

MCP SERVER

NO CODE

CLOUD HOSTED

DocSumo MCP for AI Agents

Extract structured details from invoices, bank statements, and IDs

DocSumo equips your AI agent to handle complex document data extraction. It automates processing for invoices, bank statements, and ID cards, giving you structured results right out of the box. Use it to manage end-to-end Intelligent Document Processing (IDP) workflows and audit every file processed.

A+ Quality Score 100/100

intelligent-document-processing

data-extraction

idp

invoice-processing

automated-pipelines

data-audit



The connectivity layer between AI and the world's software.



Vinkius sits between AI and every application. All communication passes through Vinkius Cloud via the Model Context Protocol (MCP) — with governance, observability, and security at every layer.

Your AI Connections Run Through Vinkius Cloud

The world's largest
managed MCP catalog

Vinkius is the connectivity layer where AI connects to the software your business already runs. We handle the hosting, the security, the credentials, the uptime — you get agents that actually do things.

We operate the world's largest managed MCP catalog. Major SaaS platforms, CRMs, databases, and cloud providers — running, monitored, production-ready. This MCP server is hosted and maintained by the Vinkius Cloud for AI Agents.

The agent doesn't manage credentials, doesn't manage uptime, doesn't manage security. Vinkius does.

— Architecture principle

Four Pillars of the Vinkius Runtime

01 — Security by design

Credentials stay encrypted at rest via AES-256. The AI agent never touches raw keys — they're injected into a sandboxed V8 isolate at runtime. Actions are logged, and connections have an emergency kill switch.

03 — Deterministic observability

Eight immutable metrics per endpoint: request volume, p95 latency, error rate, active connections, cost attribution. A live payload feed logs every tool call with mutation detection.

02 — Built on MCP Fusion

This MCP server was built with **MCP Fusion**, the open-source framework (Apache 2.0) that powers the entire Vinkius catalog. Schema-as-firewall strips undeclared fields, compiled PII redaction runs at zero overhead, and cryptographic lockfiles produce git-diffable audit trails.

04 — Autonomous operations

Servers are deployed, monitored, and patched autonomously. New capabilities and security patches ship weekly. Zero-downtime deployments ensure continuous availability across all managed MCP servers.

AES-256

Encryption at rest

Ed25519

PKI vault signatures

24h TTL

Ephemeral session keys

V8 Isolate

Sandboxed execution

One Token. Instant Access.

Every MCP server on Vinkius is accessed through a **Connection Token**. Tokens are generated in the cloud dashboard and produce a unique MCP endpoint URL. Paste this URL into any MCP-compatible client — no SDK required.

A single token can serve **multiple AI clients simultaneously**, or you can issue separate tokens per client for granular access control. Each token tracks its own request count, last activity timestamp, and can be individually enabled or revoked.

MCP ENDPOINT

`https://edge.vinkius.com/{token}/mcp`

Claude



Cursor



VS Code



Windsurf



Grok



Gemini

Security Is the Architecture

Security in Vinkius is not a feature — it's the foundation of the runtime. The gateway enforces multiple independent protection layers between AI agents and third-party APIs.

01 — Ed25519 PKI Vault

Every workspace has an Ed25519 Master Key. Session keys are generated ephemerally (24h TTL) and signed by the Master Key. Credentials never leave the vault boundary.

02 — V8 Isolate Sandboxing

Tool code runs inside isolated-vm V8 isolates with 64 MB memory caps and per-request timeouts. No filesystem access, no network access except through the SSRF-guarded fetch bridge.

03 — SSRF Guard

All outbound HTTP requests are DNS-resolved and validated before execution. Private IP ranges (10.x, 172.16-31.x, 192.168.x, AWS metadata 169.254.x) are blocked at the network layer.

05 — Cryptographic Audit Trail

Every request is signed into a SHA-256 hash chain with Ed25519 signatures. Events form a tamper-proof, SIEM-exportable forensic record.

04 — DLP & PII Redaction

A ResponseGuard pipeline intercepts every tool response. Configurable redaction patterns strip sensitive fields (emails, SSNs, card numbers) before data reaches the AI agent.

06 — Honeypot Trap System

Phantom credentials are injected into isolated environments. If a honeypot is used outside Vinkius infrastructure, the server is quarantined instantly.

Emergency Kill Switch

EU AI Act Art. 14(1)
Compliant

The kill switch is an **emergency halt** mechanism — not a simple toggle. When triggered, it executes three actions atomically:

01 — Server deactivated

The MCP server is immediately taken offline across the entire cluster.

