



ARTICLE

CONTENT, DATA, AND THE PATH TO INTEGRATION



TABLE OF CONTENTS

THE RECORDS YOUR EHR DOESN'T SEE	3
WHY ARE MEDICAL RECORDS DIFFERENT?	4
WHAT BELONGS IN THE DIGITAL RECORD?	5
GETTING BACKFILE DATA INTO THE EHR, NOT JUST A FOLDER	5
THE COST OF INCOMPLETE RECORDS	6
BALANCING SPEED AND ACCURACY	6
SECURITY AND CHAIN OF CUSTODY DURING CONVERSION	7
BRINGING IT ALL TOGETHER	7

Figure 1: Departments That Rely Most on Paper-Based Workflows

Not surprisingly, healthcare organizations have invested heavily in electronic health records (EHRs) over the past decade. A modern EHR gives clinicians a single place to view a patient's history, coordinate care, and document treatment. In addition, it replaces what used to be a fragmented mix of paper charts and disconnected systems. Even so, an EHR is only as complete as the information inside it. For many healthcare organizations, a meaningful share of patient history still lives outside the system entirely, in paper charts, historical files, and records that were never converted when the EHR went live.

This gap matters more than you might think. A clinician who proposes a treatment without full visibility into a patient's history is working with an incomplete picture, which can lead to repeated tests, delayed diagnoses, or insufficient context when making care decisions. This gap needs to be closed, but that doesn't have to mean replacing the EHR or starting over. It's most important to bring the rest of the record (i.e., the paper that came before digital adoption) into a form that the EHR can actually use. In high-stakes environments like oncology, allergy management, or emergency medicine, a missing paper note isn't just an administrative delay—it is a direct clinical liability that can lead to adverse drug events or redundant, costly, invasive procedures.

The Records Your EHR Doesn't See

Even businesses that adopted an EHR years ago might still carry a substantial paper archive. According to licensed industry research by Keypoint Intelligence®, healthcare organizations report that an average of 46%¹ of their business content is still paper-based. Furthermore, 41% of respondents state that more than half their content remains on paper. Separately, Keypoint Intelligence identified digitizing and managing legacy patient records as one of the top document-related priorities for healthcare organizations today.

Legacy records predate the EHR itself, with charts going back years or decades, as well as referral letters, consent forms, imaging reports, and records inherited through mergers, acquisitions, or physician retirements. This information sits in off-site storage, basement archives, or filing rooms that will outlast the staff members who remember what's in them.

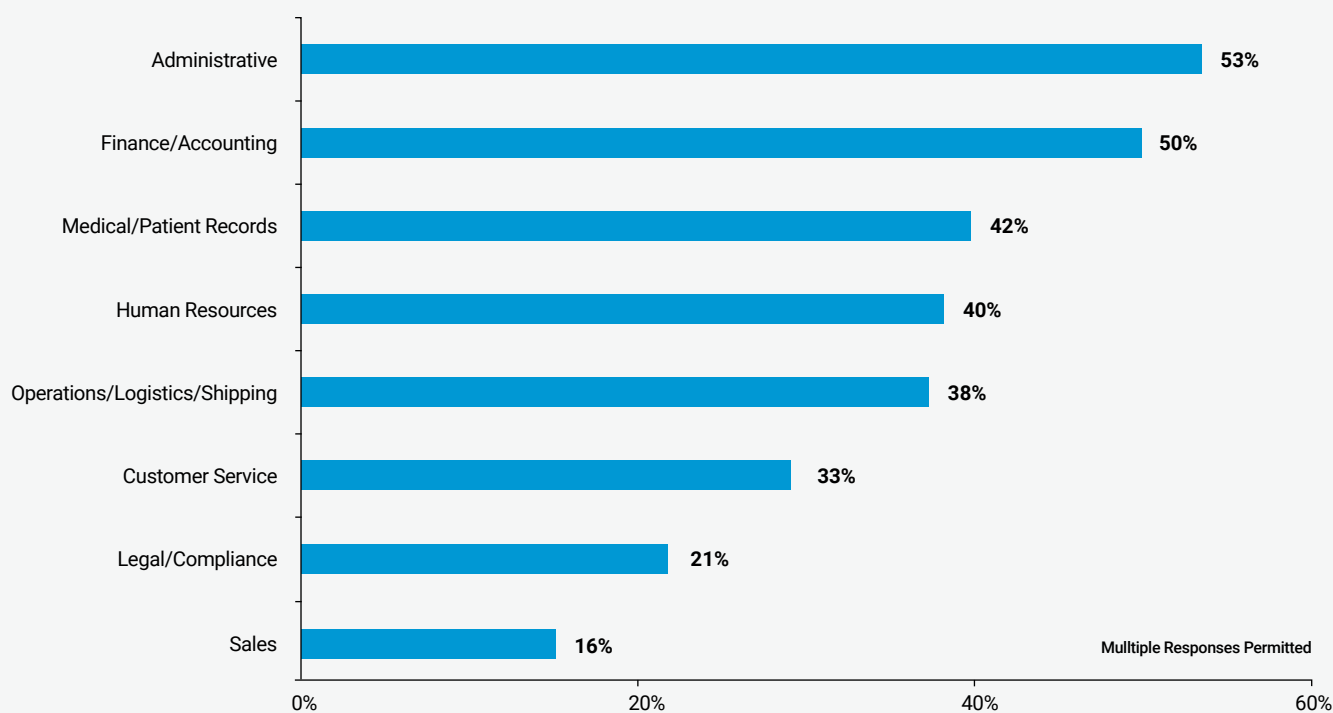
A practical, systematic first step helps address the scale of the problem: Conduct a basic archive inventory, counting boxes or linear feet by record type and estimated date range, rather than attempting to catalog every individual file. That inventory alone is usually enough to size the project realistically and identify which parts of the archive are worth tackling first.

