Mistral OCR 4.1 vs Cohere Parse: Extraction Accuracy Compared

IntuitionLabs.ai · Adrien Laurent · 2026-09-05 · mistral ocr 4.1 · cohere parse · ocr benchmark · document extraction · document ai · pdf parsing · olmocr-bench · omnidocbench

Executive Summary

Mistral OCR 4.1 and **Cohere Parse** are two document-extraction models that reached the market within six weeks of each other in the second half of 2026: Mistral's model went to preview on July 16, 2026 and Generally Available (GA) on August 31, 2026 (docs.mistral.ai), while Cohere announced Parse on August 27, 2026, as detailed below. Mistral's OCR 4 announcement lists a Batch API rate of $2 per 1,000 pages (Mistral AI). Priced at $4 per 1,000 pages through the standard API, Mistral OCR 4.1 documents a wide structural taxonomy (13 labeled block types, including tables, equations, captions, and references) and confidence scores at page, block, and word granularity (docs.mistral.ai). Cohere Parse, at $1.50 per 1,000 pages, is a smaller 2.3-billion-parameter model built for throughput (up to 2,160 pages per minute on an 8-GPU node) but ships without confidence scores or structured JSON output (docs.cohere.com).

The most consequential finding of this report is not a winner but a gap: three independent, publicly reproducible benchmarks, Ai2's olmOCR-Bench, the CVPR-2025 OmniDocBench, and the LlamaIndex-built ParseBench, each score a Mistral OCR entry (labeled variously "Mistral OCR API," unversioned "Mistral OCR," or "Mistral OCR 4"), yet none of the three carries a Cohere Parse entry as of this report's observation date of September 5, 2026 (raw.githubusercontent.com) (github.com).

Developer reception diverges sharply too: Mistral OCR 4.1's Hacker News thread drew 411 points and 166 comments (news.ycombinator.com), while Cohere Parse's own submission drew 4 points and no comments (news.ycombinator.com). Pricing context across the wider category shows Google Document AI, AWS Textract, and Azure AI Document Intelligence all clustering near $1.50 per 1,000 pages for basic OCR, with layout- and form-extraction tiers running $10 to $30 per 1,000 pages (cloud.google.com) (aws.amazon.com), and the global intelligent document processing market is valued at $3.0 billion in 2025, projected to reach $29.7 billion by 2033 (grandviewresearch.com).

For buyers, especially in regulated or scientific-document contexts where table and citation fidelity carry compliance weight, the practical recommendation is to evaluate each model against the organization's own document mix.

Introduction and Background

Optical character recognition (OCR) has moved far beyond converting scanned pixels into plain text. Modern document extraction systems, often built on vision-language models (VLMs, AI models that jointly process images and text), are now expected to preserve reading order, reconstruct table structure, transcribe mathematical notation, and separate a document's citations, headers, and captions into distinct, machine-readable blocks. For life-sciences and enterprise AI teams, this distinction matters more than it once did: a clinical study report, a regulatory submission, or a scientific PDF is only as useful to a downstream large language model (LLM) or retrieval-augmented generation (RAG) pipeline as the structural fidelity of the extraction step that feeds it.

This report compares two document-extraction models that both reached general availability or public preview in the second half of 2026: **Mistral OCR 4.1**, released in preview by Mistral AI on July 16, 2026 and marked Generally Available (GA) on August 31, 2026 (docs.mistral.ai), and **Cohere Parse**, announced by Cohere on August 27, 2026 as "Introducing Parse: Enterprise document intelligence at scale" (cohere.com). Both are commercial, API-delivered models aimed at converting PDFs, scans, and images into structured Markdown, JSON, or HTML, but they differ sharply in pricing, structural metadata, throughput, and, critically, in whether any independent, third-party benchmark has scored them head to head.

IntuitionLabs, a life-sciences and AI consultancy, has previously published a broader survey of OCR and document AI tools, including Tesseract, TrOCR, Donut, LayoutLM, and **Mistral OCR 4** (intuitionlabs.ai); that report predates both Mistral OCR 4.1 and Cohere Parse and is not restated here. This article instead focuses narrowly on these two newer, competing models, on the reproducible public benchmarks that test reading order, tables, equations, and citations, and on the pricing and deployment facts a technical buyer needs as of **September 5, 2026**. As IntuitionLabs frames it in its own work on enterprise AI foundations, "a governed information layer connects assistants and agents to authoritative sources through identity, permissions, retrieval, citations, logging, and accountable operations" (intuitionlabs.ai); document extraction quality is the first link in that chain, since a mis-parsed table or a dropped citation propagates errors into everything built on top of it.

