



Intelligent Document Processing

DocLang Disrupts Intelligent Document Processing: Why Leaders Like ABBYY Are Betting On It

An open, neutrally governed document standard is redrawing the boundaries of the IDP market. This paper sets out QKS Group's assessment of what changes, what does not, and why a category leader chose to help commoditize its own output layer.

Kunal Pakhale
Principal Analyst

TABLE OF CONTENTS

1. Executive Summary	03
2. The Shift in the IDP market	04
3. What changes, and what does not	05
4. Why the category leader leaned in	07
5. Implications for the market	08
6. What could slow this down	09
7. QKS Group’s assessment	10

1. Executive Summary

The Intelligent Document Processing market has spent a decade competing, in part, on proprietary output, each vendor turning documents into its own structured format, and each customer building custom work to consume it. In QKS Group's assessment, the arrival of a shared, openly governed document standard removes that dimension of competition and, in doing so, resets where value accrues in the category.

This is not a marginal product update. It changes the buying conversation, compresses one source of vendor lock-in, and forces every provider to compete on the parts of document intelligence that are genuinely hard. ABBYY, a long-standing category leader, has read this correctly and moved first, sponsoring the standard rather than resisting it. QKS Group views this as a position of strength, not concession, and one that signals where the durable advantage in IDP now sits: in recognition quality, document and language breadth, deployment flexibility, and data governance.

For enterprise buyers, the practical implication is immediate: the output format is no longer a lock-in worth paying a premium to accept. The questions that should drive selection are the ones above the format line.

2. The shift in the IDP market

For most of its history, the Intelligent Document Processing market has sold a single core promise: take unstructured documents- invoices, contracts, claims, statements, forms, and turn them into structured, usable data. Vendors differentiated on how well they read those documents, but also on the format of what they handed back. That output was proprietary, and it quietly created switching costs. Once an enterprise built its downstream systems around one vendor's output, moving to another meant rebuilding the plumbing.

That dynamic is now changing. The emergence of an open, neutrally governed document standard, one that any vendor can produce and any downstream system can consume, removes the format as a point of differentiation and as a source of lock-in. This is the most consequential structural change the IDP category has seen since the shift from template-based capture to machine-learning-based extraction.

The reason it matters is not technical elegance. It is economic. When output is standardized, the cost of integrating any given vendor into an enterprise stack falls sharply, and it falls for every vendor at once. That expands the addressable opportunity for the whole category; more projects clear the business case when integration is cheaper, while simultaneously stripping away one of the ways weaker vendors held onto accounts. Standardization grows the market and raises the competitive bar in the same motion.

Why now

Three pressures have converged to make this shift not just possible but necessary. First, enterprises are trying to feed documents into AI systems and agents at scale, and they are discovering that inconsistent, vendor-specific output is a primary source of friction and cost. Second, buyers have grown openly hostile to lock-in and increasingly refuse to standardize on proprietary formats. Third, the category's leaders have matured to the point where they are confident competing on capability rather than on switching costs. When those three align, an open standard stops being a threat to incumbents and becomes an instrument they can use.

3. What changes, and what does not

It is easy to overstate a shift like this. A standard does not make document intelligence easy; it makes one part of it shared. The discipline, for the analyst and for the buyer alike, is to separate what the standard actually commoditizes from what remains hard and therefore remains a source of advantage.

“The hard part of document intelligence is not the output format. It's reading a messy, real-world document correctly and standardizing that output layer helps the whole category. When every tool speaks the same language, enterprises stop rebuilding parsers and start competing on what actually matters, like recognition quality, the breadth of documents you can handle, and how you govern the data.”

- Max Vermeir, Vice President, AI Strategy, ABBYY

Max Vermeir's framing aligns closely with QKS Group's read of the market, and it is worth taking at face value because it also runs against a naive reading of ABBYY's self-interest. The output format was never where the difficulty lived. The difficulty lives in correctly interpreting a document that was designed for a human, not a machine: a smudged scan, an unusual layout, a table that spans pages, a language the model has barely seen, a form where the meaningful information is in what is not filled in. Standardizing the output does nothing to make that easier. It simply stops every vendor and every customer from re-solving the trivial part.

Table 1 - The output format is commoditized; the capability above and below it is not.

