ANCHORGRID WHITE PAPER

The Hidden Economic Cost of Repeated Document Extraction in Construction Workflows

Why repetitive extraction work consumes skilled attention across estimating, procurement, and project management workflows

2026

Executive Summary

Document extraction (the process of locating, reviewing, verifying, and retrieving information from a page within a document) is a frequent occurrence in many construction businesses, and is embedded across estimating, procurement, coordination, submittal review, and change management tasks. Construction-based document extraction regularly involves architectural drawings and project specifications, two critical classes of construction documents that routinely contain hundreds of pages. Because construction workflows rarely proceed in a straight line, this extraction often requires repeated review of the same pages across multiple stages of work and across multiple stakeholders.

Each extraction action is often considered to be operationally minor because each individual page that is referenced consumes only a small amount of time. A page is opened, a quantity is checked, a note is confirmed, or a revised condition is reviewed. In isolation, these moments rarely appear economically significant. But across a year of bids, revisions, clarifications, and procurement coordination, these small events accumulate into a meaningful share of skilled labor.

Although individual extraction events may appear straightforward, or even rote, they function as the entry point to higher-complexity work in which technical judgment is applied to the information produced. This means document extraction increasingly behaves not merely as background labor, but as a strategic operating cost at a time when document volume is expanding, revisions move faster, and experienced technical labor remains difficult to expand.

1. Document Extraction Supports Skilled Assessment

In most construction workflows, skilled assessment does not begin at the start of a workflow activity. Before a decision can be made, information must first be located, reviewed, retrieved, and/or verified. A page must be opened, the relevant condition identified, and the applicable detail confirmed. Even when the extracted information is simple, the extraction process consumes attention before interpretation begins. The work may involve counting rooms, identifying fixture locations, verifying whether a relevant condition is shown, or checking whether a prior assumption still aligns with a revised drawing.

Architectural drawings present a distinct extraction challenge, because relevant information is often distributed across symbols, schedules, legends, keyed notes, and sheet-to-sheet relationships that vary from project to project. Unlike specification sections, which are largely text-based and often follow a relatively predictable structure, drawing interpretation requires locating meaning inside project-specific visual cues that vary considerably across design teams and document sets.

Extracting data from drawings in Division 26 may involve an electrical estimator initially reviewing a reflected ceiling plan to count lighting fixtures, later reopening the same sheet to verify fixture types after an addendum, and then revisiting the electrical schedule to confirm whether circuiting assumptions remain valid. In the case of Division 08, a door estimator may first open an architectural floor plan to identify openings, later revisit the same sheet to confirm door tags, then cross-reference specification sections to determine hardware sets and fire-rating requirements. These extraction actions often include light interpretation, but their primary role is to enable subsequent tasks requiring expert attention, such as cost estimates or go/no-go bid decisions.

2. The Accumulation of Extraction Tasks Within Construction Workflows

A drawing page rarely contributes value only once. In Division 26, a reflected ceiling plan may first be reviewed by an estimator to count fixtures and identify fixture types. After award, an assistant project manager may reopen the same page to verify fixture package assumptions against procurement submittals. Later, a project manager may revisit the same sheet during a change-order discussion to determine whether fixture relocations affect branch wiring, controls, or labor scope.

In Division 08, a floor plan may first be reviewed by an estimator to count openings and assign preliminary door types. After award, an assistant project manager may reopen the page to coordinate hardware assumptions with suppliers. A project manager may later revisit the same sheet during a change-order review when opening sizes, frame conditions, or sequence changes affect scope. A detailer may then use the same opening information again to validate field conditions and confirm that the original estimate aligns with fabrication requirements.

These examples illustrate how a single page can be referenced multiple times. Repeated extraction occurs not only because the same information remains useful, but because different stakeholders return to the same information for different downstream decisions. Furthermore, construction workflows naturally generate new moments in which previously viewed information must be recovered. A page that initially supports a bid may later support procurement, submittal review, field coordination, or a change-order position.

Extraction is not the same as skilled assessment, but it affects how much attention remains available for interpreting and acting on the extracted data. When repeated retrieval expands, less time remains for pricing strategy, exclusion analysis, vendor coordination, field coordination, and scope-risk evaluation.

3. Quantifying the Economic Burden of Document Extraction

While no single firm follows identical document patterns, this section outlines an estimate of the economic cost resulting from document extraction at the firm level. We will calculate the total number of unique pages from which data must be extracted and then estimate a multiple of these pages (to represent the repeated extractions applied to each page). The resulting total number of page-reference events will be used to assess the economic impact. Our calculations will be rooted in architectural drawing pages, since such extractions are particularly tedious. The goal is not to claim a universal benchmark, but to make the order of magnitude visible.

We start with a moderate commercial subcontracting firm under the assumption that they evaluate 100 bid opportunities per year. Bid opportunities generally involve both a drawing set and a project specification set. With our focus on pages in the drawing set, an estimate of 40 relevant drawing pages per opportunity reflects the fact that many project packages contain hundreds of pages, but only a subset are directly relevant to a specific trade. This number may be lower in narrowly scoped projects and higher for more complex projects (e.g., hospitals) where drawings are denser, more fragmented, and more heavily cross-referenced.

