MCP SERVER | NO CODE | CLOUD HOSTED

Clarifai (Vision AI) MCP for AI Agents

Analyze visual data and manage image recognition workflows

Clarifai gives your AI agent full control over complex computer vision and machine learning workflows. You can run automated image predictions, list specific models and apps, audit datasets for consistency, and manage entire multi-step computational pipelines directly through natural language conversation.

A+ Quality Score 100/100

computer-vision | machine-learning | model-inference | neural-networks | image-recognition | ai-workflows

MAINTAINED BY VINKIUS CLOUD FOR AI AGENTS · JUNE 2026
vinkius.com/mcp/clarifai-vision-ai

SCAN TO VIEW

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.

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.

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.

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	Ed25519	24h TTL	V8 Isolate
Encryption at rest	PKI vault signatures	Ephemeral session keys	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.

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.

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.

06 — Honeytoken Trap System

Phantom credentials are injected into isolated environments. If a honeytoken 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.

Agent Activity

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

FinOps

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

Tool Health

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

Firewall & DLP

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

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.

Clarifai (Vision AI) MCP

6 tools available
Cloud-hosted on Vinkius

Connecting Clarifai to your AI client lets you take the guesswork out of visual AI development. Instead of writing boilerplate code or logging into a separate dashboard, your agent manages your compute environment entirely via chat. You can run automated validation inferences on images and get exact network predictions—for example, identifying bounding boxes around detected objects. Need to audit what's running? Your agent lets you list every app, model, and workflow currently in use. If you're building something complex that needs multiple models chained together, you retrieve those composed computational blocks right from the chat interface. This MCP makes advanced visual data management accessible by letting your AI client handle all the heavy lifting. You'll find this entire capability cataloged within Vinkius, giving you one place to connect and control your most sophisticated ML services.

Core Capabilities

01 — Run Automated Inference Predictions

Send an image or data input, and the MCP returns explicit network predictions detailing what was evaluated in the visual data.

02 — Audit AI Infrastructure Components

List all active applications and models to get a clear inventory of your global compute resources.

03 — Manage ML Workflow Chains

Retrieve the structure of complex computational blocks that link multiple specialized models together for multi-step tasks.

04 — Verify Training Data Assets and Concepts

Identify data structures used for training and extract semantic concepts tagging your visual datasets.

One Click on Vinkius — From Prompt to Execution

Available at **vinkius.com/mcp/clarifai-vision-ai** — connect your AI agent in three steps.

01 Subscribe to this MCP and provide your Clarifai Personal Access Token (PAT).

02 Your AI client authenticates the connection, giving it permission to manage your vision assets.

03 You ask your agent a natural language question—like 'What apps do I have?' or 'Predict what's in this image.'—and get immediate results.

The bottom line is you control complex visual AI operations through plain conversation, without needing to write any code.

Built For

This MCP targets technical roles that spend time validating or building machine learning pipelines. If you're an ML Engineer constantly monitoring models, a Data Scientist auditing training data, or a developer prototyping vision features, this is for you.

ML Engineer

Monitoring and managing active compute brains (models) and their execution contexts across different applications.

Data Scientist

Auditing datasets and concepts to ensure training data consistency before launching new visual features.

AI Developer

Testing model predictions and complex workflow logic using natural language instead of writing tedious, repetitive code.

What Changes When You Connect

01 Run automated inference directly: Use `predict_model` to get immediate, detailed predictions on images without writing a single line of prediction code.

02 Audit your whole system easily: Quickly list all active resources using `list_apps` or running `list_models`, giving you an instant overview of your compute environment.

03 Handle complex tasks simply: Instead of building multi-stage pipelines, use `list_workflows` to retrieve and manage composed computational blocks that tie multiple models together.

04 Ensure data quality: Use `list_datasets` and `list_concepts` to audit exactly what data is being used for training and what tags are applied to it.

05 Maintain logic integrity: You can track your AI setup using `list_apps` and understand how different services interact across various execution contexts.

Real-World Applications

Debugging a Failed Image Classifier

A developer gets poor results on a new feature. Instead of debugging the code, they use their agent to run `predict_model` with sample images. The detailed network predictions instantly show where the model is failing (e.g., misidentifying bounding boxes), allowing them to fix the underlying data or logic.

Preparing a New Vision Feature

A product team needs to verify if their current training set is adequate. They ask the agent to run `list_datasets` and then use `list_concepts`. This confirms they have both enough raw images *and* consistent semantic tagging across all sources.

Inventorying Production AI

An ML Engineer needs a full audit of the company's deployed visual assets. They ask their agent to execute `list_apps` and `list_models`. This provides an immediate, organized inventory of every active compute brain they manage.

Understanding Complex Pipelines

A data scientist needs to know how a critical feature is built. They use the agent to run `list_workflows`. This instantly maps out all the chained models and services required, saving hours of manual architectural review.

Patterns to Avoid

Assuming Model Availability

✗ AVOID

Writing code that assumes a specific model ID exists for prediction without first checking if it's active or correctly configured.