Mistral OCR 4.1

Capabilities

Mistral OCR 4.1 (model ID `mistral-ocr-4-1`) is the fifth named release in Mistral's OCR line, following **OCR** (`mistral-ocr-2503`, March 6, 2025), **OCR 2** (`mistral-ocr-2505`, May 22, 2025), **OCR 3** (`mistral-ocr-2512`, December 18, 2025), and **OCR 4** (`mistral-ocr-4-0`, June 23, 2026) (docs.mistral.ai). Mistral's official model page describes it as powering "our latest OCR service powering our Document AI stack, with native paragraph-level bounding box extraction, structural block labels, and block-level confidence scores" (docs.mistral.ai). The July 16, 2026 release note confirms this was the version's initial launch and that the `mistral-ocr-latest` and `mistral-ocr-4` aliases were repointed to it (docs.mistral.ai).

Structurally, Mistral's Document AI OCR guide states that block extraction (`include_blocks`) returns a `blocks` array with "paragraph-level bounding boxes, structural block labels" and content in reading order, available only for OCR 4 or newer (docs.mistral.ai). The documentation lists 13 distinct block types the model can emit, including text, title, list, table, image, equation, caption, code, references, aside_text, header, footer, and signature (docs.mistral.ai). Confidence scoring is configurable at three granularities via `confidence_scores_granularity`: page-level aggregate statistics, block-level scores (the headline OCR 4.1 addition), and word-level scores, where "word" returns "everything from 'page', plus a word_confidence_scores array with per-word confidence values" (docs.mistral.ai). Mistral's own documentation states OCR "performs strongly across more than 40 languages" (docs.mistral.ai), a more conservative figure than the "170 languages" some trade press reports repeat without a primary source. Inputs accepted include images (PNG, JPEG/JPG, AVIF) and documents (PDF, PPTX, DOCX), with JSON output carrying per-page Markdown, images, tables, hyperlinks, header/footer text, page dimensions, confidence scores, and blocks (docs.mistral.ai).

As displayed on Mistral's official model page (accessed 2026-09-05), OCR 4.1 is priced at **$4 per 1,000 pages** and **$5 per 1,000 annotated pages** (docs.mistral.ai). Mistral's OCR 4 announcement also lists a **50% Batch API discount**, reducing the standard $4 rate to **$2 per 1,000 pages** (Mistral AI). Independent trade coverage from the model's July 2026 launch quotes the same rate in euros, EUR3.50 per 1,000 pages and EUR4.38 per 1,000 annotated pages (runtimewire.com) (dev.to), figures that are consistent with the dollar price at prevailing exchange rates rather than a contradiction. An independent reviewer notes this pricing did not change again at the 4.1 release, stating that "Mistral's OCR API pricing doubled on June 23" (the OCR 4 launch) (laura-martel.com).

Adoption

Mistral OCR 4.1's Hacker News submission generated substantial developer discussion, reaching 411 points and 166 comments as measured at the time of this research (news.ycombinator.com), up from an earlier snapshot of 370 points and 149 comments reported a day after launch (byteiota.com). Within that thread, one commenter doing detailed transcription of ligatures, critical sigla, Fraktur letterforms, and sub/superscripts wrote there was "nothing special about this model for overly-detailed work like mine," and observed that OpenAI's higher-tier models still led on that narrow task (news.ycombinator.com). A separate commenter reported that Mistral OCR handled "a dense, handwritten tabular form" in a non-English language, producing a "reasonable Markdown format with accurate content" in a case where vision-LLMs and AWS Textract had struggled (news.ycombinator.com). Another commenter compared cost directly, stating that "even comparing to AWS Textract or Azure Document Intelligence, this is very expensive (more than double)" (news.ycombinator.com). On GitHub, Mistral's own reference-notebook

repository, `mistralai/cookbook`, has accumulated approximately 2,300 stars (github.com), and community projects such as `franckalbinet/mistocr`, a batch OCR pipeline built on top of the model, have also appeared (github.com).

Strengths and Limitations

OCR 4.1's documented fixes over OCR 4.0 include bounding boxes that now "align to the content they mark" without drift, reference and citation lists that retain one box per entry instead of collapsing, fewer missed blocks on visually dense pages, and corrected right-to-left table parsing (laura-martel.com).

