

Model Context Protocol - Server Implementation

Prabhu Kiran Avula

Project Report

August 9th, 2025

License: MIT LICENSE

Purpose

This project implements a custom Model Context Protocol (MCP) server designed to expand the Klavis AI ecosystem with a suite of atomic, AI-usuable tools.

While the original requirement was to build an MCP server for a single public API, I chose to extend the scope to create a multi-service content generation and research assistant, demonstrating the adaptability, scalability, and product vision expected from a founding engineer.

The implemented MCP server integrates **four distinct APIs** (**Replicate**, **Tavily**, **SerpAPI**, and **ElevenLabs**), each providing a clear, single-purpose function:

- **Replicate** – High-quality image generation from text prompts.
- **Tavily** – AI-driven deep web search and webpage summarization.
- **SerpAPI** – Fast, structured web search results for general queries.
- **ElevenLabs** – Natural-sounding text-to-speech audio generation.

By combining these services, the server enables workflows such as:

Search the web for a topic → Summarize a source, generating a script → Generate a voice narration → Create an image or visual asset to accompany it.

This approach reflects Klavis AI's **General Guidelines** principles:

- **Atomicity** – Each tool performs one explicit, well-defined action.
- **Clarity** – Tool names and descriptions are explicit for optimal LLM invocation.
- **Extensibility** – Modular architecture allows additional tools or APIs to be integrated with minimal friction.
- **Robustness** – Built-in error handling, input validation, and environment-based authentication ensure consistent reliability.

The MCP server was developed and tested primarily in **Cursor IDE (Pro)** for efficient LLM-assisted coding and debugging. Still, it is fully compatible with other MCP-capable LLM agents such as Claude Desktop.

Throughout development, I tested the system extensively on:

- Breaking news and political updates
- AI research advancements
- Creative and abstract art prompts

The **Outputs section** contains complete end-to-end examples of tool invocation sequences, logs, and generated content, confirming both functional correctness and alignment with Klavis AI's open-source contribution expectations.

Tools Implemented

Tool	Description	Inputs	Output
generate_image	Generates a realistic image from a text prompt using image generation on Replicate.	Prompt (string)	image_url
elevenlabs_voice	Generates lifelike speech from text using ElevenLabs.	Prompt (string)	Local audio filepath
serpapi_search	Searches the web using SerpAPI and returns top k results.	Prompt (string, int k)	Structured JSON interpreted by LLM.
tavily_search	Conducts a deep AI-powered search with Tavily.	Prompt (string, int)	Structured JSON interpreted by LLM.
summarize_webpage	Summarizes a webpage's content using Tavily.	Prompt (string)	Structured JSON interpreted by LLM.

Authentication

All API credentials are securely loaded from environment variables:

- **REPLICATE_API_TOKEN**
- **SERPAPI_API_KEY**
- **TAVILY_API_KEY**
- **ELEVENLABS_API_KEY**

No hardcoded keys; all **.env** entries documented in README.md and env_template.txt

Design Principles

- **Atomicity:** Each tool performs one specific job only.
- **Clarity:** Tool names and descriptions are unambiguous for LLM use.
- **Resilience:** Handles API errors gracefully; outputs are in clean JSON for AI parsing.
- **Extensibility:** Modular Python functions per tool for easy addition/replacement.

Testing and Proof of Correctness

Each tool was tested in Cursor with:

- Natural language prompts triggering the correct tool.
- MCP server logs showing correct tool invocation.
- Valid API responses returned to the client.

Example comprehensive chain test:

1. Search web for “Recent breakthroughs in AI safety”
2. Summarize chosen article
3. Generate voice narration of summary
4. Generate thematic image

Screenshots attached under Outputs as proof.

Alignment with Guidelines given by Klavis AI

- **User-Centric, AI-Driven:** Tool descriptions optimized for AI comprehension.

