MCP SERVER | NO CODE | CLOUD HOSTED

Dataiku DSS MCP for AI Agents

Manage Enterprise Data Pipelines and Model Monitoring

The Dataiku DSS MCP connects your AI client directly to your entire data science environment. You can list projects, check dataset schemas, monitor complex pipeline jobs, and audit ML model performance without leaving your chat interface. It puts full control of enterprise data workflows right into conversation.

A+ Quality Score 100/100

data-science | ml-ops | pipeline-orchestration | predictive-modeling | data-pipelines | automation

MAINTAINED BY VINKIUS CLOUD FOR AI AGENTS · JUNE 2026
vinkius.com/mcp/dataiku-dss

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.

Dataiku DSS MCP

14 tools available
Cloud-hosted on Vinkius

Need to manage collaborative data science work in a natural way? This MCP lets you talk to your Dataiku DSS instance like it’s an extension of your own brain. Instead of navigating dozens of tabs and clicking through build logs, you just ask your AI agent for what you need—whether that's listing all available projects or checking the precise schema of a raw dataset. You get immediate status updates on pipeline jobs, monitor training runs, and even trigger automation scenarios to rebuild pipelines when something breaks. It’s full command-line control over data science workflows, accessed via natural language conversation. When you connect this MCP through Vinkius, your agent gets access to the entire catalog of tools needed to manage everything from model metadata to underlying data connections.

Core Capabilities

01 — Discover and map projects

List all accessible DSS projects and retrieve detailed structural information about their datasets.

02 — Audit dataset schemas

Get the column names, data types, and full structure for any specified dataset in a project.

03 — Monitor pipeline execution

Track build tasks, training runs, and job status by listing pipeline jobs and analyzing their current state or timing.

04 — Verify data transformation logic

Retrieve the exact configuration structure for recipes—whether they use Python, SQL, or visual tools—to audit data flow.

05 — Control automation scenarios

List available automation scenarios and trigger their execution to securely rebuild pipelines or retrain models.

06 — Review deployed ML models

Identify saved machine learning models and retrieve detailed performance metrics, including the specific trained schema layers.

07 — Audit system connections

List all installed plugins and data source connections (like cloud storage or APIs) to verify organizational access rights.

One Click on Vinkius — From Prompt to Execution

Available at **vinkius.com/mcp/dataiku-dss** — connect your AI agent in three steps.

01 Subscribe to this MCP and provide your Dataiku Instance URL along with a valid API key (Personal, Project, or Global).

02 Your AI agent connects to the service endpoint managed by Vinkius.

03 You then use natural conversation to execute data science commands, like asking the system to list all projects in an environment.

The bottom line is that your AI client acts as a single command center for every aspect of your Dataiku data science workflow.

Built For

This MCP is built for the core team running an enterprise data platform. If you're a Data Scientist spending too much time clicking between tabs to check job status or verify data schemas, this saves your day. It's designed for people who need deep control over complex data pipelines and ML models.

Data Engineer

You track pipeline jobs and use natural language to verify recipe configurations across different transformation types.

MLOps Team Lead

You trigger automation scenarios and monitor deployed models in real-time, ensuring continuous deployment integrity.

Data Scientist

You check dataset schemas and compare model performance metrics without having to leave your primary research flow.

Analytics Manager

You audit project metadata and review data connections across the entire organization to maintain governance.

What Changes When You Connect

01 Instead of manually checking job status, you simply ask your agent to list jobs and get the current execution state or timing.

02 You can audit data logic by asking for the explicit configuration structures of recipes (Python, SQL, Visual), verifying transformations instantly.

03 Triggering pipelines used to require CLI commands; now you just tell your agent to run a scenario, like rebuilding datasets or retraining models.

04 Model performance review is faster. You get detailed metrics and schema layers for saved ML models without needing to open the DSS UI.

05 System oversight gets simple. You can list all data connections and installed plugins to quickly verify organizational constraints.

06 You gain full visibility into your entire data graph, from listing projects to checking dataset schemas in one conversational flow.

Real-World Applications

Investigating why a model failed

The MLOps team notices the 'Fraud-Detection' model score dropped. They ask their agent to get the model details, check performance metrics, and then use dataset_schema on the source data to see if the input structure changed.

Verifying a complex ETL job

A Data Engineer needs to confirm that an old sales forecasting pipeline ran correctly. They list the jobs, check the status of the last run using get_job, and then use get_recipe on the transformation recipe to audit the exact SQL logic used.

Setting up a new environment

An Analytics Manager needs an overview. They list all projects available, check which plugins are installed via list_plugins, and verify if cloud storage connections are properly listed using list_connections.

Resuming interrupted data flow

A Data Scientist is working on a new segmentation project. They notice the build tasks failed due to bad source data. They use run_scenario to trigger the pipeline rebuild and then check dataset_schema to confirm the raw input columns are correct.

Patterns to Avoid

Treating it like a simple file listing

✗ AVOID

Asking your agent to 'show me data for the project' will just list projects, but you won't get the schemas or job statuses.

✓ INSTEAD

To understand what's inside a dataset, always use the `dataset_schema` tool after running `list_datasets`. If you need status updates, start with `list_jobs`.

Forgetting context dependencies

✗ AVOID

Trying to check model performance without knowing which project or dataset it relates to. The agent will fail because the scope is missing.

✓ INSTEAD

Always use `get_project` first to confirm your current working scope, then proceed with listing models (`list_models`) and checking their metrics using `get_model`.