02 — All tokens revoked

Every connection token is invalidated. Total lockout — reconnection blocked until new tokens are issued.

03 — WebSocket connections killed

Active connections terminated via Redis pubsub broadcast. Propagates to every runtime node in the cluster.

Full Visibility. Zero Guesswork.

The Vinkius cloud dashboard includes a full MCP Governance suite — real-time analytics and security controls for production AI operations.

Control Plane

KPI dashboard with request volume, latency, success rate, token consumption, and AI-generated operational briefings.

FinOps

Cost tracking per tool, payload compression savings, budget optimization signals, and consumption trends.

Firewall & DLP

PII redaction activity, sensitive data protection counters, and security event timeline.

Agent Activity

Which AI clients are connecting, how often, and what they're doing — real-time session tracking.

Tool Health

Slowest and most error-prone tools, with actionable root-cause insights and performance baselines.

Incident Log

Error trends, failure rates, status-code breakdowns, and forensic audit trail access.

Get started at cloud.vinkius.com — connect your AI agent in under 60 seconds.

DocSumo MCP

10 tools available

Cloud-hosted on Vinkius

DocSumo connects advanced document intelligence directly into your AI workflow. You can now automate pulling complex data from invoices, bank statements, or even government ID cards. Instead of manually reviewing files, your agent handles the heavy lifting, structuring messy information for you to use immediately.

If your job involves processing high volumes of varied documents, this MCP lets you manage everything from start to finish. You can monitor entire document pipelines, see exactly which fields were extracted, and even check a record's confidence score if it needs human eyes. Finding the right connection point used to be tough, but Vinkius makes it easy; you connect once to access DocSumo's full suite of features directly through your preferred AI client.

This means finance teams can pull invoice details instantly for bookkeeping, compliance officers can audit IDs and bank statements for verification status, and operations leads can keep a real-time pulse on the entire document processing health.

Core Capabilities

01 — Pull structured data from documents

Your agent pulls specific details—like invoice numbers or total amounts—from complex file types.

03 — Review documents needing manual checks

The system flags records where the AI wasn't confident enough and needs human review.

05 — List all document types available

Your agent shows you every type of document (invoices, bank statements) configured within the system.

02 — Audit and track failed extractions

You find out exactly which files didn't get processed correctly, along with error messages.

04 — Check overall processing health

You get a quick, high-level summary of how many documents were processed successfully in total.

One Click on Vinkius — From Prompt to Execution

Available at vinkius.com/mcp/docsumo — connect your AI agent in three steps.

- 01** Connect DocSumo's MCP to your AI client and authorize it using your API key.
- 02** Tell your agent what kind of data you need, like 'all outstanding invoices from last month' or 'any documents needing review'.
- 03** The system processes the request, pulls the structured data, and presents actionable results in a conversation.

The bottom line is that it turns unstructured files into clean, usable data points using simple chat commands.

Built For

This MCP is critical for anyone dealing with high volumes of paper-based or digital documents. It helps compliance officers who can't afford audit gaps, finance teams drowning in receipts, and operations leads needing a single view of document status.

Accounts Payable Specialist

You use the MCP to quickly pull invoice details from vendor files so you can process payments without manual data entry.

Compliance Officer

You audit processed ID cards and bank statements using the agent, instantly verifying if documentation meets regulatory standards.

Operations Manager

You monitor document processing success rates across multiple pipelines to troubleshoot bottlenecks before they impact deadlines.

What Changes When You Connect

- 01** Stop manual data entry. Instead of copy-pasting numbers from PDFs into spreadsheets, your agent extracts the exact figures directly.

-
- 02** Maintain compliance control by using tools like `list_documents_awaiting_review` to ensure no sensitive file is missed or unchecked.
-
- 03** Get a clear view of operational status with `quick_idp_health_audit`, letting you troubleshoot processing failures instantly rather than waiting for end-of-day reports.
-
- 04** Speed up bookkeeping by using `get_document_extraction_data` to pull verified invoice totals and vendor names on demand, without opening the file.
-
- 05** Keep track of your assets with `list_processed_documents`, giving you a full audit trail of everything that passed through the pipeline.
-

Real-World Applications

Auditing Compliance Documents

A compliance officer needs to verify 50 ID cards against current policy. They ask their agent to run `list_documents_awaiting_review` and get a list of all flagged IDs, which they then review in bulk using the MCP.

Finding Specific Missing Invoices

An operations lead is looking for an invoice from 'XYZ Corp' that was processed last week. They use the agent to `search_documents_by_filename` and immediately pull up the correct document record.

