



WHITE PAPER

How to Choose an IDP Solution for AP Automation

A practical decision framework for finance leaders evaluating intelligent document processing.

PREPARED BY
DocXtract

PUBLISHED
2026

AUDIENCE
CFOs, Controllers & AP Leads

— EXECUTIVE SUMMARY

The decision is not whether to automate. It is what to automate with.

Most finance leaders have already accepted that manual accounts payable processing is a cost they no longer want to carry. The harder question is which intelligent document processing (IDP) solution to build their AP automation on, because that choice determines whether the project delivers a step change or becomes another half-automated workflow that still needs people to fill the gaps.

This white paper gives finance leaders a structured way to make that decision. It moves past vendor marketing and feature checklists to the seven dimensions that actually predict whether an IDP solution will work on your invoices, integrate with your systems, and hold up at your volume. For each dimension it provides the specific questions to ask, the answers that signal real capability, and the answers that signal future rework.

THE CORE ARGUMENT

The accuracy number on a vendor's homepage tells you almost nothing. What matters is accuracy on your document mix, how the solution handles the invoices it has never seen, who owns the integration work, and what happens at the exceptions. Evaluate those, and the right choice becomes obvious.

CONTENTS

01	Why the IDP choice decides the AP outcome	03
02	What AP automation actually requires	03
03	The seven-dimension evaluation framework	04
04	Template OCR vs. intelligent extraction	06
05	The finance leader's evaluation checklist	07

Why the IDP choice decides the AP outcome

Accounts payable is a document problem before it is a payment problem. Every invoice that enters the process has to be read, understood, validated, and turned into structured data before any workflow, approval, or ERP posting can happen. The intelligent document processing engine is the layer that does this. It is the foundation the entire automation sits on.

This is why the IDP choice matters more than the workflow tooling around it. A sophisticated approval workflow built on a weak extraction engine simply routes bad data faster. If the engine cannot read a third of your invoice formats, no amount of workflow design closes that gap. People close it, by hand, and the automation quietly becomes assisted manual processing.

\$10–15

Cost to process one invoice manually

~14.6 days

Average manual invoice cycle time

~2%

Error rate in manual data entry

Industry benchmarks: Levvel Research and Ardent Partners, cited in 2025 AP performance reviews. Automation typically reduces per-invoice cost to \$2–3 and cycle time to 3–5 days, but only when the underlying extraction is reliable enough to run straight through.

What AP automation actually requires

Before evaluating any vendor, it helps to be precise about what end-to-end AP automation demands from the document layer. Four things are non-negotiable:

1

Read any invoice, from any vendor, at any quality.

Your vendors do not coordinate their invoice formats. The engine must handle the full range, including scans, photos, and regional formats, without a template per layout.

2

Extract the complete data set, not just the header.

Header fields, line items, tax breakdowns, and totals all need to be captured accurately and structured for the ERP.

3

Validate before posting.

Field formats, tax IDs, duplicate checks, and vendor master cross-references should be checked at the point of extraction, not discovered downstream.

4

Integrate with the systems you already run.

Structured data has to flow into your ERP and accounting stack through real connectors or an API, not a manual export.

WHY THIS MATTERS FOR EVALUATION

Every dimension in the framework that follows maps back to one of these four requirements. If a solution is strong on workflow but weak on reading your actual invoice mix, it fails the first and most important test.

The seven-dimension evaluation framework

Score each candidate solution across these seven dimensions. Together they predict whether an IDP solution will deliver real AP automation or partial automation that still leans on your team.

1 Accuracy on your document mix

A homepage accuracy number is measured on the vendor's test set, not your invoices.

The only accuracy figure that matters is the one measured on a representative sample of your own invoices, including your messiest vendors and lowest-quality scans. Ask for a test on your documents before you commit to anything.

ASK THE VENDOR

- What is field-level accuracy on a sample of our real invoices, not a demo set?
- How is accuracy measured: per document, per field, or per character?
- Can we run a test on our own documents before buying?

2 Handling of unseen formats

New vendors and changed layouts are constant. What happens when one arrives?

This is the single biggest predictor of long-term success. Template-based systems break or need reconfiguration on every unseen layout. Intelligent extraction reads new formats on arrival because it understands document structure rather than matching fixed positions.

ASK THE VENDOR

- When a vendor we have never seen sends an invoice, does it process automatically or need setup?
- Who builds and maintains templates, and how long does a new format take?
- What is the ongoing maintenance burden as our vendor base grows?

3 Depth of extraction

Reading the invoice number is easy. Line items and tax are where tools fall short.

