

Open in app ↗

≡ Medium

🔍 Search

✍ Write



★ Member-only story

# 13 MCP Servers for Data Wranglers (To Turn Data to Gold)



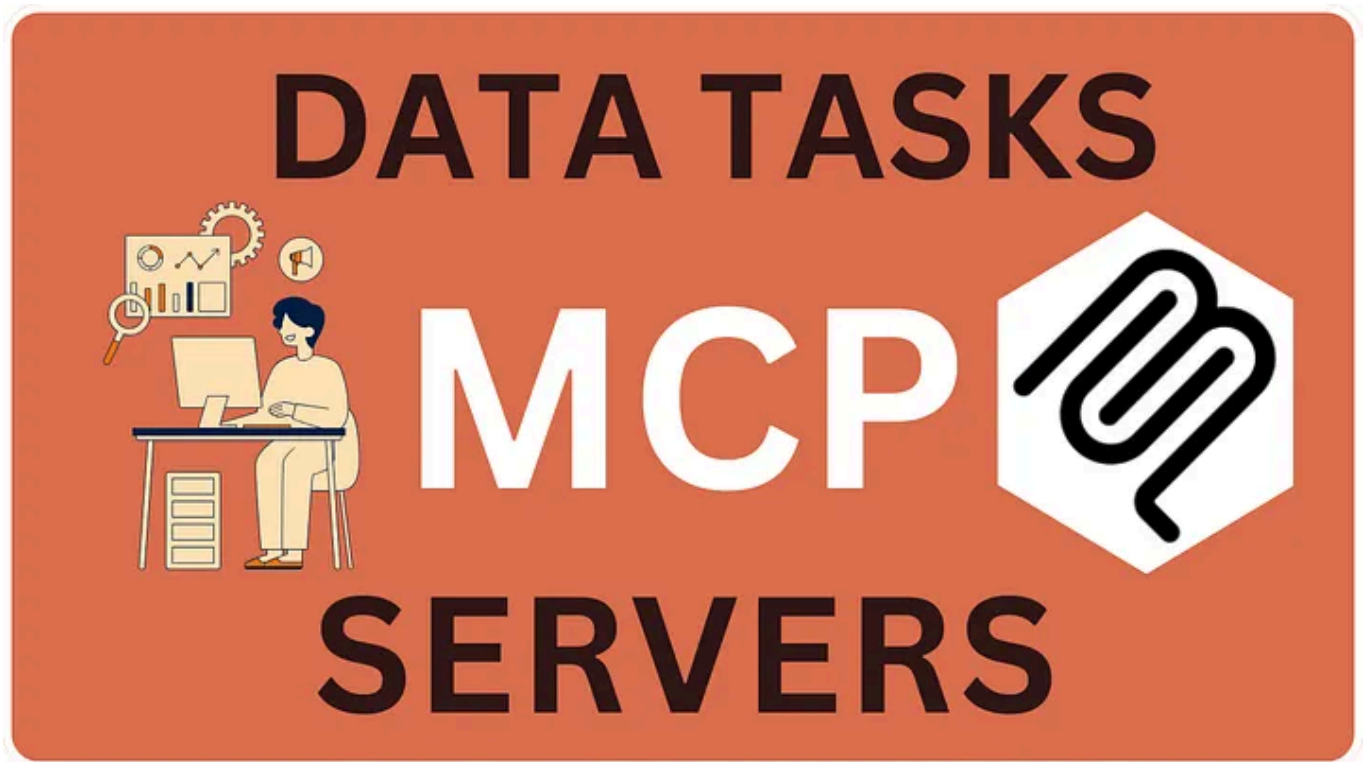
Joe Njenga

Follow

9 min read · Jun 29, 2025

👏 317

💬 2



MCP Servers for Data Wranglers — Featured Image

Fast data processing is one of the best use cases of MCP.

If you're a data wrangler, most of your time is wasted moving data around, cleaning it, and getting it into a format that makes sense.

MCP servers can now help you tap into your data sources and LLMs to analyse massive amounts of data within the shortest time possible.

Whether you're dealing with messy CSV files, complex APIs, or data scattered across multiple platforms, *these servers eliminate the manual work that kills your productivity.*

Recently, I worked on a project using Bright Data MCP server and discovered the immediate potential for data wranglers who use MCP.

I've tested dozens of MCP servers over the past few months.

Most revolve around coding and data, but here's the reality: many are either too basic for real-world scenarios or too complicated to implement quickly.