Why Are Medical Records Different?

Not every department carries this weight equally. Keypoint Intelligence's licensed IT Decision-Maker research found that medical and patient records were among the most paper-reliant content types in healthcare. Roughly 42%² of healthcare organizations cited this choice, which was more than double the average across other surveyed industries. That gap says something about why healthcare records resist digitization more than other document types.

Figure 1: Departments That Rely Most on Paper-Based Workflows

Which of the following departments within your organization rely on paper-based workflows the most?



N = 41 Healthcare Respondents in the United States

Source: Office Group Printing and Scanning Survey; Keypoint Intelligence 2025

Patient records could include decades of handwritten notes, signatures, and annotations that don't scan cleanly into structured data. Retention requirements mean older files can't simply be discarded once a patient stops actively receiving care, and those requirements vary by record type and state. A single record may also need to be referenced by multiple departments (e.g., imaging, billing, referring physicians, legal), so the format has to remain usable well beyond the point of initial capture.

Because these constraints vary so much by record type, a helpful strategy is to sort the archive into a few general categories (e.g., active clinical records, closed cases, records nearing the end of their retention window) before deciding how each should be handled. That sorting process usually reveals that only a fraction of the archive needs the full weight of clinical-grade digitization.



What Belongs in the Digital Record?

Not every page in a legacy archive needs the same level of treatment. Full integration that is indexed, tagged, and searchable inside the EHR makes the most sense for information that clinicians reference often. These might include recent visit summaries, active medication lists, known allergies, and chronic condition histories. Older or rarely accessed material, closed cases, historical billing records, or documents tied to patients no longer receiving care may only need to be preserved and retrievable, not fully structured into the clinical workflow.

Attempting to apply a single standard to the entire archive typically wastes effort in one direction or leaves real gaps in the other. Full integration for records that nobody references will cost more than it returns, while archiving everything at the same shallow level can make active clinical information harder to find than it should be. By using the categories from the sorting step above, medical teams can decide upfront which records need full clinical integration versus simple digital archiving. As a result, efforts can be directed to processes that actually improve care.

Getting Backfile Data into the EHR, Not Just a Folder

Digitizing a paper record and integrating it are two different things. A scanned PDF sitting in a folder will be more accessible than the original paper, but it still requires someone to search, open, and interpret it manually. In other words, the information is technically digital but functionally still disconnected from the rest of the record. Real integration means ensuring that the content is indexed, tagged, and structured appropriately. It should surface inside the EHR itself, be searchable alongside everything that is entered digitally, and remain visible to a clinician without a separate lookup step.

A practical approach involves prioritizing by use, not by volume. Teams should identify which patient files are still being actively referenced (e.g., recent discharges, ongoing chronic care patients, active referrals) and digitize those first. Older, inactive records can be handled at a later time. This ensures that the highest-value information is immediately usable without requiring a multi-year project before any benefit is realized. It also gives staff members and leaders an early, visible result that helps sustain support for the rest of the effort.

