MCP SERVER | NO CODE | CLOUD HOSTED

Collection Catalog Deduplicator MCP

Clean up messy inventory and fix duplicate catalog entries.

Collection Catalog Deduplicator MCP helps you maintain clean, accurate data across your inventories. Your AI client uses this MCP to find redundant entries, merge them into single master records, and check the overall health of your catalog. It handles the heavy lifting of identifying duplicates via ISBN, UPC, or descriptive attributes so your data stays reliable.

A+ Quality Score 100/100

deduplication | collection | data-integrity | inventory-management | cataloging

MAINTAINED BY VINKIUS CLOUD FOR AI AGENTS · OCTOBER 2026
vinkius.com/en/ai-agent-connect/collection-catalog-deduplicator

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.

Collection Catalog Deduplicator MCP

4 tools available
Cloud-hosted on Vinkius

Managing large collections often leads to messy, redundant data. You might end up with three different entries for the same book or two separate records for the same piece of equipment. This MCP gives your AI agent the ability to scan your datasets, spot these overlaps, and fix them without you having to manually hunt through spreadsheets.

You can use it to group identical items by matching identifiers like ISBN or UPC, or by comparing titles and editions. Once the duplicates are found, your agent can collapse them into a single, high-fidelity master record. It also provides a way to check the statistical health of your entire collection, giving you a clear picture of how much redundant data is cluttering your system. Whether you are managing a library, a retail inventory, or a private collection, this tool ensures your records remain accurate and easy to navigate.

Core Capabilities

01 — Duplicate Detection

Your agent identifies redundant entries using ISBN, UPC, or title matches.

02 — Record Merging

Your agent combines multiple entries into one accurate master record.

03 — Data Health Audits

Your agent calculates statistical reports on your catalog's integrity.

04 — Owner Filtering

Your agent pulls specific records belonging to a single owner.

One Click on Vinkius — From Prompt to Execution

Available at **vinkius.com/en/ai-agent-connect/collection-catalog-deduplicator** — connect your AI agent in three steps.

01 Connect the MCP to your preferred client like Claude or Cursor via Vinkius.

02 Ask your agent to scan the collection for duplicates.

03 Review the groups of identical items identified by the agent.

04 Instruct the agent to merge the duplicates into master records.

05 Run an integrity report to confirm the cleanup was successful.

Connect the MCP to your AI client to start cleaning your data immediately.

Built For

This MCP is built for anyone managing large, item-based datasets that require high accuracy.

Inventory Managers

They hand off the task of cleaning up duplicate stock entries to their AI agent.

Librarians

They use the tool to ensure catalog records for books and media are unique.

Data Analysts

They use the integrity reports to assess the quality of collection datasets.

What Changes When You Connect

01 Reduces data redundancy by merging duplicate entries.

02 Identifies errors using specific identifiers like ISBN and UPC.

03	Provides statistical health checks for entire catalogs.
04	Simplifies ownership tracking within large datasets.

Real-World Applications

Library Catalog Cleanup

An agent finds multiple entries for the same book title and merges them into one record.

Retail Inventory Audit

A manager uses the integrity tool to see how many duplicate SKUs exist in the system.

Private Collection Management

A collector filters their database to see only items owned by a specific person.

Database Integrity Checks

A user runs a report to find the duplicate ratio across a massive dataset.

Patterns to Avoid

Manual record hunting

AVOID

You spend hours scrolling through a massive spreadsheet to find where two different entries for the same book exist.

✓ INSTEAD

Let your agent scan the dataset for overlaps using `find_duplicates` to group identical items automatically.

Ignoring data health

AVOID

You assume your inventory is accurate without actually knowing how many duplicate SKUs are cluttering the system.

✓ INSTEAD

Run `calculate_catalog_integrity` to get a statistical report on your redundancy levels and overall data health.

Partial data cleanup

AVOID

You find a duplicate but leave the two separate records active, which keeps your counts wrong.

✓ INSTEAD

Use `merge_records` to collapse those multiple entries into one single, accurate master record.

The Right Fit

Connect this MCP if you manage large, item-based datasets like libraries or retail inventories where accuracy is non-negotiable. Do not connect it if your data is small enough to manage manually or if your records don't share common identifiers like ISBN or UPC. The decision comes down to whether redundant entries are actively corrupting your data integrity.

Collection Catalog Deduplicator Tools

These tools give your agent the ability to scan, audit, and clean up redundant entries across your entire dataset.

#	TOOL	DESCRIPTION
01	`find_duplicates`	This tool scans your records to group identical items based on matching criteria.
02	`calculate_catalog_integrity`	This tool provides a statistical overview of your collection's health and redundancy levels.
03	`filter_by_owner`	This tool isolates all records belonging to a specific person within the dataset.
04	`merge_records`	This tool collapses a group of duplicate records into one single master record.

See It in Action

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

U Find all duplicate records in this list: [{'id': '1', 'title': 'The Great Gatsby', 'edition': '1st', 'identifier': '12345'}, {'id': '2', 'title': 'The Great Gatsby', 'edition': '1st', 'identifier': '12345'}]

I found 1 group of duplicates containing 2 records.

U What is the integrity report for my collection?

Your collection has 100 total entries, 5 duplicates, and 95 unique items, resulting in a duplicate ratio of 0.05.

U Show me all items owned by Alice.

Alice owns 3 items: 'Vintage Camera', 'Leather Journal', and 'Fountain Pen'.

Frequently Asked Questions

01 **How does the MCP find duplicates?**

The MCP identifies duplicates by matching specific identifiers like ISBN or UPC, or by comparing descriptive attributes like titles and editions.

02 **Can I use this with Claude or Cursor?**

Yes, you can connect this MCP to any MCP-compatible client including Claude, Cursor, and Windsurf.

03 **What happens when I merge records?**

The merge_records tool takes a group of duplicate entries and collapses them into a single, high-fidelity master record.

04 **How do I check if my catalog is clean?**

You can use the calculate_catalog_integrity tool to get a statistical overview of your collection's health.

05 **Do I need to host the MCP myself?**

No, Vinkius hosts and manages the MCP for you. You just connect your client and start working.

06 **How does the tool identify duplicates?**

The `find_duplicates` tool identifies matches if the identifier (like an ISBN) is identical, or if both the title and edition match exactly when no identifier is present.

07 **Can I merge multiple records into one?**

Yes, you can use `merge_records` to collapse a group of duplicate IDs into a single master record that preserves the most complete metadata.

08 **How can I check the health of my collection?**

You can use `calculate_catalog_integrity` to get a statistical report including total entries, duplicate counts, and the duplicate ratio.

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 `"collection-catalog-deduplicator": { "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

Collection Catalog Deduplicator 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 Collection Catalog Deduplicator. 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	October 2026
MCP Server	Collection Catalog Deduplicator MCP
Server ID	`01a0f6a2-f698-7369-ac09-0d995a2bc46a`
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/en/ai-agent-connect/collection-catalog-deduplicator.

© 2026 Vinkius · vinkius.com · support@vinkius.com