What you need are servers that work out of the box and handle the data challenges you face.

This list contains the 13 MCP servers that are suitable for a wide range of data processing tasks.

## **Web Data Collection MCPs**



Web Data Collection MCPs Illustration / By Author

## 1. Bright Data MCP

Bright Data MCP handles the heavy lifting when you need data from websites that don't want to share it easily.

*This server connects you to Bright Data's proxy network, so you can scrape sites without getting blocked or throttled. I've used it to pull product data from e-commerce sites and social media metrics that would normally require hours of manual work.*

The setup takes about five minutes, and then you're pulling structured data from almost any website. If dealing with CAPTCHAs or IP bans that slow your data collection workflows, you should try this MCP.

### Key Features

- Residential proxy network with millions of IPs
- Built-in CAPTCHA solving and bot detection bypass
- JSON output format ready for analysis
- Rate limiting controls to avoid getting banned
- Works with dynamic JavaScript-heavy sites
- Supports bulk data collection tasks

Bright Data MCP makes web scraping very easy, as just chatting with your LLM.

**Git Link:** [Bright Data MCP](#)

## **2. Crawl4AI RAG MCP Server**

Crawl4AI RAG takes a smarter approach to web data collection by understanding content context before extracting it.

*Unlike basic scrapers that grab everything, this server analyzes page structure and pulls only relevant information. I've found it particularly useful when researching competitors or gathering market intelligence from multiple sources.*

Its RAG component means it can answer questions about the content it crawls, turning raw web data into actionable insights without additional processing steps.

## Key Features

- Context-aware content extraction
- Built-in RAG capabilities for instant analysis
- Handles multiple page formats automatically
- Content summarization and question answering
- Batch processing for large-scale research
- Filters out ads and irrelevant content

The advantage here is that the scraped content you get is meaningful data already processed.

**Git Link:** [Crawl4AI RAG MCP Server](#)

## 3. Firecrawl MCP Server

Firecrawl MCP specializes in turning messy websites into clean, structured data that you can use.

*This server excels at handling modern web applications with complex JavaScript and dynamic content loading. I've used it to extract data from SaaS dashboards and internal tools that traditional scrapers can't touch.*

The output comes formatted as clean markdown or JSON, so you don't spend time cleaning up the extracted content before analysis.

## Key Features

- JavaScript rendering for dynamic sites
- Clean markdown and JSON output formats
- Handles authentication and login flows
- Extracts images and media files
- Respects robots.txt and rate limits
- API-first design for easy integration

Firecrawl MCP is a tool you want to use when other scrapers fail you.

Git Link: [Firecrawl MCP Server](#)

## **Database Connection MCPs**



Database Connection MCPs Illustration / By Author

## 4. Universal Database MCP Server

Universal Database MCP connects to almost any database without making you remember different syntax for each one.

*This server speaks PostgreSQL, MySQL, SQLite, and Oracle through a single interface. I've used it to pull reports from legacy systems and modern databases using the same commands, which saves hours of context switching.*

The query builder handles complex joins and aggregations, so you can focus on getting insights.

### Key Features

- Supports PostgreSQL, MySQL, SQLite, Oracle, and more

- Unified query interface across all database types
- Built-in query optimization and caching
- Visual query builder for complex operations
- Connection pooling for better performance
- Export results to CSV, JSON, or Excel formats

Universal Database MCP eliminates the need to learn multiple database dialects. One server, one syntax, all your data sources connected.

**Git Link:** [Universal Database MCP Server](#)

## **5. Multi-Database MCP Server**

Multi-Database MCP takes a different approach by letting you query multiple databases simultaneously in a single operation.

*Instead of connecting to databases one at a time, this server creates a unified view across your entire data infrastructure. I've found it invaluable when building dashboards that need data from sales, marketing, and operations databases.*

The cross-database joins work seamlessly, and you can correlate data that is stored in separate systems without complex ETL processes.

### **Key Features**

- Simultaneous connections to multiple databases
- Cross-database queries and joins



- Real-time data federation
- Schema discovery and mapping
- Automatic data type conversion
- Transaction support across different sources

This MCP gives you a single view of your entire data in a unified access.

**Git Link:** [Multi-Database MCP Server](#)

## **6. MotherDuck & DuckDB MCP Server**

MotherDuck MCP brings analytical database power directly into your Claude conversations without server setup headaches.

*DuckDB runs analytics queries faster than traditional databases, especially on large datasets. You can use it to crunch through millions of rows of transaction data that would burn out other systems.*

The cloud integration through MotherDuck means you get enterprise-grade performance without managing infrastructure or worrying about scaling issues.

