MCP SERVER | NO CODE | CLOUD HOSTED

Data.gov MCP for AI Agents

Analyze and retrieve US federal open datasets by topic and keyword

Data.gov connects your AI client directly to over 300,000 open datasets from US federal agencies. You can search by topic, find specific organizations like NASA or USDA, and pull detailed metadata on everything from climate models to student performance data. It's the single source for public records research.

A+ Quality Score 98.33/100

open-data | federal-datasets | public-records | data-catalog | research-data

MAINTAINED BY VINKIUS CLOUD FOR AI AGENTS · JUNE 2026
vinkius.com/mcp/datagov

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.

Data.gov MCP

13 tools available
Cloud-hosted on Vinkius

Need to dig into government data? This MCP lets your AI agent converse with the entire Data.gov catalog. You don’t need to mess with API keys or write complex queries just to get started. Instead, you talk naturally about what you're looking for—say, 'Give me all datasets on water quality in Florida.' The system handles the search across hundreds of agencies and topics.

It provides structure by allowing you to explore data groups like 'Health' or 'Finance,' or drill down into specific agencies like the Census Bureau. If your research requires understanding what kind of files are available, it shows formats ranging from JSON to CSV. Connecting this through Vinkius gives your AI client access to all these resources in one place, letting you spend time analyzing data instead of building connections.

Core Capabilities

01 — Search for datasets across multiple topics

Find specific public datasets by searching keywords and filtering results by organization or tags.

02 — Identify relevant government organizations

List federal agencies, like the EPA or NOAA, to see what kind of data they publish.

03 — Browse datasets by general topic area

Explore pre-grouped topics such as agriculture, climate, or public safety to narrow your focus quickly.

04 — Get detailed information on a dataset

Retrieve the full metadata, license details, and resource download links for any specific data set ID.

05 — Check available data formats

List all possible file types (CSV, JSON, XML) so you know what format your final dataset will be in.

One Click on Vinkius — From Prompt to Execution

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

01 Subscribe to this MCP on Vinkius and connect it to your AI client. No API key is needed because all the data is public domain.

02 Ask your agent a question like, 'Show me all datasets related to education spending.'

03 The agent executes the necessary search across the catalog and returns structured results with titles, sources, and download options.

The bottom line is you get conversational access to massive, complex public data without writing a single API call.

Built For

Anyone who works with large-scale information—academic researchers needing source material, journalists investigating policy gaps, or developers building applications that rely on official government statistics. If your job involves public records, this MCP is for you.

Academic Researcher

Needs to find and cross-reference multiple federal datasets—say, correlating EPA climate data with USDA agricultural yields—for a dissertation.

Investigative Journalist

Must pull public records on everything from census demographics to local health department reports for an exposé piece.

Data Scientist

Builds visualization dashboards that require clean, structured data inputs from multiple government sources (e.g., NOAA and Census Bureau).

What Changes When You Connect

01 Search over 300,000 records instantly: Use `search_datasets` to find precise data sets across agriculture, health, or climate without knowing the exact dataset ID.

02 Understand where the data comes from: You can list all organizations using `list_organizations`, letting you trace a topic back to its original federal source (e.g., NASA).

03 Explore by theme, not just keywords: Use `get_group_datasets` or `list_groups` to browse entire domains like 'Education' and see related data sets.

04 Know the file format upfront: Before running a query, use `list_resource_formats` to confirm if your data is available as clean JSON or CSV for immediate use in code.

05 Get complete context on any dataset: Running `get_dataset` gives you all the metadata—the license, the date it was last updated, and what fields are included.

Real-World Applications

Correlating climate trends with agricultural yields

A researcher needs to compare NOAA's historical temperature data with USDA crop reports. They ask their agent to search for both datasets, using `search_datasets` and then getting the details via `get_dataset` for reliable download links.

Verifying a news report on public safety statistics

A journalist needs hard data. They ask their agent to look up all available datasets tagged 'public-safety' using `get_tag_datasets`, then use `get_organization` to see which federal agency published the most recent information.

Building a dashboard on educational spending

A developer needs structured data for a new app. They ask their agent to list all groups, find 'education', then use `get_group_datasets` to pull titles and download links from relevant departments.

Mapping US environmental policy changes

A student needs to understand the scope of federal resources. They ask their agent to list all available resource formats using `list_resource_formats`, then use `search_datasets` filtered by 'EPA' and 'water quality'.

Patterns to Avoid

Treating it like a simple keyword search

✗ AVOID

Asking the agent, 'Show me everything about climate.' This results in hundreds of generic links because you didn't narrow down the scope or source.

✓ INSTEAD

Instead, prompt: 'Using Data.gov, find datasets related to global warming published by NOAA and filter them only for JSON format.' This forces the use of `search_datasets` combined with specific filtering capabilities.

Assuming data is ready for direct coding

✗ AVOID

Just getting a list of titles from an organization's page without checking the license. You might download a dataset that has restricted usage rights.

✓ INSTEAD

Always ask the agent to run `get_dataset` first. This verifies the license information and gives you the full metadata before you start writing code.

Ignoring the topic structure

✗ AVOID

Searching vaguely for 'health' without checking if it falls under a specific group or organization, leading to irrelevant results.