What the standard commoditizes	What remains a source of advantage
The serialization format of extracted output	Recognition and extraction quality on messy, real-world documents
Custom parsing and integration glue between vendor and downstream systems	Breadth of document types, layouts, and languages handled out of the box
The lock-in that came from proprietary output	Deployment flexibility - on-premise, containerized, cloud, at enterprise scale

What the standard commoditizes

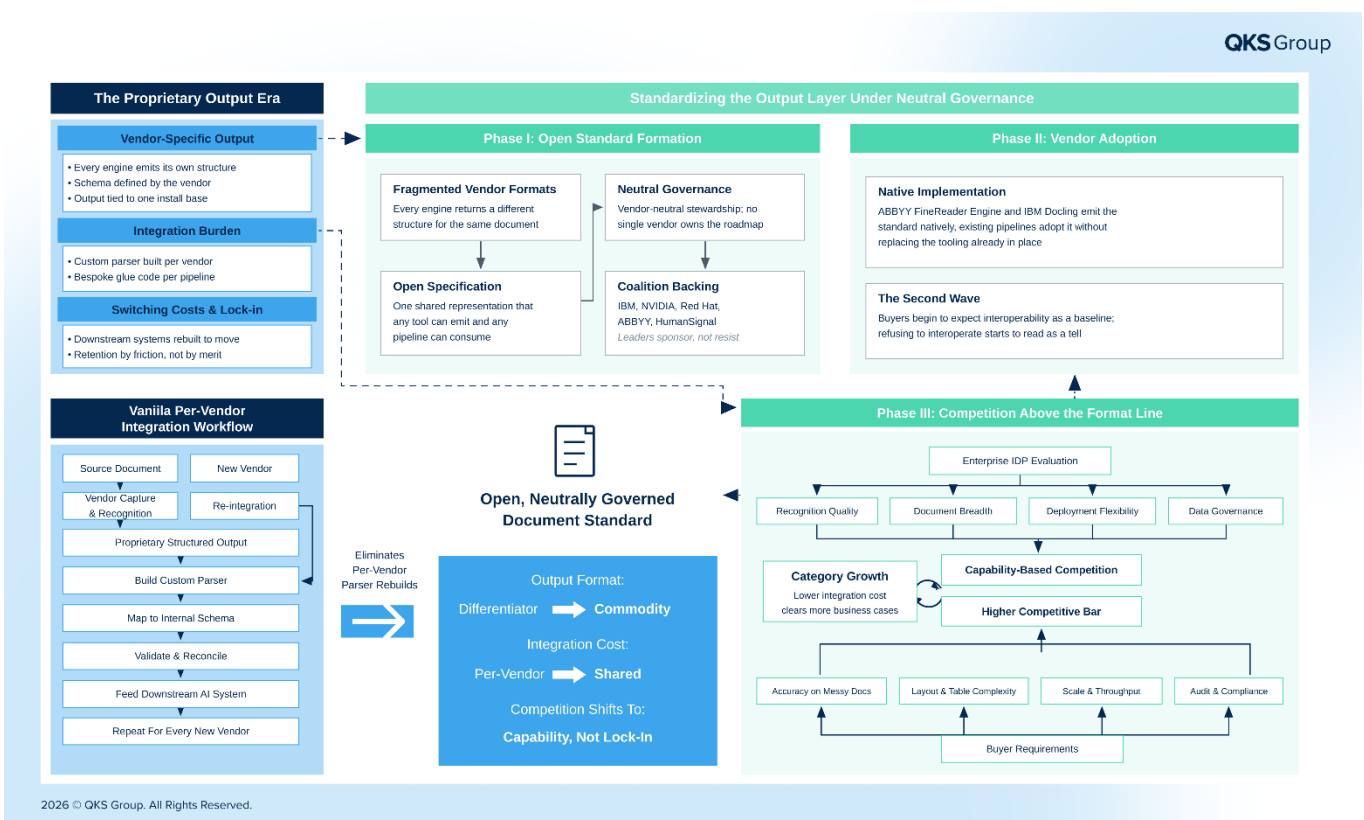
One axis of vendor-versus-vendor differentiation

What remains a source of advantage

Data governance, auditability, and compliance handling

This distinction is the heart of the assessment. A vendor whose advantage rested on proprietary output has genuinely lost something. A vendor whose advantage rests on the right-hand column has lost nothing, and has gained a larger, more accessible market to sell into. That difference determines who benefits from the disruption and who is exposed by it.

Figure 1 sets out the full picture: the per-vendor integration model the market is leaving behind, the phased emergence of an openly governed alternative, and the competitive ground on which the category now competes.



Standardizing the output layer moves competition above the format line

4. Why the category leader leaned in

The instinctive reading of an incumbent sponsoring a standard that commoditizes part of its own product is that the incumbent is on the defensive. QKS Group reads it the opposite way. ABBYY is not a challenger hoping a standard will level the field. It is an established leader, repeatedly recognized at the top of independent IDP evaluations, with a large enterprise install base and decades of investment in the hard, upstream capability that the standard does not touch. For a company in that position, sponsoring the standard is a rational offensive move.

“A proprietary format would have tied this market to a vendor's install base. A neutrally governed one grows the entire category, and ABBYY grows with it. Sponsoring DocLang is a bet that we win by making the format free and competing on everything above it.”

- Max Vermeir, Vice President, AI Strategy, ABBYY

This is the clearest articulation of the strategic logic, and it is worth unpacking because it generalizes beyond ABBYY. The bet has three parts, and each holds up under scrutiny.