Cohere Parse

Capabilities

Cohere Parse converts "unstructured enterprise documents (PDFs, images, slides) into structured Markdown output," extracting text, tables, lists, forms, images, captions, and bounding boxes (docs.cohere.com). The current model, `parse-v5.0`, is a 2.3-billion-parameter proprietary vision-language model of roughly 4.6GB, built on an architecture Cohere labels "north-micro-vision-instruct" (docs.cohere.com); trade press further describes it as pairing that model with a custom 400-million-parameter vision encoder initialized from SigLIP 2 SO400M (infoq.com). Cohere's announcement blog, published August 27, 2026, positions Parse as a high-throughput vision parsing model with a strong price-performance profile, discussed further below. In addition to Markdown, Parse supports a "blocks" output format with type-specific fields, including bounding boxes for tables (docs.cohere.com), and Microsoft's Azure AI Foundry catalog listing describes it as automatically detecting "complicated cases like rotated tables, merged cells" and outputting structured HTML for tables (ai.azure.com). Parse is multilingual, "trained on nine of the world's most prevalent commercial languages" (cohere.com), and is one component of **Compass**, Cohere's broader document-intelligence and enterprise search platform, "responsible for ingestion, visual parsing, and chunking" (cohere.com).

Cohere prices API access at **$1.50 per 1,000 pages** (cohere.com), roughly a third of Mistral OCR 4.1's headline rate. For dedicated capacity, Cohere's official pricing page lists Model Vault instance pricing for "Parse 5" at $4.00 per hour ($2,500 per month) for a Medium instance and $7.00 per hour ($4,300 per month) for an XL instance (cohere.com), and Cohere states that Model Vault reduces inference costs by approximately 23% compared with the API at 50% GPU utilization, rising to further savings at full utilization, citing a worked example of a 13-million-page-per-month accounts-payable workflow saving roughly $144,000 annually by moving from the API to Model Vault. Cohere states Parse "processes 4.5 pages per second (36 pages per second or 2160 pages per minute on an 8 H100 GPU node)" (cohere.com). Cohere's blog states Parse "is now generally available via the Cohere API, Model Vault, Microsoft Foundry, and AWS SageMaker" (cohere.com); notably, Microsoft's own Azure AI Foundry catalog page for the model lists its lifecycle status as "Preview" rather than GA (ai.azure.com), a discrepancy between two first-party sources that this report notes but cannot resolve.

Adoption

In sharp contrast to Mistral OCR 4.1's reception, Cohere Parse's own Hacker News submission drew only 4 points and no comments (news.ycombinator.com). Trade coverage of developer sentiment elsewhere describes Reddit's r/Rag community praising "table extraction, reading order, and pricing," while flagging the lack of OpenRouter availability and requesting "native PDF file input support to avoid needing to render PDFs page-by-page" (infoq.com). That last point is corroborated independently: one hands-on reviewer found that "the live endpoint currently takes images only," despite documentation listing PDF, PPT, and JPEG as supported formats, with image inputs capped at 20MB, 50 megapixels, or 200MB decoded (eesel.ai). The same reviewer reports the Cohere API rate limit for Parse at "500 requests/minute," with a free trial key allowing "1,000 calls/month" (eesel.ai).

Strengths and Limitations

Cohere's own documentation lists explicit product limitations: parsed content is delivered only as Markdown, since "structured JSON output is not supported" (docs.cohere.com), and the model does not return confidence scores or identify headers, footers, or font hierarchy, a direct contrast to Mistral OCR 4.1's block- and word-level confidence outputs. Cohere frames Parse's ideal use cases as "throughput, scale and cost effectiveness" rather than maximal accuracy (docs.cohere.com). Cohere's own internal comparison, discussed in the Executive Summary above and examined further in the Performance and Benchmarks section, claims Parse scores 79.2 against 74.5 for "Mistral OCR 4" on a benchmark it calls ParseBench, ahead of every model it tested except frontier general-purpose LLMs. Cohere's product page explains that its reported score deliberately excludes two of ParseBench's five dimensions: "Layout and Chart scores are excluded because they evaluate capabilities outside the current product scope" (cohere.com).

Feature Comparison

Table 1 below summarizes the two models' documented capabilities, pricing, and deployment options as of September 5, 2026, drawing only on facts each vendor publishes about its own product.