### **Key Features**

- Lightning-fast analytical queries on large datasets
- Cloud-native through MotherDuck integration
- Columnar storage for optimal performance

- Built-in support for Parquet, CSV, and JSON files
- SQL-compatible with advanced analytics functions
- Automatic scaling based on query complexity

MotherDuck MCP turns complex analytical queries into simple conversations.

**Git Link:** [MotherDuck & DuckDB MCP Server](#)

## **7. MCP Server for MySQL**

MySQL MCP Server focuses exclusively on MySQL databases and does better than generic database connectors.

*This server understands MySQL's specific features like stored procedures, triggers, and advanced indexing strategies. I tested it by optimizing a query performance on a large MySQL installation.*

The MySQL-specific optimizations mean queries run faster, and you can access features that other database connectors ignore.

### **Key Features**

- MySQL-specific optimizations and features
- Support for stored procedures and triggers
- Advanced indexing and query plan analysis

- Replication status monitoring
- Binary log parsing for change tracking
- Performance schema integration

When you're working exclusively with MySQL, this specialized server beats generic solutions every time.

Git Link: [MCP Server for MySQL](#)

## Document & Formatting MCPs



Document & Formatting MCPs Illustration / By Author

## 8. Excel MCP Server

Excel MCP Server turns spreadsheet operations into simple click-through conversations instead of endless menus and formulas.

*This server reads, writes, and manipulates Excel files without opening Excel at all. I've used it to process monthly reports that used to take hours of manual formatting and formula updates.*

It works with complex spreadsheets, making it easy to process huge amounts of data.

### Key Features

- Read and write .xlsx and .xls files directly
- Create complex formulas through natural language
- Batch process multiple spreadsheets
- Generate charts and pivot tables programmatically
- Preserve formatting and styles
- Handle large datasets without memory issues

Excel MCP Server eliminates the tedious parts of spreadsheet work.

Git Link: [Excel MCP Server](#)

## 9. Markdownify MCP Server

Markdownify MCP Server converts messy documents into clean, readable markdown format that works everywhere.

*HTML pages, Word documents, PDFs with weird formatting — this server cleans them up and outputs consistent markdown. I've used it to standardize documentation from different sources into a unified format for team wikis.*

The conversion preserves important structure like headers, lists, and links while stripping out the formatting noise that makes documents hard to read.

## **Key Features**

- Converts HTML, Word, and PDF to clean markdown
- Preserves document structure and hierarchy
- Removes formatting clutter and inconsistencies
- Handles tables and embedded images
- Batch processing for multiple documents
- Customizable output formatting rules

This MCP has the best cleanup power of saving hours of manual editing.

**Git Link:** [Markdownify MCP Server](#)

## **10. AntV Chart Generator**

AntV Chart Generator creates professional data visualizations without requiring design skills or complex charting libraries.

*This server takes your data and generates publication-ready charts using AntV's visualization engine. You can use it to create executive dashboards and client reports with a professional design outlook.*

The chart includes all types of charts from basic bar charts to complex statistical visualizations, and the output works in presentations, web pages, or printed reports.

### **Key Features**

- Professional chart templates for all data types
- Interactive visualizations with hover effects
- Multiple export formats (SVG, PNG, PDF)
- Responsive designs that work on any screen size
- Statistical chart types like box plots and heatmaps
- Custom color schemes and branding options

AntV Chart Generator bridges the gap between raw data and compelling visual stories.

Resource Link: [AntV Chart Generator](#)

## **Specialized Research Tools MCPs**



Specialized Research Tools MCPs Illustration / By Author

## 11. Financial Datasets MCP

Financial Datasets MCP connects you to real-time market data and historical financial information.

*This server pulls stock prices, earnings reports, economic indicators, and company financials through a simple interface. I tested it to build investment research reports and a market analysis app for a YouTube tutorial.*

The data comes clean and structured, so you can jump straight into analysis instead of spending time formatting and validating financial information.

### Key Features

- Real-time stock prices and market data

- Historical financial statements and ratios
- Economic indicators and macro data
- Earnings reports and analyst estimates
- Sector and industry comparisons
- Custom financial calculations and metrics

Financial Datasets MCP gives you access to professional-grade financial data right in your LLM.

**Resource Link:** [Financial Datasets MCP](#)

## **12. BrowserTools MCP**

BrowserTools MCP automates browser tasks that take long research hours.