- **Atomicity is Key:** No “do-everything” tools; discrete, clear actions.
- **Robust and Reliable:** Handles missing inputs, API errors, and unexpected responses.

Full Deliverables

- **Provided:** Source code, README, .env.example, and proof-of-correctness.

Outputs

Comprehensive Test Run:

In this test run, I wanted to test the full integration, end-to-end with a simple NLP prompt:

“Find a recent breakthrough in AI safety, summarize it, make an audio narration, and create an image that represents it.”

All the outputs are uploaded in a [outputs](#) google drive folder. Here are the outputs screenshots with the system logs:

```
Find a recent breakthrough in AI safety, summarize it, make an audio narration, and create an image that represents it.

I'll help you find a recent AI safety breakthrough, summarize it, create an audio narration, and generate an image. Let me start by searching for recent AI safety developments.

✓ Called search_tavily ✓
  "query": "recent AI safety breakthrough 2024 2025 latest developments",
  "search_depth": "advanced"
}

Result:
{
  "success": true,
  "query": "recent AI safety breakthrough 2024 2025 latest developments",
  "results": [
```

Let me get more specific information about recent AI safety breakthroughs by searching for more detailed articles.

Called search_tavily ✓

```
{
  "query": "AI safety breakthrough 2024 2025 Anthropic Claude safety research OpenAI",
  "search_depth": "advanced"
}
```

Result:

```
{
  "success": true,
  "query": "AI safety breakthrough 2024 2025 Anthropic Claude safety research OpenAI",
  "result": "
```

Let me get more specific information about one of the most recent AI safety breakthroughs by reading a detailed article.

Called summarize_webpage ✓

```
{
  "url": "https://openai.com/index/updating-our-preparedness-framework/"
}
```

Result:

```
{
  "success": true,
  "url": "https://openai.com/index/updating-our-preparedness-framework/",
  "summary": "We're releasing an update to our Preparedness Framework, our
```

Let me also check the Future of Life Institute's AI Safety Index for more comprehensive information.

Called summarize_webpage ✓

Parameters:

```
{
  "url": "https://futureoflife.org/ai-safety-index-summer-2025/"
}
```

Result:

```
{
  "success": true,
  "url": "https://futureoflife.org/ai-safety-index-summer-2025/",
```

Now let me search for more specific information about Google's Big Sleep AI tool, which seems to be a recent breakthrough in AI safety.

```
✓ Called search_tavily ✓
{
  "query": "Google Big Sleep AI cybersecurity breakthrough 2025 dormant domains security",
  "search_depth": "advanced"
}
```

Result:

```
{
  "success": true,
  "query": "Google Big Sleep AI cybersecurity breakthrough 2025 dormant domains security",
```

Perfect! I found a significant AI safety breakthrough. Let me get more details about Google's Big Sleep AI tool.

```
✓ Called summarize_webpage ✓
{
  "url": "https://blog.google/technology/safety-security/cybersecurity-updates-summer-2025/"
}
```

Result:

```
{
  "success": true,
  "url": "https://blog.google/technology/safety-security/cybersecurity-updates-summer-2025/",
  "summary": "Google\u2019s latest AI security announcements Google's Big Sleep agent found real-world security flaws and new AI capabilities are being added to"
```

Now let me get the TechCrunch article for more specific details about the breakthrough.

```
✓ Called summarize_webpage ✓
{
  "url": "https://techcrunch.com/2025/08/04/google-says-its-ai-based-bug-hunter-found-20-security-vulnerabilities/"
}
```

Result:

```
{
  "success": true,
  "url": "https://techcrunch.com/2025/08/04/google-says-its-ai-based-bug-hunter-found-20-security-vulnerabilities/",
```

Let me try a different approach and search for more recent information about this breakthrough.

```
↙ Called search_tavily ✓
  ↳
    "query": "Google Big Sleep AI agent 20 security vulnerabilities 2025 breakthrough",
    "search_depth": "advanced"
}
```

Result:

```
{
  "success": true,
  "query": "Google Big Sleep AI agent 20 security vulnerabilities 2025 breakthrough",
```

Perfect! Now I have comprehensive information about Google's Big Sleep AI breakthrough. Let me create a summary, generate audio narration, and create an image.