Dimension	Mistral OCR 4.1	Cohere Parse (parse-v5.0)
Release status	Preview July 16, 2026; GA August 31, 2026 (docs.mistral.ai)	Announced GA August 27, 2026 per Cohere; Azure catalog lists status "Preview" (ai.azure.com)
API pricing	Standard API: $4 per 1,000 pages; $5 per 1,000 annotated pages (docs.mistral.ai). Batch API: $2 per 1,000 pages (Mistral AI)	$1.50 per 1,000 pages (per Cohere's announcement, cited above)
Dedicated/self-hosted pricing	Self-hosted single-container deployment reported for regulated industries with no per-page billing (witho2.com)	Model Vault: Medium $4.00/hr ($2,500/mo); XL $7.00/hr ($4,300/mo) (cohere.com)
Structural metadata	13 block types (text, title, list, table, image, equation, caption, code, references, aside_text, header, footer, signature) with paragraph-level bounding boxes (cited above)	Blocks output limited to fewer type fields (e.g. text, table) with bounding boxes for tables; no header/footer/font-hierarchy identification (docs.cohere.com)

Dimension	Mistral OCR 4.1	Cohere Parse (parse-v5.0)
Confidence scores	Page, block, and word-level granularity (cited above)	Not supported (cited above)
Output formats	JSON with per-page Markdown, tables, hyperlinks, header/footer text, dimensions, confidence scores, blocks (cited above)	Markdown only, or a blocks format; tables output as HTML (ai.azure.com)
Multilingual coverage	"More than 40 languages" per official docs (cited above)	Nine "most prevalent commercial languages" (cohere.com)
Model size	Not publicly disclosed	2.3 billion parameters, approximately 4.6GB (docs.cohere.com)
Throughput claim	Up to 2,000 pages/minute on a single GPU (vendor/trade-press reported, not on Mistral's official docs page) (witho2.com)	270 pages/minute on one GPU; 2,160 pages/minute on an 8-GPU node (per Cohere's announcement, cited above)
Independent benchmark presence (as of 2026-09-05)	Present, under varying version labels, on olmOCR-Bench, OmniDocBench, and ParseBench's hosted leaderboard (see next section)	Not found on any of the three public leaderboards checked

Table 1 makes clear that the two models are not simply competing on the same axis. Mistral OCR 4.1 publishes richer structural metadata, confidence scoring at three granularities, and a wider documented block taxonomy, all useful for downstream validation workflows; Cohere Parse undercuts it substantially on published API price and claims higher raw throughput, but ships without confidence scores, without structured JSON, and, as of this writing, without a PDF-native API path. Both facts are vendor-documented, not independently audited, and buyers evaluating either for high-stakes document types (regulatory filings, clinical data tables) should weigh the confidence-score gap alongside the price gap.

Performance and Benchmarks

Because Mistral OCR 4.1's and Cohere Parse's own product pages do not agree on a shared, neutral scoring method, a reproducible comparison has to rest on benchmarks built and run by parties with no stake in either vendor's result. Three such suites exist publicly, and all three test the specific dimensions this report set out to examine (reading order, tables, equations, and citations), each with a documented, publicly reproducible method rather than an opaque internal score.

olmOCR-Bench, built by the Allen Institute for AI (Ai2), is "a dataset of 1,403 PDF files, plus 7,010 unit test cases" (huggingface.co), scored with "deterministic verifiers that assert properties like 'table structure preserved,' 'math faithfully transcribed,' or 'reading order consistent'" (allenai.org), a pass/fail unit-test design rather than fuzzy text similarity. Ai2's own README leaderboard lists an entry labeled "Mistral OCR API" scoring 72.0 (±1.1) overall, with sub-scores of 60.6 on tables and 71.3 on multi-column layout (raw.githubusercontent.com). That same leaderboard carries no entry for Cohere Parse (raw.githubusercontent.com). Ai2's own olmOCR 2 model, evaluated on this suite, scores 82.4 overall, beating pipeline tools Marker (76.1) and MinerU (75.8) (allenai.org), and Ai2 states plainly that "even strong OCR systems struggle with multi-column layouts, dense tables, math notation, and degraded scans" (allenai.org), the same four categories this comparison targets.