Confirm the engine captures the full structure your ERP needs: line-item tables, multi-line descriptions, tax breakdowns, and totals, with the validation that keeps bad data out of your books.

ASK THE VENDOR

- Does it extract complete line-item tables, or only header fields?
- Does it handle tax fields and region-specific requirements such as GST and HSN?
- What validation runs before data reaches the ERP?

4 Integration with your stack

Extraction that does not reach your ERP cleanly is only half a solution.

Clarify exactly how validated data lands in your ERP and accounting systems, whether through pre-built connectors or an API, and crucially, who does the integration work.

ASK THE VENDOR

- Do you have pre-built connectors for our ERP and accounting systems?
- Is there a documented REST API, and who builds the integration: us or you?
- What does a typical integration timeline look like?

5 Exception handling and straight-through rate

The value is in what flows through untouched. The cost is in what does not.

Ask for the realistic straight-through processing rate on a mix like yours, and understand exactly what happens to the exceptions: how they surface, how a human resolves them, and whether the system learns.

ASK THE VENDOR

- What straight-through rate is realistic on our invoice mix?
- How are low-confidence fields flagged and routed for review?
- Does the system improve as it processes more of our documents?

6 Time to value and deployment

A solution that takes months to configure delays every benefit you are buying.

Deployment speed separates solutions that need extensive training and template-building from those that work on day one. Pin down the realistic timeline from contract to processing live invoices.

ASK THE VENDOR

- How long from signed contract to processing live invoices?
- Is upfront template-building or model training required before go-live?
- What does your team do versus what ours has to do?

7 Scale, security and total cost

The price per page is only part of the real cost of ownership.

Factor in volume handling, data security and compliance posture, and the full cost including integration, maintenance, and the human effort the solution still requires.

Template OCR vs. intelligent extraction

Most of the seven dimensions resolve to one underlying architectural difference. Traditional OCR matches fields to fixed positions on a known template. Intelligent extraction reads a document by understanding its structure and context, the way a person does. This difference determines how a solution behaves on every dimension that matters at scale.

Dimension	Template-based OCR	Intelligent extraction
Unseen invoice format	Breaks or needs a new template	Reads it on arrival
New vendor onboarding	Template build per layout	No setup required
Maintenance burden	Grows with vendor count	Stays flat
Accuracy on messy scans	Degrades sharply	Holds up
Line-item tables	Often partial	Full structure
Time to deploy	Weeks of configuration	Days
Cost trajectory	Rises with variety	Predictable

THE DECISION RULE

If your vendor base is small, stable, and uniform, template OCR can work. For everyone else, the long tail of formats is exactly where template systems fail and where the manual workarounds quietly return. Intelligent extraction is what keeps automation actually automated as you scale.

What "good" looks like

A strong IDP solution for AP automation reads any invoice without templates, captures complete structured data including line items and tax, validates before posting, integrates through real connectors and an API, surfaces only genuine exceptions for human review, deploys in days rather than months, and holds its cost trajectory as your volume and vendor variety grow. Score your candidates against that standard.

The finance leader's evaluation checklist

Use this in vendor conversations. A solution that earns a confident yes on most of these is built for real AP automation. **Hesitation on the first three is the clearest warning sign.**

- 1 Accuracy proof.** They will test on our own invoices and report field-level accuracy before we commit.
- 2 Unseen formats.** A new vendor's invoice processes automatically, with no template build.
- 3 Extraction depth.** Full line items, tax fields and totals are captured and validated, not just header data.
- 4 Integration ownership.** Pre-built connectors or a documented API exist, and the vendor owns the integration work.
- 5 Straight-through rate.** A realistic touchless rate is quoted for a document mix like ours.
- 6 Exception path.** Low-confidence fields are flagged and routed clearly, and the system learns over time.
- 7 Time to value.** We can be live on real invoices in weeks, not months, with minimal work on our side.
- 8 Total cost clarity.** All-in cost is transparent, including integration, maintenance and residual manual effort.

- 9 Scale and security.** Throughput, data security, residency and compliance meet our requirements at peak volume.

SEE IT ON YOUR INVOICES

See how DocXtract scores on your invoices.

DocXtract is an AI-powered intelligent document processing platform built for exactly this evaluation: 98%+ field-level accuracy, no templates, deployment in days, and pre-built connectors for the systems you already run. Test it on your own documents and see the framework in action.

Book a demo →

docxtract.io | connect@docxtract.io

© 2025 DocXtract. This white paper is provided for informational purposes for finance and procurement leaders evaluating intelligent document processing. Benchmark figures are drawn from cited third-party industry research and represent general ranges, not guarantees of specific outcomes.

DocXtract