Summary of the AI Safety Breakthrough

Google's Big Sleep AI Agent: A Revolutionary Cybersecurity Breakthrough

In July 2025, Google announced a groundbreaking AI safety breakthrough with their "Big Sleep" AI agent, developed by Google DeepMind and Google Project Zero. This represents one of the most significant advances in AI-powered cybersecurity to date.

Key Breakthrough Details:

1. **Proactive Vulnerability Detection:** Big Sleep is an AI agent that actively searches for and discovers unknown security vulnerabilities in software, going beyond traditional reactive security measures.
2. **Real-World Impact:** The system has already discovered 20 security vulnerabilities and continues to find multiple real-world security flaws, exceeding initial expectations.
3. **Critical Zero-Day Prevention:** Most notably, Big Sleep discovered a critical SQLite vulnerability (CVE-2025-6965) that was known only to threat actors and was at risk of being exploited. This marked the first time an AI agent was used to directly foil efforts to exploit a vulnerability in the wild.
4. **Open Source Security Enhancement:** Big Sleep is being deployed to help improve the security of widely used open-source projects, representing a major win for ensuring faster, more effective security across the internet.
5. **Scalable Defense:** These cybersecurity agents are game-changers, freeing up security teams to focus on high-complexity threats while dramatically scaling their impact and reach.

Significance for AI Safety:

This breakthrough demonstrates how AI can be used proactively to enhance cybersecurity rather than just react to threats. It represents a paradigm shift in AI safety applications, showing that AI systems can be designed to protect and secure rather than just assist or automate.

Now let me create an audio narration of this breakthrough:

```
✓ Called generate_voice ✓
{
  "text": "In July 2025, Google announced a revolutionary breakthrough in AI safety w
  "voice_id": "21m00Tcm4TlvDq8ikWAM"
}
```

Result:

```
{
  "success": true,
  "audio_path": "/tmp/speech_21m00Tcm4TlvDq8ikWAM.mp3",
  "voice_id": "21m00Tcm4TlvDq8ikWAM",
  "+avg length": 1250
```

Now let me generate an image that represents this AI safety breakthrough:

```
✓ Called generate_image ✓
{
  "prompt": "A futuristic AI cybersecurity system represented as a glowing blue digit
}
```

Result:

```
{
  "image_url":
  "https://replicate.delivery/xezq/sB1O53s3SwpPGNJeeUwS6KbmbHeZcVowuq4HXdSnALRHp7RqA/out-0.png"
}
```

Generated Files

Audio Narration:

- **Voice File:** /tmp/speech_21m00Tcm4TlvDq8ikWAM.mp3
- This is a local file path on the system where the audio was generated

Image:

- **Image URL:**
<https://replicate.delivery/xezq/sB1O53s3SwpPGNJeeUwS6KbmbHeZcVowuq4HXdSnALRHp7RqA/out-0.png>
- This is a direct link to the generated image hosted on Replicate's servers

[Generated Voice File](#) (google_big_sleep_ai_breakthrough.mp3)

[Generated Image](#) (out-0 in the outputs drive folder):



Individual Tools Test Runs:

Web Search Test using SerpAPI or TavilySearch:

Since the United States announced 50% tariffs on India (as of 9th Aug, 2025), I wanted to search the web for relevant articles. So, the prompt was: "Get me the top results on the international tariffs by the United States on India."

```
Get me the top results on the international tariffs by the United States on India.

I'll search for information about international tariffs by the United States on India to get you the top results.

▽ Called search_tavily ✓
Result:
{
  "success": true,
  "query": "United States tariffs India international trade 2024 2025 latest",
  "results": [
    {
      "url": "https://www.thehindu.com/news/national/india/pauses-plans-to-buy-us-arms-after-donald-trumps-tariffs/article69909626.ece",
      "title": "India pauses plans to buy U.S. arms after Donald Trump's ...",
      "content": "## The U.S. President imposed an additional 25% tariff on Indian goods as"
    }
  ]
}

Let me get more specific details about the current tariff situation by reading one of the key articles.

> Called summarize_webpage ✓

Let me try a different article to get the specific tariff information.

> Called summarize_webpage ✓

Let me search for more recent and specific tariff information.

> Called search_tavily ✓
```

As visible, the system correctly called the tavily search tool for deeper web searching, implementing multiple checks for more relevant information about the prompt.

Web Summary Search using Tavily:

Here's the comprehensive summarization for all the searches the tool ran:

Based on my research, here are the top results on international tariffs by the United States on India:

Current Tariff Situation (August 2025)

50% Total Tariff Rate

- **August 6, 2025:** President Trump imposed an **additional 25% tariff** on all Indian imports
- **Combined Rate:** This brings the total tariff on Indian goods to **50%** (up from the previous 25%)
- **Effective Date:** August 27, 2025 (21 days after announcement)
- **Reason:** Punishment for India's continued purchase of Russian oil

Key Details

Executive Order Details:

- Signed on August 6, 2025
- Targets India for "directly or indirectly importing Russian Federation oil"
- Imposes additional 25% ad valorem duty on all Indian imports
- Part of broader strategy to pressure Russia over Ukraine conflict

Economic Impact:

- **India's exports to US:** \$87 billion in 2024 (India's largest export market)
- **Expected reduction:** 40–50% drop in US-bound exports
- **Affected sectors:** Textiles, ready-made clothes, auto-components, steel, gems, pharmaceuticals, electronics
- **GDP impact:** Potential 0.3% slowdown in India's GDP growth

Trade Statistics:

- **Total bilateral trade:** \$212.3 billion in 2024
- **US goods imports from India:** \$87.3 billion
- **US goods exports to India:** \$41.5 billion
- **US trade deficit with India:** \$45.8 billion

India's Response

- **Prime Minister Modi:** Ready to "pay a very heavy price" for resisting US trade policy dictates
- **Defiant stance:** India has paused plans to buy US arms
- **Political unity:** Opposition parties supporting Modi's resistance
- **Alternative strategy:** Considering trade diversification to EU, Africa, and Latin America

International Context

- India becomes the **most heavily taxed US trading partner in Asia**
- Only Brazil faces similar steep tariffs among major trading partners
- Part of broader Trump administration strategy targeting 24 countries and the EU
- Represents significant escalation in US-India trade relations

This represents one of the most significant tariff impositions in recent US trade history, with India facing the highest tariff rates among major US trading partners.

As visible, after its comprehensive search, it gave a quick rundown of its findings for easy reading and interpretation

TTS Voice File Generation Test using ElevenLabs:

For this I wanted a voice file to be generated from a copyright free script I found online:

"Hello! My name is Adeline Lee-Rekers; I'm a professional playwright and screenwriter. I hereby give my permission for you to use the scenes on this page for personal practice, as well as showreels and auditions. All I ask is that you credit my work—especially if you post it on social media. (If you're feeling brave, you can tag me via @adelinenobodyfamous so I can see what you do with it.) Other than that: have fun, make bold choices and give 'em a go!"