Quickly Reconciling Bank Statements

A finance analyst needs to reconcile last month's expenses. They ask their agent to `list_latest_extraction_results` for bank statements, pulling all transaction dates and amounts into a single chat summary.

Verifying a Specific Record

A user needs to validate the grand total on a specific invoice ID. They ask their agent, which then uses `get_document_extraction_data` to provide the structured number and its confidence score.

Patterns to Avoid

Assuming all files are clean

X AVOID

Manually processing a batch of documents without first checking for errors. You risk missing data or submitting inaccurate reports based on failed extractions.

✓ INSTEAD

Always run ``list_failed_doc_extractions`` before starting major processes. This ensures your agent spots any files that couldn't be read, preventing bad data from entering your system.

Over-relying on manual dashboards

X AVOID

Spending hours clicking through separate dashboard tabs to check document success rates or process status for different departments.

✓ INSTEAD

Use ``quick_idp_health_audit`` with your agent. It pulls a consolidated, real-time summary of the entire pipeline's health in one request.

Ignoring metadata limits

X AVOID

Trying to process an enormous volume of documents without knowing your account's usage capacity, leading to unexpected service interruptions.

✓ INSTEAD

First, use ``get_docsumo_account_metadata``. This tells you what your current usage limits are, so you can plan processing jobs safely.

The Right Fit

Use DocSumo if your core problem is moving data from unstructured documents (PDFs, images) into structured databases or reports. Specifically, connect it if you need to audit compliance records or automate repetitive bookkeeping tasks involving bank statements and invoices. Don't use this MCP if your only goal is simple text summarization; for that, a general-purpose agent will do. However, if the data source is always messy, varied, or requires complex field identification (like matching an invoice number to a specific vendor), then DocSumo is mandatory because it handles the document intelligence layer.

DocSumo MCP for AI Agents: Automating Finance Data Extraction from Documents

Currently, finance teams spend hours dealing with paper stacks and digital PDFs. They have to open dozens of files, copy the invoice number into a spreadsheet, paste the total amount in another cell, and manually verify if the bank statement date matches the receipt date. It's tedious, error-prone, and slows down closing books.

With this MCP, your agent handles that entire process conversationally. You simply ask for 'all vendor payments over \$500 from last quarter,' and it pulls out the verified invoice numbers, amounts, and dates into a clean list—no copy/pasting required.

DocSumo MCP for AI Agents: Improving Compliance Audits with IDP

For compliance roles, the process is worse. You can't just trust a single source of truth; you need to cross-reference multiple IDs and documents (like passports against utility bills). This means tedious manual checks for discrepancies and logging every verification step.

Now, your agent manages that audit trail. It pulls data from various sources and lets you check confidence scores instantly. You get an auditable, structured report showing exactly what was verified and which items need a human look.

DocSumo: 10 Tools for Intelligent Document Processing (IDP) Data Extraction

Use these tools to pull specific metadata, audit failure lists, get success summaries, and search through all your processed document records.

#	TOOL	DESCRIPTION
01	<code>get_docsumo_account_metadata</code>	Retrieves usage limits and general status information for your DocSumo account.
02	<code>list_failed_doc_extractions</code>	Finds a list of documents that could not be processed by the system.
03	<code>get_document_extraction_data</code>	Pulls the structured data points extracted from one specific document you name.
04	<code>quick_idp_health_audit</code>	Provides a fast, high-level summary of how well your entire document processing pipeline is running.
05	<code>list_successfully_parsed_docs</code>	Gives you a list of all documents that finished the process and passed verification checks.
06	<code>list_docsumo_document_types</code>	Lists every document format (like invoices or bank statements) that is currently configured in DocSumo.
07	<code>list_processed_documents</code>	Shows a list of all documents processed by the system, which you can filter by type.
08	<code>list_latest_extraction_results</code>	Identifies and lists the most recently processed document results across all categories.
09	<code>list_documents_awaiting_review</code>	Finds documents that need human eyes because the AI wasn't confident enough with the data.
10	<code>search_documents_by_filename</code>	Searches for processed files using a specific keyword found in the document name.

See It in Action

Real prompts you can use once this MCP is connected to your AI agent through Vinkius Cloud.

- U** Pull all the bank statements processed last month and list the top 5 vendors.



Summary of Processed Bank Statements (Last Month)

VENDOR NAME	TOTAL AMOUNT PULLED	DATE VERIFIED
Chase Bank	\$12,450.00	2023-11-28
Wells Fargo	\$9,120.55	2023-11-27
UtilityCorp	\$350.00	2023-11-26

Need the detailed line items for Chase Bank? Let me know.