Every time a drawing page is referenced for data extraction, we estimate an average extraction duration of 7 minutes. This assumption for each page-reference event reflects typical manual extraction activities such as counting rooms, locating fixture types, confirming opening tags, measuring distances, or verifying whether a revision changes a previously noted condition. A simple room count in a small office building may take only a few minutes. A room count in a hospital, however, may take materially longer because

departments, room types, alternate layouts, medical equipment, and repeated plan areas can make boundaries and counting rules harder to interpret.

Similarly, measuring and verifying distances often introduces variability because scale interpretation, missing dimensions, overlapping callouts, rotated plans, or sheet-to-sheet continuation may require repeated checking. These examples are why the model uses an average rather than treating every page-reference event as identical.

For each page, we assume a multiple of 3 to arrive at the total page-reference count. A page may be referenced once during estimating, again during clarification or addenda review, and again during procurement, project management, or change-order evaluation. In some workflows, particularly those involving detailers, supplier coordination, or multiple post-award revisions, the same page may be referenced more often. This multiple varies considerably across the construction industry. For a door estimator on a hospital project, 3 might be low. For a small office tenant improvement, it might be high. Nevertheless, our experience suggests that a multiple of 3 is reasonable as a central estimate. Using these assumptions, the annual burden for a moderate subcontractor is calculated below.

Exhibit 1: Moderate Commercial Subcontractor Volumes

Input / Output	Illustrative Assumption
Annual bid opportunities	100
Relevant drawing pages per opportunity	40
Page-reference multiple	3
Total page-reference events per year	12,000
Minutes per page-reference event	7
Estimated annual labor hours	**~1,400**
FTE equivalent (2,000 hours/year)	**~0.7 FTE**

The calculation is intentionally conservative in two respects. First, it focuses primarily on relevant drawing pages, even though similar extraction patterns occur across specifications, schedules, addenda, submittals, and downstream change documentation. Second, it treats the annual burden as distributed across technical staff rather than as one dedicated role. In practice, the time may be spread across estimators, assistant project managers, project managers, coordinators, suppliers, and detailers. The same extraction dynamics apply across trades and firm sizes, so the aggregate burden is substantial even under conservative assumptions.

4. Why This Becomes an Operating Constraint

Document extraction becomes a strategic operating cost when it begins to limit what a business can do with the technical staff it already has. The cost is not only the wage cost of time spent in drawings and specifications. It is fewer bids reviewed deeply, less time for scope-risk analysis, slower responses to revisions, and more pressure on the same estimators, project managers, coordinators, and detailers.

The true cost is not labor expense alone. It is what cannot happen while skilled staff have to spend significant time to locate, verify, and recover information needed for downstream decisions. Even the first extraction from a page can require time and attention: opening the correct drawing, finding the relevant detail or note, confirming the applicable quantity, checking a schedule relationship, or determining

whether a revision affects a prior assumption. When similar page-reference events accumulate across bids, revisions, procurement, submittals, and change-order review, that burden begins to affect higher-value assessment.

The burden compounds when teams must repeatedly recover information already seen before. When a team spends time reopening drawings, finding the same detail or note, verifying the same quantity, or checking the same schedule relationship, that time is not available for higher-value assessment.

This burden is also cognitive. Daniel Kahneman's work on attention and effort describes cognitive capacity as a limited resource: tasks requiring concentration draw from a finite pool of available attention, and competing demands reduce what remains for other forms of judgment. In construction workflows, reopening a drawing, reacquiring context, checking a note, and comparing a revision each consume part of that limited capacity before higher-level assessment begins. The cost is therefore not only the minutes spent on extraction, but the recurring use of scarce cognitive bandwidth that could otherwise support higher-value decisions.[8,9]

> *The true cost is not labor expense alone. It is what cannot happen while skilled staff have to spend significant time to locate, verify, and recover information needed for downstream decisions.*

To be clear, not all manual review should be reduced. Some document review is where expert judgment enters the workflow: testing an assumption, reconciling a scope gap, challenging a quantity, or interpreting an ambiguous detail can improve bid quality and reduce risk. The burden identified here is narrower. It is the time and attention consumed by locating, reacquiring, checking, and validating information as a precondition to that judgment. Even when a given stakeholder extracts a piece of information only once, the process still consumes cognitive capacity. When similar page-reference events are aggregated across projects, revisions, workflows, and stakeholders, they become a strategic operating cost.

This matters because many construction businesses operate with limited technical capacity. Estimators, assistant project managers, project managers, coordinators, and detailers often carry overlapping responsibilities across bids, procurement, submittal review, field coordination, and change management. A few minutes of extraction may appear harmless in one moment, but repeated across workflows, those minutes draw from the same pool of skilled attention needed for pricing strategy, scope-risk analysis, supplier negotiation, and project execution decisions. The burden may not appear as one visible failure, but its effects can surface indirectly: slower turnaround, narrower review coverage, more late-stage compression, and reduced time for strategic assessment.