[Generated Voice File](#) (adeline_lee_rekkers.mp3)

```
"Hello! My name is Adeline Lee-Rekers; I'm a professional playwright and screenwriter. I hereby give my permission for you to use the scenes on this page for personal practice, as well as showreels and auditions. All I ask is that you credit my work—especially if you post it on social media. (If you're feeling brave, you can tag me via @adelinenobodyfamous so I can see what you do with it.) Other than that: have fun, make bold choices and give 'em a go!"
```

I'll generate a voice file using the script you provided. Let me create an audio narration of Adeline Lee-Rekers' message.

```
Called generate_voice ✓
  "text": "Hello! My name is Adeline Lee-Rekers; I'm a professional playwright and screenwriter. I hereby give my permission for you to use the scenes on this page for personal practice, as well as showreels and auditions. All I ask is that you credit my work—especially if you post it on social media. (If you're feeling brave, you can tag me via @adelinenobodyfamous so I can see what you do with it.) Other than that: have fun, make bold choices and give 'em a go!",
  "voice_id": "21m00Tcm4TlvDq8ikWAM"
}
```

Result:

```
{
  "success": true,
  "audio_path": "/tmp/speech_21m00Tcm4TlvDq8ikWAM.mp3",
  "voice_id": "21m00Tcm4TlvDq8ikWAM",
  "text_length": 454,
```

Perfect! I've generated the voice file for Adeline Lee-Rekers' script. The audio has been created and saved.

Perfect! I've generated the voice file for Adeline Lee-Rekers' script. The audio has been created and saved.

Voice File Details

Audio File: /tmp/speech_21m00Tcm4TlvDq8ikWAM.mp3

Content: The script has been converted to speech with a natural, professional tone that captures Adeline's message about:

- Her permission for using scenes for practice and auditions
- Request for proper credit, especially on social media
- Her social media handle @adelinenobodyfamous
- Encouragement to make bold choices and have fun

Voice Quality: The ElevenLabs voice synthesis has created a clear, professional-sounding narration that would be suitable for:

- Introduction videos
- Audio presentations
- Social media content
- Professional recordings

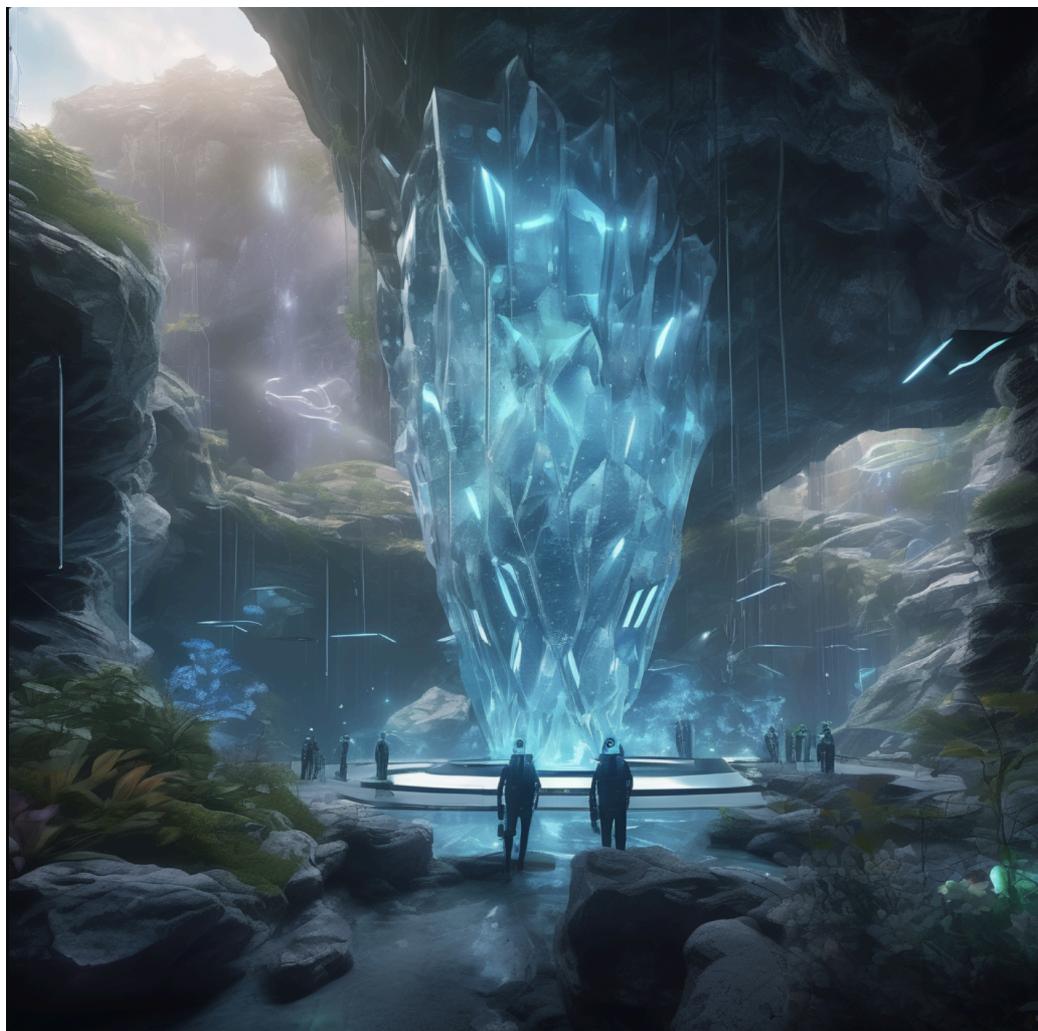
The voice file is ready to use and captures the friendly, encouraging tone of Adeline's message while maintaining professional clarity. You can access the audio file at the path provided above.