- U** Check if there are any ID cards that need manual review.



🔔 Action Required: Manual Review List 🔔

We found **3** documents needing human verification due to low confidence scores. Please check these immediately:

- `ID_Scan_009.jpg` : Low match on name/address.
- `Passport_BETA.pdf` : Poor image quality, date unreadable.
- `Utility_Record_78.png` : Discrepancy found in serial number.

Reviewing these files ensures our compliance record is spotless.

U What are the most recently processed documents and their total value?



Latest Document Extraction Results

- **Document:** Invoice_ABC-123.pdf (Processed: 5 minutes ago)
 - * Vendor: Global Supplies Inc.
 - * Grand Total: **\$4,500.00** (Confidence: 99%)
- **Document:** Statement_XYZ.pdf (Processed: 3 hours ago)
 - * Period: Oct 1 - Oct 31
 - * Total Amount Pulled: **\$22,500.00** (Confidence: 98%)

Frequently Asked Questions

01 How does the DocSumo MCP help with invoice processing?

It automates the difficult process of pulling structured data from messy invoices. Instead of manually typing out the vendor name, invoice number, and total amount, your agent pulls these verified fields directly into your chat conversation.

02 Is DocSumo good for compliance auditing?

Yes. It helps you maintain audit trails by allowing you to flag documents that need human review using the ``list_documents_awaiting_review`` tool. This ensures no critical ID or document is overlooked.

03 What if I have different types of files, like receipts and bank statements?

The MCP handles various formats. You can list all available document types to see what it supports—from basic receipts to complex bank statements—and process them all through one consistent workflow.

04 Can I check the status of a whole batch of documents?

Absolutely. You can use ``quick_idp_health_audit`` to get a single summary report on your entire pipeline's success rate, so you know immediately if there are bottlenecks.

05 I need to search for an old document by name; how does the DocSumo MCP help?

You can use the agent to ``search_documents_by_filename``. You just give it a keyword, and it finds all relevant processed documents, giving you direct access to their extracted data.

06 Is this better than using a separate dedicated document management tool?

This MCP integrates the intelligence layer directly into your existing workflow. It doesn't just store files; it extracts and structures the usable data, making the information actionable immediately within your agent conversation.

Go Live in 60 Seconds

Get your connection token from cloud.vinkius.com, then paste the endpoint URL into any MCP-compatible client.

YOUR MCP ENDPOINT

```
https://edge.vinkius.com/[TOKEN]/mcp
```

CLIENT

WHERE TO CONFIGURE



Claude AI

Profile → Customize → Connectors → "+" → Add custom connector → Paste endpoint



Cursor

Settings → Features → MCP Servers → "+ Add New MCP Server" → Type: SSE → Paste endpoint



VS Code

Ctrl/Cmd+Shift+P → "MCP: Add Server" → add `"docsumo": { "url": "..." }`



Windsurf

MCP Settings → `mcp_settings.json` → Add endpoint URL



ChatGPT

Settings → Tools & plugins → Add MCP server → Paste endpoint



Gemini

Extensions → Add MCP Server → Paste endpoint URL

ASK AN AI
ABOUT THIS

Let your preferred AI
explain this MCP server



Ask ChatGPT



Ask Claude



Ask Perplexity



Ask Gemini



Ask Grok



READY TO CONNECT

DocSumo is live on Vinkius Cloud.

Get your connection token, paste it into your AI agent, and start building. No SDK. No deployment. Just results.

[Start at cloud.vinkius.com](https://cloud.vinkius.com) →

vinkius.com · support@vinkius.com

INDEPENDENT PLATFORM DISCLAIMER

Vinkius is an independent platform and is not affiliated with, endorsed by, sponsored by, verified by, or otherwise authorized by DocSumo. All third-party trademarks, logos, and brand names are the property of their respective owners. Their use in this document is strictly for informational purposes to identify service compatibility and interoperability.

DOCUMENT INFORMATION

Generated	June 2026
MCP Server	DocSumo MCP
Server ID	019d7587-7253-73c1-878b-6fe528a502e2
Platform	Vinkius Cloud for AI Agents
Endpoint	https://edge.vinkius.com/{token}/mcp

LICENSE & USAGE

This document is generated automatically by the Vinkius PDF Engine. Content reflects the MCP server configuration at the time of generation and may change as updates are deployed. For the most current information, visit vinkius.com/mcp/docsumo.