Over-relying on manual auditing

✗ AVOID

Manually comparing the recipe configuration against what is actually running in production. This takes hours of clicking.

✓ INSTEAD

Use `list_recipes` to see all available transformations, and then use `get_recipe` to pull the precise YAML/JSON structure directly into your chat for instant comparison.

The Right Fit

Use this MCP if you need deep, programmatic control over data pipeline governance. You're managing complex, production-grade data science workflows and need an agent to monitor job states (`get_job`), audit recipes (`get_recipe`), or trigger critical automation steps (`run_scenario`). Don’t use it if your goal is simple data lookup; for that, a standard database query tool will be faster. If

you just want to view documentation, look for a knowledge base MCP instead. This is for operational control: knowing what ran, why it failed, and how to fix it.

Dataiku DSS MCP: Auditing Data Pipelines and Recipes

Right now, auditing a data pipeline means jumping between the DSS UI, downloading logs, cross-referencing Python scripts, and manually comparing schema versions across multiple tabs. It's slow, error-prone detective work.

With this MCP, you simply ask your agent to audit the transformation logic. You can use `list_recipes` to see every workflow available and then drill down with `get_recipe` to pull the precise configuration structure into a readable chat format. The painful clicking stops; you get actionable data instantly.

Dataiku DSS MCP: Monitoring ML Model Performance

Previously, checking model health required navigating to the 'Models' tab, finding the specific deployment, and then running a separate performance report. It was a multi-step process just for a status update.

Now, you can ask your agent to list models (`list_models`) and immediately request detailed metrics using `get_model` . You get the algorithm used, the trained schema layers, and key performance indicators right in the chat—no dashboard navigation required.

14 Tools for Data Science Workflow Management

Use these tools to control every aspect of your DSS environment, from listing projects to triggering complex data transformations.

#	TOOL	DESCRIPTION
01	`list_projects`	Lists all DSS projects that your API key has access to.
02	`get_project`	Retrieves metadata, settings, and tags for a specific Dataiku project.
03	`list_datasets`	Lists every dataset contained within a specified project.
04	`dataset_schema`	Provides the complete column names and data types for any given dataset schema.
05	`list_recipes`	Lists all defined recipes, which are your data transformation workflows, in a project.
06	`list_jobs`	Shows all pipeline jobs associated with a project, covering build tasks and model training runs.
07	`get_job`	Gets the current status, timing data, and outputs for a specific job run.
08	`list_scenarios`	Retrieves a list of available automation scenarios within a project.
09	`list_models`	Lists all machine learning models that have been saved or deployed in the project.
10	`get_model`	Retrieves metadata, algorithm details, and performance metrics for a specific ML model.
11	`run_scenario`	Triggers an automation scenario execution, which can rebuild pipelines or retrain models.
12	`list_plugins`	Lists all DSS plugins that have been installed in the environment.
13	`list_connections`	Shows a list of data connections, including configured databases, cloud storage accounts, or APIs.
14	`get_recipe`	Retrieves the full configuration and settings for a specific data transformation recipe.

See It in Action

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

U List all my active projects and tell me which one has the 'Customer' tag.

Found 4 Projects:

- `Churn-Prediction-Q1` : (Tags: Customer, High Priority)
- `Sales-Forecasting` : (Tags: Sales, Quarterly)
- `Fraud-Detection-Live` : (Tags: Security, Live)
- `Staging-Sandbox` : (No tags)

U Check the schema for the raw user logs dataset in the Fraud project.

Schema retrieved successfully:

COLUMN NAME	DATA TYPE	DESCRIPTION
`timestamp`	Date	Time of action completion
`user_id`	String	Unique ID for user
`ip_address`	String	IP address used in transaction
`action_type`	String	e.g., login, purchase, view

Total columns: 4

U I need to trigger the pipeline rebuild for my sales data.

Attempting to run scenario 'REBUILD_PIPELINE' in project 'SALES-FORECASTING'.

Status: Triggered.
Details: The system is now rebuilding underlying datasets and recipes. This process can take 15–20 minutes depending on data volume. I will notify you when the run is complete.

Frequently Asked Questions

01 **How do I check if my dataiku projects are connected to external databases?**

The MCP allows you to list all data connections and installed plugins. This lets you quickly audit your entire environment by seeing which cloud storage, APIs, or SQL databases are linked to your DSS instance.

02 **Can this MCP help me monitor if a model is performing well?**

Yes, you can list saved machine learning models and then request detailed performance metrics. This helps data scientists compare schema layers and track changes in prediction quality directly through conversation.

03 **Does Dataiku DSS MCP let me run manual pipeline jobs?**

Absolutely. You can use the tools to list all available pipeline jobs, check their status using `get_job`, and even trigger a full rebuild or retraining cycle via automation scenarios.

04 **What if I need to audit the SQL logic in my data transformations?**

You can retrieve recipes by listing them first, then using the specific tool to pull the explicit configuration structure. This allows you to verify exact Python or SQL code without opening the DSS interface.

05 **How do I find out what projects I have access to?**

You simply ask your agent to list all accessible DSS projects. It provides a comprehensive overview, including project metadata and tags, so you know exactly what resources are available for your team.

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 `"dataiku-dss": { "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

Dataiku DSS 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

DOCUMENT INFORMATION

Generated	June 2026
MCP Server	Dataiku DSS MCP
Server ID	`019d7582-315c-7179-a27e-efc75014bf8f`
Platform	Vinkius Cloud for AI Agents
Endpoint	`https://edge.vinkius.com/{token}/mcp`

LICENSE & USAGE