

# Nanonets OCR-3: The new state-of-the-art OCR model built for the agentic stack

OCR-3 ranks #1 on global OCR benchmarks. It features bounding boxes, confidence scores, native VQA, and more capabilities built-in for the agentic stack.

	Nanonets OCR-3	SPECIALIZED VLMS				GENERAL VLMS	
		Chandra OCR 2	LightOn OCR-2	Deepseek-OCR2	Mistral OCR 3	Gemini 3.1 Pro	GPT-5.4
<b>OlmOCR Benchmark</b>	<b>87.4</b>	<b>85.9</b>	83.2	76.3	81.7	79.6	81.0
ArXiv Math	89.2	90.2	89.6	81.9	85.4	70.6	83.1
H&F	96.6	92.5	-	95.6	93.8	-	-
Long/Tiny	93.4	92.1	91.4	88.7	88.9	90.3	82.6
Multi-Col	87.6	83.5	84.8	83.6	82.1	79.2	83.7
Old Scans	49.6	49.8	42.2	33.7	48.8	47.5	43.9
Scans Math	88.9	89.3	85.6	68.8	68.3	84.9	82.3
Tables	94.2	89.9	89.0	78.1	86.1	84.9	91.1
<b>OmniDocBench (v1.5)</b>	<b>90.5</b>	85.5	-	87.7	85.3	85.3	85.3

**San Francisco, California Apr 9, 2026 (Issuewire.com)** - [Nanonets has released OCR-3](#), the latest model in their family of visual language models (VLMs) specialized for OCR and document understanding tasks. It is specifically built for agents, RAG, and other LLM-based use cases.

Based on benchmarks, it is the most accurate OCR model today.

- 93.1 on OLM-OCR (Global #1)
- 85.9 on IDP Leaderboard (Global #1)
- 90.5 on OmniDocBench

The team also released scores on domain-specific document benchmarks:

- 94.5% on FinanceBench (dense SEC 10-K filings)
- 96.0% on DocBench Legal (multi-column court filings)
- 90.1% on HealthcareBench (clinical notes and forms)

OCR-3 provides bounding boxes and confidence scores as metadata in its outputs, two features that are sorely missed in document pipelines built on foundational models or other VLMs.

- **Confidence scores:** Every extraction includes a reliability score. You can pass high-confidence outputs directly, and route low-confidence outputs to a human or a larger model.
- **Bounding boxes:** The model outputs exact spatial coordinates for every element. These are useful for RAG citations, feeding agents precise document regions, source highlighting, and observability in the agentic stack.

Additionally, the model natively supports visual question-answering with answers backed by source evidence from documents. It also provides an RAG mode that converts documents into chunks of structured markdown optimized for RAG applications and agents.

The model API exposes five endpoints:

- /parse - returns structured markdown with exact reading order.
- /extract - returns a type-safe object based on your custom schema.
- /split - separates large batches of mixed documents.
- /chunk - creates context-aware document chunks for RAG.
- /vqa - returns grounded answers with exact source regions.

[Nanonets](#) OCR-3 is a 35-billion parameter Mixture-of-Experts (MoE) model, which makes it twice as fast as OCR-2. It is trained on a dataset of 11 million documents.

Developers can access the [OCR-3 model API](#) starting today.



Nanonets

\*\*\*\*\*@nanonets.com

<https://nanonets.com>

Source : Nanonets

[See on IssueWire](#)