The Cost of Incomplete Records

When patient history isn't fully available inside the EHR, the cost shows up in places that are easy to overlook. For example:

- Clinicians may order tests that were already performed simply because the results live in a paper file that no one thought to check.
- Referral coordinators may spend time tracking down records that exist but aren't easily accessible where they're needed.
- Patients may be asked to repeat information they've already provided because their history isn't visible to the person asking.

These costs rarely show up on a single line item, which is part of why they're easy to underestimate. It may be helpful to track a small sample of patient encounters over a few weeks, noting any point where staff had to search for, request, or re-collect information that should have already been on file. This type of lightweight audit turns an abstract efficiency argument into specific, countable instances that tend to carry far more weight when making the case for prioritizing backfile work alongside other technology investments.

Balancing Speed and Accuracy

Backfile conversion projects often face pressure to move quickly, especially when a large archive is involved. In a clinical setting, however, a misfiled record or a data entry error carries real consequences. It might result in a note attached to the wrong patient, a missed allergy flag, or an incomplete medication history.

Building validation directly into the conversion process helps catch these issues early on, before they multiply across a larger archive. Sampling a percentage of converted records for accuracy review, confirming that index fields match the source document, and involving clinical staff in spot-checking high-risk record types all reduce the chance that speed will come at the expense of reliability. A short pilot phase, converting a small batch and reviewing results before scaling up, is often the fastest way to uncover problems with indexing rules or document quality.



Security and Chain of Custody During Conversion

A backfile project creates a temporary window where a patient's information exists in two places at once: The original paper and the new digital copy. That window needs its own safeguards, separate from the ongoing security practices already in place for day-to-day operations.

Access to records during conversion should be limited to authorized staff and logged, just as access to the finished digital record would be. Once a document has been verified as accurately captured, organizations need a clear, documented process for retiring the paper original in line with their retention policy. Both versions should not persist indefinitely. The conversion window itself is a security boundary, not just the before and after. This distinction matters most when outside vendors or temporary staff members are involved in the conversion work.

Bringing It All Together

Backfile digitization doesn't have a fixed endpoint. New paper will continue to enter healthcare firms through referrals, outside records, and patients transitioning from other providers, even in organizations that are otherwise fully digital. Treating digitization as an ongoing capability keeps the EHR complete as new content arrives instead of leaving it accurate only to the day when the initial project wrapped.

Businesses that build this strategy into standard practice, while revisiting it as new paper accumulates, are better positioned to keep the full patient record where clinicians actually need it. The strategic objective is to help ensure patient history is complete, accessible, and actionable at the point of care, supporting better-informed clinical decisions.

If you're ready to take the next step, connect with Canon to explore what's possible for your organization.

For more information about how Canon U.S.A. can help your organization with your digital transformation goals, call 1-844-50-CANON or visit usa.canon.com/business.

¹ Source: Office Group Print and Scan Survey; Keypoint Intelligence 2025 (N=41 Healthcare respondents in the United States)

² Source: Digital Transformation IT Decision-Maker Survey; Keypoint Intelligence 2025 (N=50 Healthcare respondents in the United States)





1-844-50-CANON

usa.canon.com/business



The opinions expressed represent our interpretation and analysis of information generally available to the public or released by responsible individuals in the subject companies. We believe that the sources of information on which our material is based are reliable and we have applied our best professional judgment to the data obtained.

This material is prepared specifically for Canon U.S.A.

Canon is a registered trademark of Canon Inc. in the United States and elsewhere. Keypoint Intelligence is an independent authority on the document imaging industry, whose proprietary research is licensed by Canon for this white paper. All other referenced product names and marks are trademarks of their respective owners and are hereby acknowledged. Neither Canon Inc. nor Canon U.S.A., Inc. represents or warrants any third-party product or feature referenced hereunder. The authors and publishers of this content are not engaged in the rendering of professional advice or services including, without limitation, legal or regulatory advice or services. Individuals and organizations should perform their own research and conduct their own due diligence concerning the products and suggestions discussed in this white paper. Canon does not make any warranties concerning the accuracy or completeness of the opinions, data and other information contained in this content and, as such, assumes no liability for any errors, omissions or inaccuracies therein, or for an individual's or organization's reliance on such opinions, data or other information. Canon U.S.A., Inc., One Canon Park, Melville, NY 11747