced costs in recalling files from storage, to name a few.

Prioritizing by access frequency is one approach. Consider what makes the most sense for your organization (e.g., Would it make more sense to prioritize by department? Or, by some combination of factors?). The key is breaking down a seemingly insurmountable endeavor into smaller, more manageable projects, and tackling them in an order that makes sense for the priorities of your specific organization.

Understand Your Legal, Regulatory Compliance, and Business Needs with Respect to Your Information Before You Scan

An effective IG program applies a consistent lifecycle for all information in your organization, no matter its form or system. Consult your retention policy, schedule, and relevant stakeholders to understand what obligations you have for information that exists in paper form *before* you begin scanning. (If you identify gaps in your understanding or current approach, a scanning project is an opportunity to address these issues.) If you do not understand what information you have, why you must retain it, or for how long, scanning may effectively replicate information management problems that exist in your paper files into your digital information environment. This can be costly and increase risk.

An obvious example is scanning files you no longer need or that are nearing the end of their lifecycle. If, after consultation with relevant stakeholders (e.g., legal, RIM, business, etc.) and consideration of applicable policies, holds, and your retention schedule, it is determined that there is no legal, regulatory compliance, or business need to retain specific files, scanning only creates another copy of information that should properly be a candidate for defensible deletion. Similarly, if files are about to reach the end of their lifecycle in the near term and will be eligible for defensible deletion soon, it may make more sense to manage them for the remainder of their life *as paper.* Scanning is expensive in terms of time and resources and scanning documents at or near the end of their lifecycles can be a costly and avoidable mistake. Scanning also creates a copy and increases the risk that your information won't be consistently

handled because it exists in more than one place and form. Here, you would have a paper and electronic form to delete.

Scanning creates a digital copy of your paper files, but it is the information itself that must be governed. Understanding what is in your paper files and your obligations with respect to that information allows you to include that information in the digital systems in which you are capturing those scanned files (e.g., tagging or extracting metadata to establish retention categories and trigger dates allows you to enforce your retention schedule in the digital system). While it might be possible to shred paper after a conversion project, in many cases organizations are obligated to, or choose to, keep paper files along with the digitally captured copy. Making sure that paper file tracking and management information is up to date helps ensure that your information is handled consistently according to your organization's legal, regulatory, and business needs, in both paper and digital form. If you maintain information in paper and digital forms, consider cross-referencing to ensure that changes in legal, regulatory compliance, and business obligations are reflected consistently across both.

Consider Shredding Unnecessary Paper or Finding Better Storage for What You Keep

A scanning project creates an opportunity to destroy unneeded paper. If you have determined that you have no reason or obligation to keep specific paper files or paper copies of what you have scanned, getting rid of boxes offers obvious savings in terms of storage costs and effort in double-bookkeeping, described above for information that exists in multiple forms. Consider *securely shredding* what you no longer need, again, *only after* you have determined you may do so, and confirm the approach your organization wants to take. For remaining files, move them to secure storage.

Learn More About Capture

The above are just a few things to keep in mind as you start a legacy paper file scanning project. To learn more about capture and considerations for imaging and digitization of paper, download the free ARMA/Access white paper, *Effective Capture: The Foundation of Information Governance and Digital Transformation*, or view the ARMA/Access webinar "Information Governance and Digital Transformation Must Begin with a New Understanding of Capture."