The burden is also difficult to see because it is fragmented. When an estimator reopens drawings late in a bid cycle, an assistant project manager rechecks the same information during procurement, and a project manager later revisits that information during a change-order discussion, the cumulative labor is rarely recognized as one connected cost.

When higher-value assessment is delayed, compressed, or selectively skipped, repeated extraction begins influencing bid quality, responsiveness, and technical capacity. At that point, document extraction is no longer just a background task. It becomes an operating variable that affects how much skilled work the business can absorb without proportional headcount growth.

5. Why This Matters More Now

The economic burden of repeated document extraction is becoming more important because the environment around construction information has changed. Firms are being asked to process more information, respond to more revisions, and make decisions faster, while the supply of experienced technical labor remains difficult to expand.

This labor constraint matters because many of the roles affected by repeated document extraction are not easily replaced or scaled. Industry workforce surveys continue to show persistent difficulty filling construction roles, including salaried roles that support estimating, project management, supervision, and coordination. Estimators, project managers, coordinators, and detailers often carry overlapping responsibilities across bids, procurement, submittal review, field coordination, and change management. When these people spend more time recovering information that has already been reviewed, less of their attention remains available for the judgment-based work that depends on their experience.[1,2]

At the same time, construction information has become easier to distribute but not necessarily easier to interpret. Drawing sets, specifications, addenda, submittals, schedules, BIM-related outputs, and downstream change documentation now move quickly through digital workflows. Yet the underlying work of locating the relevant information, understanding its trade-specific meaning, checking it against related documents, and determining whether it has changed remains a largely human burden in many firms.[3]

This is why the problem is not solved by OCR alone. Construction drawings and specifications are not simply collections of readable text. They contain trade-specific meaning distributed across symbols, schedules, legends, keyed notes, details, sheet references, specification sections, addenda, and revision histories. In estimating workflows especially, the challenge is not merely to read information on a page, but to preserve the context, source location, and relationships that allow extracted information to support pricing, scope review, quantity takeoff, and downstream project decisions.[5–7]

As document volume grows and revisions move faster, repeated retrieval consumes a larger share of the same limited attention that firms rely upon for complex assessment. The result is not simply a software problem. It is a capacity problem: how much skilled assessment remains available after repetitive extraction work has already consumed part of the day.

Conclusion

Construction businesses face many sources of pressure, including labor shortages, compressed schedules, procurement volatility, and evolving project complexity. Within that broader environment, repeated extraction events represent an often-overlooked operating burden because they repeatedly occupy the same skilled attention needed for higher-value assessment.

A page is reopened, a quantity is verified, a note is checked, a revision is compared. Each event appears minor. Across bids, revisions, clarifications, procurement coordination, submittal review, and change-order analysis, however, these events accumulate into a meaningful share of skilled labor.

The strategic importance of document extraction therefore lies not only in direct labor cost, but in how it governs where technical attention remains available inside the business. A firm may not experience extraction burden as a single crisis, but it may experience the downstream consequences through slower review cycles, reduced technical coverage, and less time available for decisions that affect cost, risk, and scope.

Manual review remains essential in many cases, especially where field conditions, ambiguous drawings, trade coordination, or contractual interpretation require expert assessment. But repetitive extraction increasingly deserves to be minimized where it does not materially improve high-level decisions. The opportunity is not to remove skilled judgment from construction workflows, but to preserve it by reducing the repeated extraction work that prevents that judgment from being applied where it matters most.

AnchorGrid addresses this problem by turning construction drawings and specifications into structured, reusable project information, helping teams reduce the repeated extraction work that consumes skilled attention before higher-value judgment can be applied.

References

1. Associated General Contractors of America and Arcoro. “2024 Workforce Survey Analysis.” Associated General Contractors of America, 2024.
2. Associated General Contractors of America and Sage. “The 2024 Construction Hiring and Business Outlook.” Associated General Contractors of America, 2024.
3. National Institute of Building Sciences Digital Twin Integration Subcommittee. “Digital Twins for the Built Environment.” National Institute of Building Sciences, 2024.
4. U.S. Census Bureau. “2023 County Business Patterns: NAICS 238 — Specialty Trade Contractors.” U.S. Census Bureau, 2025.
5. Billd. *The Construction Estimating Process: A Breakdown for Subcontractors.* Chris Doyle, updated 2026.
6. Procore Technologies. *Construction Cost Estimating: A Step-By-Step Guide.* Procore, 2024.
7. Whole Building Design Guide. *Cost Estimating (Level 4 and Level 5 Construction Documents).* National Institute of Building Sciences.
8. Kahneman, Daniel. *Attention and Effort.* Englewood Cliffs, NJ: Prentice-Hall, 1973.
9. Brann, Amy. “Stop Overloading the Wrong Part of Your Brain at Work.” *Harvard Business Review*, October 16, 2025.

About AnchorGrid

AnchorGrid builds infrastructure and workflow tools that help commercial construction teams turn drawings and specifications into structured, reusable project information. Its technology uses trained detection models to locate and extract element-level data from architectural drawings, including doors, fixtures, room counts, and other trade-specific elements, and returns structured results that can support estimating, procurement, project management, and software development workflows. The goal is not to replace expert judgment, but to preserve more of it for the decisions where experience matters most.