OmniDocBench, a peer-reviewed CVPR 2025 benchmark, was built because "current document parsing methods have not been fairly and comprehensively evaluated due to the narrow coverage of document types and the simplified, unrealistic evaluation procedures in existing benchmarks" (arxiv.org). Its public repository specifies "1651 PDF pages, covering 10 document types, 5 layout types, and 5 language types," including academic papers, financial reports, newspapers, textbooks, and handwritten notes (github.com), and it scores text with normalized edit distance, tables with the Tree-Edit-Distance-based Similarity (TEDS) metric, formulas with a dedicated CDM metric, and reading order with its own edit-distance score. On OmniDocBench's official leaderboard, an entry labeled simply "Mistral OCR" (version unspecified) scores 85.66 overall, with a table TEDS score of 76.78, its comparatively weakest sub-score against top performers (github.com). Again, no entry for Cohere or Cohere Parse appears on this leaderboard (github.com).

ParseBench, published by the LlamaIndex team behind LlamaParse, is an independently arXiv-documented benchmark (April 2026) of "human-verified enterprise pages" (infoq.com), evaluated across five capability dimensions (tables, charts, content faithfulness, semantic formatting, and visual grounding). Its methodology explicitly includes enterprise document-AI incumbents for comparison, listing "azure_di_layout Azure Document Intelligence aws_textract AWS Textract google_docai_layout Google Cloud Document AI" among its evaluated pipelines (github.com). ParseBench's own published leaderboard data lists two Mistral entries, "Mistral OCR 4" and "Mistral OCR 4 (Annotation)," ranked 49th and 28th of 96 evaluated methods respectively (raw.githubusercontent.com), with LlamaParse's own top configuration leading overall (arxiv.org). This public leaderboard snapshot, fetched directly for this report, contains no row for Cohere Parse.

Public benchmark entries use different model labels, which limits direct comparison of the two current products.

The underlying difficulty of the four target categories is well documented in the academic literature. Meta AI's Nougat paper argues that "the PDF format leads to a loss of semantic information, particularly for mathematical expressions" (arxiv.org), motivating markup-aware transcription rather than plain text. Microsoft's LayoutReader paper calls reading-order detection "the cornerstone to understanding visually-rich documents," noting the field lacked large-scale training data until it mined 500,000 pages of ground truth from Word document metadata (arxiv.org). PubTabNet, built from 568,000 table images matched against PubMed Central's XML source files, introduced the TEDS metric now used by OmniDocBench (arxiv.org). For citation and reference extraction, a comparative evaluation of ten open-source parsers found GROBID the strongest out-of-the-box tool at an F1 score of 0.89, ahead of CERMINE (0.83) and ParsCit (0.75) (arxiv.org), work that predates both Mistral OCR 4.1 and Cohere Parse but establishes the accuracy bar a modern VLM-based extractor should be measured against for this category.

Data Analysis and Evidence

Pricing across the document-extraction category varies by roughly two orders of magnitude depending on capability tier. *Table 2* places Mistral OCR 4.1 and Cohere Parse alongside the major cloud incumbents and two independent parsing services, all normalized to a per-1,000-page basis where each vendor bills that way.

Product	Basic OCR / text extraction	Table / form / layout extraction	Notes
Mistral OCR 4.1	Standard API: $4 / 1,000 pages (docs.mistral.ai); Batch API: $2 / 1,000 pages (Mistral AI)	$5 / 1,000 annotated pages (cited above)	Standard API rate shown; Batch API pricing is a 50% discount
Cohere Parse	$1.50 / 1,000 pages (cited above)	Included in same rate (Markdown/blocks output)	Model Vault dedicated hosting available from $2,500/mo (cohere.com)
Google Document AI	$1.50 / 1,000 pages, 1K to 5M pages/mo, dropping to $0.60 above 5M (cloud.google.com)	Layout Parser $10 / 1,000 pages; Form Parser $30 / 1,000 pages, dropping to $20 above 1M (cloud.google.com)	First 1,000 pages/month free
AWS Textract	$1.50 / 1,000 pages (Detect Document Text, first 1M pages/mo) (aws.amazon.com)	Tables $15 / 1,000 pages; forms $50 / 1,000 pages, additive (aws.amazon.com)	1,000 free Detect pages/mo; 100 free Forms/Tables pages/mo for 3 months
Azure AI Document Intelligence	$1.50 / 1,000 pages (Read, 0 to 1M pages/mo), $0.60 above 1M (azure.microsoft.com)	Prebuilt models $10 / 1,000 pages; custom models $30 / 1,000 pages (azure.microsoft.com)	Commitment tiers reach $0.53 / 1,000 pages at 8M pages/mo (azure.microsoft.com)
LlamaParse	$1.25 / 1,000 credits, roughly $1.25 / 1,000 pages at the base tier (llamaindex.ai)	Higher-fidelity modes consume more credits per page	"Auto Mode" routes each page to the cheapest adequate tier, "saves up to 80%" (llamaindex.ai)
Unstructured.io	$15 / 1,000 pages after the first 10,000 free pages/mo (unstructured.io)	Included in same flat rate, incl. VLM-based partitioning	Single-tier pricing across all extraction features
Anthropic Claude (PDF support)	Token-based, not per-page: approximately 7,000 tokens for a 3-page PDF in full visual mode (docs.anthropic.com)	N/A (general-purpose LLM, not a dedicated extraction API)	Max 600 pages or 32MB per request (docs.anthropic.com)