*This server controls web browsers programmatically, and you can automate repetitive research tasks across multiple websites. I've used it to gather pricing information from competitor sites and extract data from web applications that don't offer APIs.*

The automation handles complex interactions like login flows, form submissions, and navigation sequences that would take hours to complete manually.

### **Key Features**

- Automated web browser control and navigation



- Form filling and submission automation
- Screenshot capture for documentation
- Cookie and session management
- Multi-tab workflows and parallel processing
- Custom wait conditions for dynamic content

BrowserTools MCP turns tedious web research into set-and-forget automation.

Resource Link: [BrowserTools MCP](#)

### **13. BioMCP (Biomedical Database Integration)**

BioMCP connects to major biomedical databases and research repositories, making scientific literature and clinical data accessible through simple queries.

*This server searches PubMed, clinical trial databases, and genomic repositories without requiring specialized knowledge of each database's query syntax.*

The integration spans multiple scientific databases, so you can cross-reference findings and build comprehensive research reports from authoritative medical sources.

#### **Key Features**

- PubMed and scientific literature search

- Clinical trial database integration
- Genomic and proteomic data access
- Drug interaction and pharmaceutical databases
- Medical terminology and ontology support
- Citation formatting and reference management

BioMCP opens up the world of biomedical research to non-specialists and data wranglers looking for such data

Resource Link: [BioMCP](#)

## **Final Thoughts**

These 13 MCP servers cover every stage of the data wrangling process, from collection to analysis to presentation.

If you're struggling with web scraping tasks, grab Bright Data or Firecrawl.

These MCP servers are powerful tools designed to streamline and accelerate your data analysis workflow.

It's time to turn that data into gold!

# Let's Connect!

If you are new to my content, my name is Joe Njenga

Join thousands of other software engineers, AI engineers, and solopreneurs who read my content daily on Medium and on YouTube where I review the latest AI engineering tools and trends.

If you would like to connect directly, you can reach out :

**AI Integration Software Engineer (10+ Years Experience )**

Software Engineer specializing in AI integration and automation. Expert in building AI agents, MCP servers, RAG...

njengah.com

Follow me on Medium | YouTube Channel | X | LinkedIn

- Data Analysis
- Mcp Protocol
- Mcp Server
- Data Wrangling
- Ai Data Collection



Written by Joe Njenga

4.4K followers · 99 following

Software & AI Automation Engineer, Tech Writer & Educator. Vision: Enlighten, Educate, Entertain. One story at a time. Work with me: mail.njengah@gmail.com

Follow



## Responses (2)



Bgerby

What are your thoughts?



Jes Hunsballe

Jul 31



The markdownify link is dead and the github user does not seem to have such a repo. There is another markdownify mcp server at <https://github.com/zcaceres/markdownify-mcp>



[Reply](#)



Khelan Patel

Jul 8



Great roundup, finally, someone made data wrangling sound less like digital janitorial work. At Techlusion, we've been helping teams build custom MCP-powered workflows that eliminate the mess and deliver insights [straight to the point](#). If you're... [more](#)



[Reply](#)

## More from Joe Njenga



 Joe Njenga

## I Tested Upgraded Claude Code 2.0 (And Discovered 7 New Features...

Claude Code just dropped version 2.0, and they packed it with the most requested...

★ Sep 30 🖱 359 💬 12  ⋮

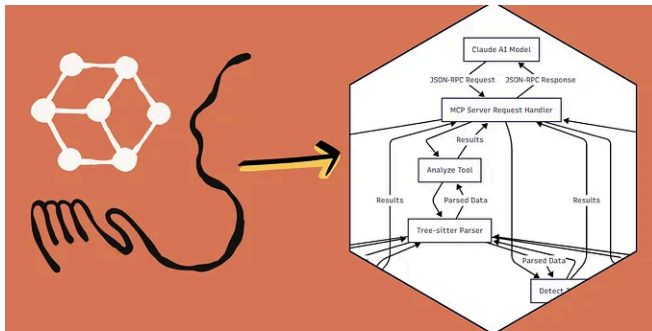


 Joe Njenga

## 12 Little-Known Claude Code Commands That Make You a Whiz

What if I told you there are some Claude Code commands that, although not very popular,...