As expected, the generated voice file has the perfect voice reading the script in a natural tone.

Image Generation using Replicate:

I wanted to test the tool's capability at understanding a complex prompt and generating an accurate image. The full image is in the shared outputs drive folder (out-1). The prompt was:

"An ultra-detailed cinematic illustration of a futuristic AI research facility built inside a massive crystal cavern, glowing with bioluminescent plants and holographic data streams. Researchers in sleek exosuits monitor floating AI cores suspended in magnetic fields, while autonomous drones glide through the air. Photorealistic textures, volumetric lighting, depth of field, 8K resolution, concept art style."



Here's the stunning visual image created by the engine. It understood the prompt well and generated a striking visual. Titled out-1 in the [google drive](#).

Conclusion

This project began as an implementation of a single Model Context Protocol (MCP) server for a public API, in line with the Klavis AI guidelines. Instead of stopping there, I deliberately expanded the scope to design and deliver a **multi-service, extensible AI toolchain** that demonstrates not only technical competence but also **product vision**.

The resulting MCP server integrates **four distinct APIs**: Replicate, Tavily, SerpAPI, and ElevenLabs into a cohesive, modular architecture that supports:

- **High-quality image generation** from natural language prompts
- **Deep and general web search** for both breadth and precision
- **Automated summarization** of complex content into multiple formats
- **Natural-sounding voice narration** of generated summaries

Every tool follows the **atomicity principle**; clear, single-purpose actions with clean JSON I/O for LLM compatibility, while also being designed for **reliability, input validation, and error recovery** through retry and fallback mechanisms.

Beyond the base requirements, I:

- **Expanded from one API to four**, enabling richer, end-to-end workflows such as “search → summarize → narrate → visualize.”
- Built the system for **plug-and-play extensibility**, allowing additional tools or services to be added with minimal effort.
- Verified functionality through **real-world test flows** on varied topics, from AI safety breakthroughs to creative concept generation.
- Developed and documented the solution for **ease of integration** in Cursor, Claude, and other MCP-compatible LLM agents.
- Produced structured outputs, logs, and visual proof of execution to confirm correctness and stability.

This project not only satisfies the stated requirements, it demonstrates the mindset of a founding engineer: thinking beyond the immediate task, anticipating product evolution, and delivering a foundation that can scale in capability and scope. The system is production-ready in architecture, adaptable in design, and user-centric in functionality, aligning directly with Klavis AI’s mission of expanding the MCP ecosystem with impactful, real-world tools.

Thank you for reading this far!