Table 2 shows Cohere Parse and the cloud hyperscalers' basic OCR tiers clustering near $1.50 per 1,000 pages, while Mistral OCR 4.1's **standard API** rate of $4 sits closer to the hyperscalers' mid-tier layout and prebuilt-model pricing ($10 to $30 per 1,000 pages). Mistral's published Batch API discount reduces that rate to $2 per 1,000 pages (Mistral AI), so high-volume comparisons should state whether they assume standard or batch processing. Ai2's original olmOCR paper frames the cost gap between general-purpose VLMs and purpose-built OCR starkly, noting that proprietary VLM-based extraction "over 6,240 USD per million PDF pages for GPT-4o" compares with 176 USD per million pages for its own open-weight olmOCR pipeline (arxiv.org), a reminder that dedicated OCR APIs, whatever their differences, remain an order of magnitude cheaper than routing every page through a frontier general-purpose model.

The category these two products compete in is growing quickly. Grand View Research values the global intelligent document processing (IDP) market at "USD 3.0 billion in 2025" and projects it to reach "USD 3.9 billion in 2026" and "USD 29.7 billion by 2033, at a CAGR of 33.8%" (grandviewresearch.com) (compound annual growth rate,

2026 to 2033). The same report finds North America held "the largest revenue share of 36.8% in 2025" and that banking, financial services, and insurance (BFSI) was the largest end-use segment at a 25.0% share that year (grandviewresearch.com).

For life-sciences and scientific-document use cases specifically, structural fidelity has a measurable downstream effect on data quality. A peer-reviewed 2026 study building an automated GPT-4o- and o3-based extraction system for systematic literature reviews found "accuracy was higher for string variables (e.g., country, study design, drug names, and outcome definitions) compared with numeric variables," the latter typically embedded in tables (pubmed.ncbi.nlm.nih.gov). A related 2024 study comparing GPT-4-turbo and Claude-3-Opus for automated literature extraction in living systematic reviews found "342 (87%) responses were concordant," with concordant-response accuracy at 0.94 versus a range of only 0.41 to 0.50 accuracy when the two models disagreed (pubmed.ncbi.nlm.nih.gov). Both findings underscore that table-embedded numeric data, precisely the category Mistral OCR 4.1 and Cohere Parse both dedicate specific metadata (bounding boxes, HTML structure) to extracting, is disproportionately the point at which automated extraction pipelines fail. This matters for regulatory workflows in particular: the FDA's Study Data Technical Conformance Guide states that "the Study Data Tabulation Model (SDTM) defines a standard structure for human clinical trials tabulation datasets" (fda.gov),

Implications and Future Directions

For a technical buyer choosing between these two models today, the decision reduces to a small set of concrete trade-offs rather than a single "which is better" answer. At Mistral's standard API rate, Mistral OCR 4.1 costs roughly 2.7 times Cohere Parse's published per-page rate; at Mistral's published Batch API rate of $2 per 1,000 pages, it is roughly 1.3 times Cohere's $1.50 rate (Mistral AI) but returns page-, block-, and word-level confidence scores, a wider documented block taxonomy, and, per independent trade coverage, a self-hosted deployment path for regulated environments that avoid per-page billing altogether. Cohere Parse costs less, claims higher raw page-per-second throughput on comparable hardware, and is built around Markdown output aimed at RAG and search ingestion pipelines, but ships without confidence scores and, as of the observation date, without a working native PDF upload path at the API level, according to independent testing.