- **The format was never the moat.** A leader whose defensible value is recognition quality, breadth, and deployment can give away the output format without exposing anything that actually protects its position. The moat is upstream and downstream of the format line, not on it.
- **Neutral governance grows the pie.** A proprietary format would have chained the category's growth to any single vendor's reach. An openly governed one lowers integration cost across the whole ecosystem, and the vendor with the strongest underlying capability captures a disproportionate share of a larger market.
- **It neutralizes the buyer's strongest objection.** Lock-in is the objection that stalls competitive deals. By backing a neutral standard, a leader removes that objection from the table and forces the comparison onto the ground where it is strongest: capability.

Read together, the two quotes describe a coherent posture: give away the commodity, compete on the craft. It is the behavior of a vendor that expects to win in a standardized world, not one hiding from it. In QKS Group's view, that confidence is itself a signal buyers should weigh, a leader willing to compete without the crutch of lock-in is usually a leader confident in what remains.

5. Implications for the market

For enterprise buyers

The single most important consequence is that the output format should drop out of the selection criteria. It is no longer a reason to accept a premium or tolerate a weaker product. The questions that should now drive an IDP decision are the ones the standard cannot answer for the buyer: How accurately does the vendor read the documents the enterprise actually has, in the condition they are actually in? How broad is its coverage of the relevant document types and languages without custom work? How does it deploy into the target environment and scale to real volumes? How does it handle the governance, auditability, and compliance obligations the industry imposes? A standardized output layer makes it easier, not harder, to run a genuine bake-off on these questions, because a proof of concept no longer requires bespoke integration to evaluate each vendor.

For IDP vendors

The competitive pressure rises. Vendors who leaned on proprietary output to retain accounts will feel the floor move; those who invested in the hard capability will find a larger, more accessible market. QKS Group expects a second wave of vendors to adopt the standard once its consumption side gains momentum, not out of enthusiasm, but because buyers will begin to expect it, and because refusing to interoperate will read as a tell. The vendors who move early and frame it as confidence, as ABBYY has, will control the narrative; the ones who resist will look like they are protecting a moat that no longer holds.

For the analyst scorecard

For firms that evaluate this market, QKS Group among them, the shift has a practical effect on how vendors should be weighed. Output format and integration convenience decline as differentiators. Recognition quality, breadth, deployment, and governance rise in relative weight. Evaluation frameworks that still reward proprietary output or penalize interoperability will need to be recalibrated to reflect where value has actually moved.

6. What could slow this down

An honest assessment has to account for the ways this shift could stall, because standards are not self-executing and adoption is never guaranteed.

- **Adoption is a two-sided problem.** A shared format is only valuable if both the vendors producing output and the systems consuming it commit to it. Early sponsorship by a leader helps, but the tipping point comes when the systems that consume documents, the AI platforms and agents downstream, begin to expect the standard by default. Until then, momentum is a question rather than a fact.
- **Neutral governance must be demonstrated, not asserted.** The credibility of an open standard rests on its stewardship visibly serving interoperability over any single sponsor's product priorities. The founding participants are also commercial competitors. The early signals are encouraging, but sustained, visibly neutral governance is the thing to watch over the coming quarters.
- **The claims deserve independent validation.** The performance and efficiency benefits attributed to the standard are, so far, largely vendor-reported. They are plausible and directionally consistent with what QKS Group would expect, but independent replication across different systems and document sets is what turns a persuasive claim into an established fact. A genuinely open standard should make that validation easier, and buyers should ask for it.
- **It is early.** This is a young standard with a small number of early adopters. That is normal, and the openness is a strength, but any recommendation today is a recommendation about a credible direction of travel and a serious founding coalition, not about a settled, battle-tested standard. That distinction belongs in any honest buyer conversation.

7. QKS Group's assessment

In QKS Group's view, the open standardization of the document output layer is the most significant structural change in Intelligent Document Processing in years, and it is a healthy one. It removes a source of friction and lock-in that never served buyers, it grows the addressable market, and it forces the category to compete on the capabilities that actually determine whether a document-driven AI project succeeds.

ABBYY's decision to sponsor rather than resist is the clearest leading indicator of where value is moving. When the category's leader voluntarily gives away the output format and stakes its position on everything above it, it is telling the market plainly that the format was never the point. QKS Group reads that as confidence, and as well founded. The vendors who thrive in the next phase of this market will be the ones who, like ABBYY, treat the standard as a floor to build on rather than a moat to defend.

For enterprise leaders, Kunal Pakhale's (Principal Analyst at QKS Group) recommendation is straightforward: "Treat interoperable output as a baseline expectation, not a premium feature. Redirect the scrutiny once applied to formats and integration onto recognition quality, breadth, deployment, and governance. And evaluate vendors on their willingness to compete in the open, because in a standardized market, that willingness is itself a measure of how much they trust their own capability."

About this paper

This white paper reflects the independent assessment of QKS Group, a vendor-neutral technology research and advisory firm, on the Intelligent Document Processing market. Direct quotations from Max Vermeir, Vice President of AI Strategy at ABBYY, are included with attribution.