✓ INSTEAD

First, use `list_groups` to see the main categories. Then, refine your search by calling `get_group_datasets` for 'Health' before running a general keyword search.

The Right Fit

Use this MCP if your project requires comprehensive access to official, non-commercial US government data across diverse topics like finance, health, and environment. It excels when you need to validate information against primary federal sources. Don't use it if you are looking for niche local or state-level records; Data.gov focuses on the national level.

Don't use this MCP if your only goal is basic web scraping of a single website. For that, a general web crawling tool is better. Instead, focus on using its specific tools like `get_tag` and `list_tags` to understand the *structure* of the data—this helps you build a much more precise query for your agent.

Data.gov MCP: Finding US Public Records with Federal Datasets

Today, finding comprehensive public records is a massive time sink. Researchers and journalists spend hours navigating dozens of agency websites, manually checking different databases for the same topic. It's an endless cycle of logging in, clicking through tabs, and copy-pasting links just to verify if the data exists.

With this MCP, you simply ask your agent what you need—like 'All records on renewable energy.' The system handles the cross-referencing across multiple federal sources and provides a consolidated list of available datasets and their download options. You get instant, structured access.

Data.gov MCP: Understanding Data Structure with Topic Groups

Manually identifying which specific agency holds the data you need—is it the Census Bureau or the Department of Commerce? You have to check each one individually, and they often use different naming conventions for similar topics.

This MCP solves that by allowing you to browse high-level topic groups using `list_groups`. It shows you all related datasets under 'Finance' or 'Health,' grouping them logically so you know exactly where to look next. You don't have to guess.

Data.gov: All 13 Tools for Federal Dataset Research

These tools let your agent search the catalog structure, retrieve metadata on specific groups or organizations, and perform deep data queries across federal sources.

#	TOOL	DESCRIPTION
01	`get_dataset`	Retrieves full metadata for a specific dataset ID or name, including licenses and download options.
02	`get_group`	Gets details about broad topic groups like 'climate' or 'health', listing associated datasets.
03	`get_organization`	Provides information on a specific federal agency, including its contact info and published data sets.
04	`get_organization_datasets`	Lists all available datasets published by one particular organization you specify.
05	`get_status`	Checks the overall API status of Data.gov, providing total counts for groups, organizations, and tags.
06	`get_tag`	Returns details about a specific tag (like 'public-safety'), along with how many datasets use it.
07	`get_tag_datasets`	Retrieves all dataset titles and download links associated with a given descriptive tag.
08	`get_group_datasets`	Lists all datasets that fall under a specific topic group, such as 'education'.
09	`list_groups`	Returns a list of every major topic group available on the site (e.g., finance, climate).
10	`list_organizations`	Lists all federal agencies that have contributed data to Data.gov.
11	`list_resource_formats`	Displays a list of all possible file formats (JSON, CSV, etc.) available for download.
12	`list_tags`	Shows every active tag used across the catalog and how many datasets are associated with it.
13	`search_datasets`	Performs a comprehensive search using free text, filtering by organization or group to find specific public records.

See It in Action

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

U Search for climate change data from NOAA.

Climate Data Report: US Federal Sources

• **NOAA Climate Data:** Comprehensive temperature and sea level readings. (Format: JSON, Updated: 2024-05-15)
• **EPA Greenhouse Gas Emissions:** Yearly emissions by state. (Format: CSV, Updated: 2023-01-01)

You can download both datasets directly via the provided links for immediate analysis.

U What education records are available?

📚 Education Datasets Found:

The following data sets cover student performance and spending:

Dataset Name	Organization	Key Topics	Format
NAEP Assessment Data	Dept. of Ed.	Student Grades, Spending	CSV
College Scorecard	Dept. of Ed.	Enrollment, Cost	JSON

Start by reviewing the full metadata for any title above.

U Show me all datasets related to water quality in Florida.

💧 **Water Quality Datasets (Florida Focus):**

We found 8 relevant records from various sources:

- **State Pollutant Levels:** EPA, measuring pH and turbidity. (Latest: 2024-06)
- **River Basin Flow Rates:** USGS, tracking flow volume. (Format: XML)

Please specify if you need data for a certain year to narrow down the search.

Frequently Asked Questions

01 **Do I need an API key?**

No! Data.gov data is public domain and freely accessible. No authentication required.

02 **How many datasets are available?**

Data.gov catalogs 300,000+ datasets from over 200 federal agencies including NASA, USDA, EPA, NOAA, Department of Education, Census Bureau and many more.

03 **What organizations publish data?**

Over 200 federal agencies including NASA, USDA, EPA, NOAA, Department of Education, Census Bureau, Department of Transportation, FBI, CDC, FDA and many more.

04 **What formats are available?**

Common formats: CSV, JSON, XML, Shapefile, GeoJSON, PDF, HTML, RDF, KML, ZIP. Use list_resource_formats to see all available formats.

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 `"datagov": { "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

Data.gov 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 Data.gov. 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	Data.gov MCP
Server ID	`019d842d-24cf-7357-ae36-96b9618fc2a7`
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/datagov.