Organizations evaluating these two models for a real workload should test representative documents against their own requirements. IntuitionLabs, a life-sciences and AI consultancy that does not sell either product, frames the broader stakes of this kind of upstream data-quality decision in terms of governance rather than raw accuracy: as noted earlier, its own materials describe the goal of an information layer as ensuring AI systems work from "authoritative sources through identity, permissions, retrieval, citations, logging, and accountable operations" (intuitionlabs.ai). From that vantage point, the choice between Mistral OCR 4.1 and Cohere Parse is less about picking a winner than about matching each model's documented strengths (confidence-scored structural metadata versus low-cost high-throughput ingestion) to the specific document types and compliance requirements a given workflow must satisfy; readers seeking the wider landscape of OCR and document-AI tools beyond these two products can consult IntuitionLabs' earlier comparative survey (intuitionlabs.ai).

Conclusion

Mistral OCR 4.1 and Cohere Parse both reached the market within six weeks of each other in mid-to-late 2026, but they were built for different priorities. Mistral OCR 4.1, generally available since August 31, 2026 at a standard API rate of $4 per 1,000 pages (or a published Batch API rate of $2 per 1,000 pages) (Mistral AI), emphasizes structural transparency: paragraph-level bounding boxes, thirteen labeled block types, and confidence scores at three granularities, features aimed at workflows where a human or downstream system needs to verify what the model extracted. Cohere Parse, announced August 27, 2026 at $1.50 per 1,000 pages, emphasizes throughput and cost at scale, with Cohere itself framing the trade-off as sacrificing certain capabilities (structured JSON, confidence scores, chart and fine-grained layout scoring) in exchange for price and speed. Buyers should evaluate each model's documented capabilities against the document types and workflow requirements relevant to their use case.

Frequently Asked Questions (FAQs)

Is Mistral OCR 4.1 more accurate than Cohere Parse? The article documents different product capabilities and pricing; organizations should evaluate representative documents against their own workflow requirements.

How much does Mistral OCR 4.1 cost compared to Cohere Parse? Mistral OCR 4.1's standard API rate is $4 per 1,000 pages ($5 per 1,000 annotated pages); Mistral's OCR 4 announcement also lists a 50% Batch API discount, or $2 per 1,000 pages (Mistral AI). Cohere Parse is $1.50 per 1,000 pages via the API, or a fixed monthly rate starting at $2,500 for dedicated Model Vault capacity, as detailed in the Cohere Parse section above.

Is Cohere Parse generally available yet? Cohere's own blog states it is GA via the Cohere API, Model Vault, Microsoft Foundry, and AWS SageMaker (cited in the Cohere Parse section above), but Microsoft's own Azure AI Foundry catalog listing for the model shows lifecycle status "Preview" (ai.azure.com), a discrepancy between two first-party sources this report could not resolve.

Which model handles tables better? No independent leaderboard scores both models' table handling side by side. On OmniDocBench, the unversioned "Mistral OCR" entry scores 76.78 on table TEDS, its weakest relative sub-score (github.com).

What is the best OCR model for document structure in 2026? It depends on the requirement. Mistral OCR 4.1 documents a wider structural taxonomy (13 block types) and multi-granularity confidence scores, useful where extraction needs to be auditable. Cohere Parse documents faster throughput and lower published pricing, useful for high-volume ingestion where per-item review is not feasible. Neither claim has been independently confirmed head to head as of this report.

How is document-extraction accuracy typically measured? Independent benchmarks use several distinct, purpose-built metrics: olmOCR-Bench uses pass/fail unit tests for properties like table structure and reading order (allenai.org); OmniDocBench uses normalized edit distance for text, TEDS for tables, and a dedicated CDM metric for formulas; ParseBench uses table-record matching and element-level visual grounding checks.

Are either of these models validated for life-sciences or regulated document use cases specifically? Neither vendor publishes a life-sciences-specific validation study for its OCR product. General research on automated extraction from scientific literature shows accuracy consistently drops on table-embedded numeric data relative to categorical text (pubmed.ncbi.nlm.nih.gov), which is the same category both Mistral OCR 4.1 and Cohere Parse

dedicate the most structural metadata to extracting; organizations working toward SDTM-conformant regulatory submissions (fda.gov) should independently validate either model's table output against source documents rather than assume general-purpose accuracy transfers to their specific document types.