✓ INSTEAD

First, always run `list_models` to confirm the correct parameters are available. Then, use `predict_model` with the validated ID to ensure your inference attempt succeeds.

Ignoring Data Lineage

✗ AVOID

Developing a new feature and only testing it on live data without understanding which specific concepts or datasets trained the model.

✓ INSTEAD

Before building, run `list_datasets` followed by `list_concepts`. This verifies both the raw image sources and the semantic tags used to train your visual data.

Over-Complicating Workflows

✗ AVOID

Trying to manually stitch together multiple models and services in code without a clear architectural map of dependencies.

✓ INSTEAD

Use `list_workflows` first. This tool lets you see the pre-composed, proven computational blocks that tie several specialized AI limits together, simplifying your deployment.

The Right Fit

You should use this MCP if your work involves managing complex visual data or running automated machine learning inference on images. Specifically, if you need to check model availability (`list_models`), run predictions (`predict_model`), or audit the entire infrastructure (using `list_apps` and `list_workflows`). Don't use it if you only need simple text generation; a general-purpose language MCP will handle that better. If your goal is purely data storage management without ML components, look for a dedicated database connector instead.

Clarifai (Vision AI) MCP: Managing Computer Vision Inference

Manually setting up complex visual pipelines means logging into multiple developer consoles. You have to check model status, pull the correct dataset IDs, and then manually construct API calls for every single prediction step. It's a tedious cycle of checking dashboards and copying configuration parameters.

With this MCP, you just ask your agent to run an inference. Your AI client handles all the background checks—it validates the models, manages the compute limits, and returns the full JSON analysis in one conversational response. You get immediate, actionable visual results.

Clarifai (Vision AI) MCP: Auditing ML Infrastructure with Vision Data

Today, understanding your entire ML setup requires running several separate queries just to list what apps are active or how many models exist across different projects. It’s a fragmented view of your compute environment.

Now, you ask the MCP agent for an inventory. In one conversational turn, it delivers a complete listing of all applications and models. You get a single source of truth about your entire visual AI stack.

6 Vision Data Operations Tools in Clarifai (Vision AI)

Use these tools to list apps, run predictions on images, map datasets, and manage entire machine learning workflows.

#	TOOL	DESCRIPTION
01	`list_apps`	Lists all the specific Clarifai applications that manage your compute limits.
02	`list_models`	Extracts structural details about the computer vision parameters driving your AI features.
03	`predict_model`	Runs an automated inference on a model, returning explicit network predictions and classifications for an image.
04	`list_workflows`	Retrieves the exact structure of composed computational blocks that tie multiple AI models together.
05	`list_datasets`	Identifies and maps data structures used for training your visual nodes.
06	`list_concepts`	Extracts the semantic concepts that are explicitly attached as tags to your datasets.

See It in Action

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

U Show me my available vision apps and any custom models I've deployed for face recognition.

🔎 **Clarifai App Inventory**

- General-Vision (Active)
- Face-Recognition (Active)
- Image-Moderation (Testing)

👤 **Models Found:**

1. `general-v2`: High-level classification model.
2. `face-rec-v3`: Optimized for identifying known individuals.

U Can you predict what's in this photo, and tell me the confidence level?

✅ **Inference Complete**

The image contains:

- **Person**: 98.5% (Bounding Box: [x1, y1, x2, y2])
- **Outdoors**: 94.0%
- **Tree**: 78.2%

I've also attached the full JSON payload detailing all bounding boxes and scores.

U What datasets are available for my training? I need to check concept tagging.

Dataset Audit: Custom-Trainer

- `training-v1` (500 images)
- `validation-v1` (100 images)

I also found 3 core concepts attached: **'outdoor'**, **'person'**, and **'vehicle'**. These are the tags used across all provided data.

Frequently Asked Questions

01 **How do I use the Clarifai (Vision AI) MCP to test a new image model?**

You simply ask your agent to run an inference on a specific model. You don't worry about API endpoints or payload formatting. Just tell it what you want to predict, and it runs the automated validation inference for you.

02 **Can the Clarifai (Vision AI) MCP help me understand my entire ML system?**

Yes. You can list all your active apps and models in one go. This gives you a clear, comprehensive map of every single compute resource running for your visual AI efforts.

03 **Is the Clarifai (Vision AI) MCP useful for auditing my training data?**

Absolutely. You can list datasets and concepts to ensure that your training data is consistent and correctly tagged before you launch a new feature, which prevents major ML bugs later on.

04 **Does the Clarifai (Vision AI) MCP handle complex multi-step workflows?**

Yes. It retrieves composed computational blocks, meaning it understands how multiple different models need to talk to each other sequentially. You don't have to build that logic yourself.

05 **What if I want to know which apps are currently active in my account?**

Just ask your agent to list the applications. It quickly gives you an inventory of every bounded Clarifai app, helping you organize and audit all your current ML deployments.

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 `"clarifai-vision-ai": { "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

Clarifai (Vision AI) 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 →

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 Clarifai (Vision AI). 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	Clarifai (Vision AI) MCP
Server ID	`019d7570-bfc9-7062-a7e9-d4a69d73d425`
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/clarifai-vision-ai.