★ Sep 21 🖱 247 💬 4  ⋮

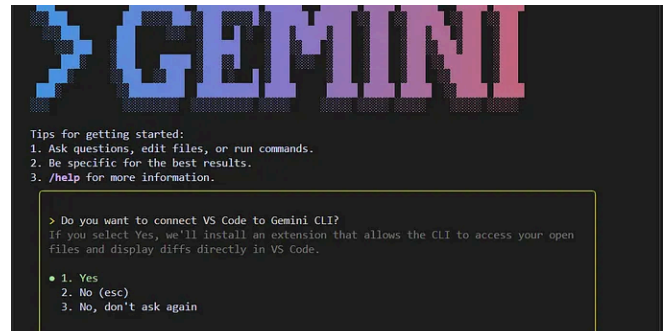


 Joe Njenga

## I Asked Claude 4.5 to Build Me a Complex MCP Server (It Was So...

If you want to build your custom MCP server, stop wasting time thinking, coding, or trying...

★ 3d ago 🖱 359 💬 6  ⋮



 Joe Njenga

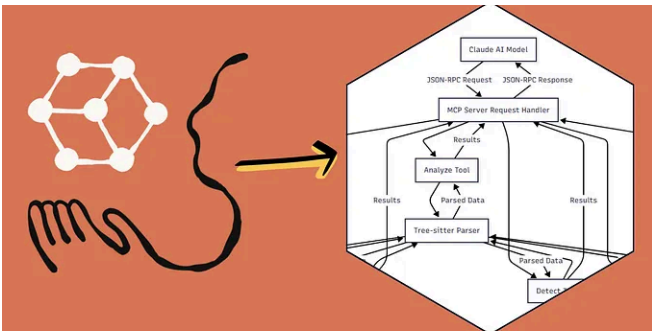
## I Tested The Updated Gemini CLI (And Found These New Features...

Gemini CLI has these newly updated features that are making it edge closer to dethroning...

★ Sep 12 🖱 158 💬 5  ⋮

See all from Joe Njenga

# Recommended from Medium



Joe Njenga

## I Asked Claude 4.5 to Build Me a Complex MCP Server (It Was So...

If you want to build your custom MCP server, stop wasting time thinking, coding, or trying...

★ 3d ago 🖱️ 359 💬 6 📖 + ⋮



Platform Engineers

## How AI Agents Cut Cloud Costs by 60%: The Platform Engineer's...

The \$482 Billion Cloud Waste Problem Nobody's Solving

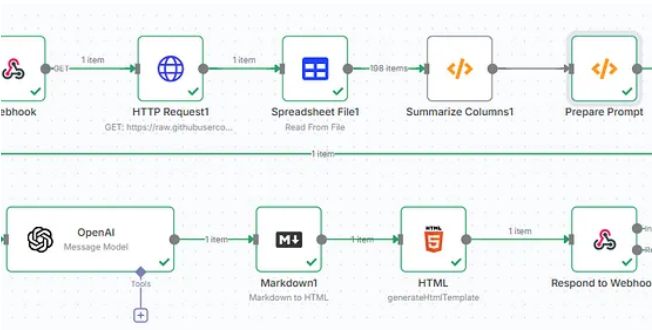


In Stackademic by Blend Visions

## 5 Best MCP Servers for Effortless Vibe Coding in 2025 🚀

The time has come for you to maximize your coding process through enhanced...

★ Apr 27 🖱️ 979 💬 12 📖 + ⋮



Malini Anbazhagan

## I Tried Building an AI Agent for My Data—Here's What I Built Instead

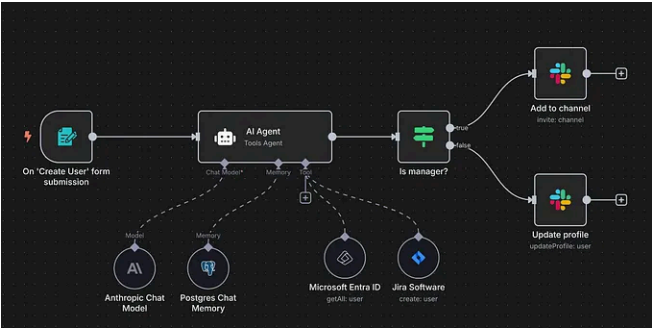
As a Data Analyst, I've always dreamed of taking my workflows to the next level — ...



 In Vibe Coding by Alex Dunlop

## This Just Became the Most Important Tool in Frontend (Here'...

The goldmine I've always wanted solved



 In Coding Nexus by Code Coup

## 10 n8n Tricks Beginners Wish They Knew Earlier

Everyone loves saying n8n is beginner-friendly. Honestly? Not really.

See more recommendations