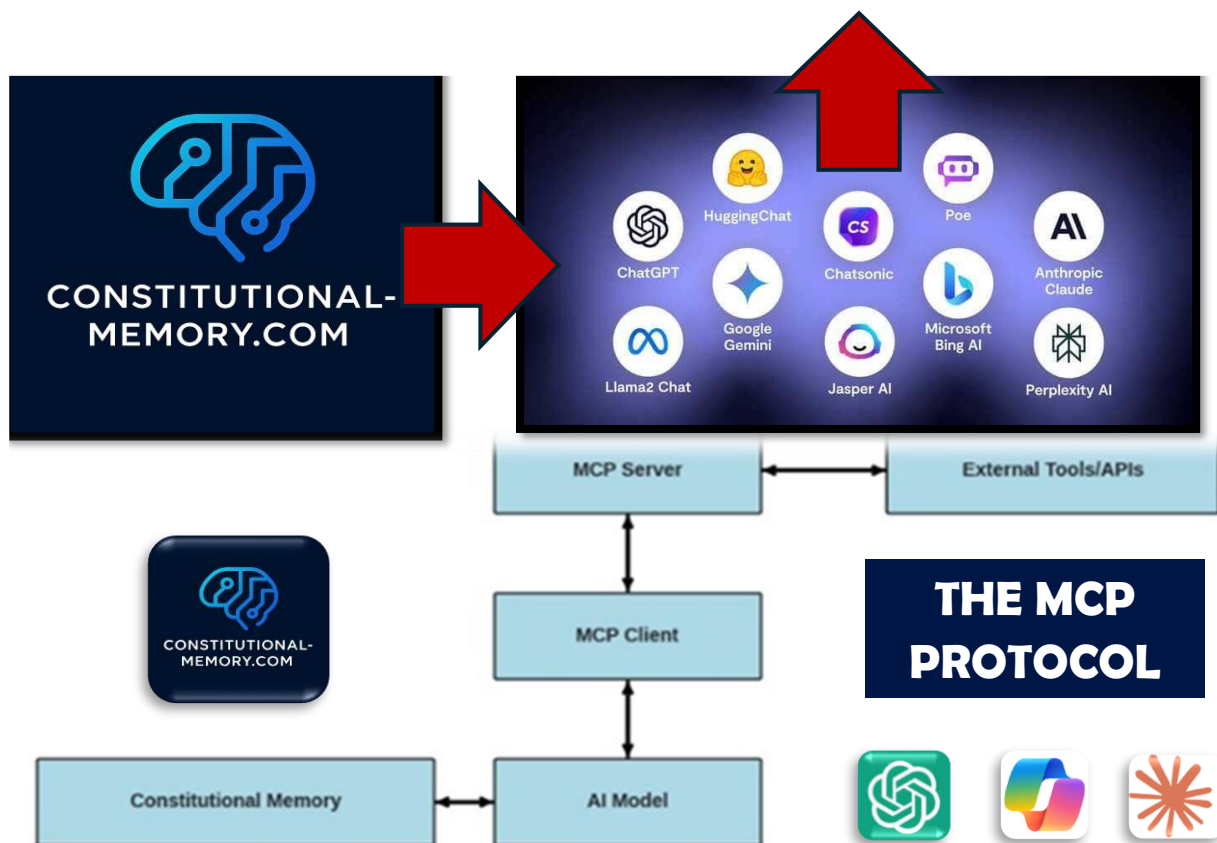
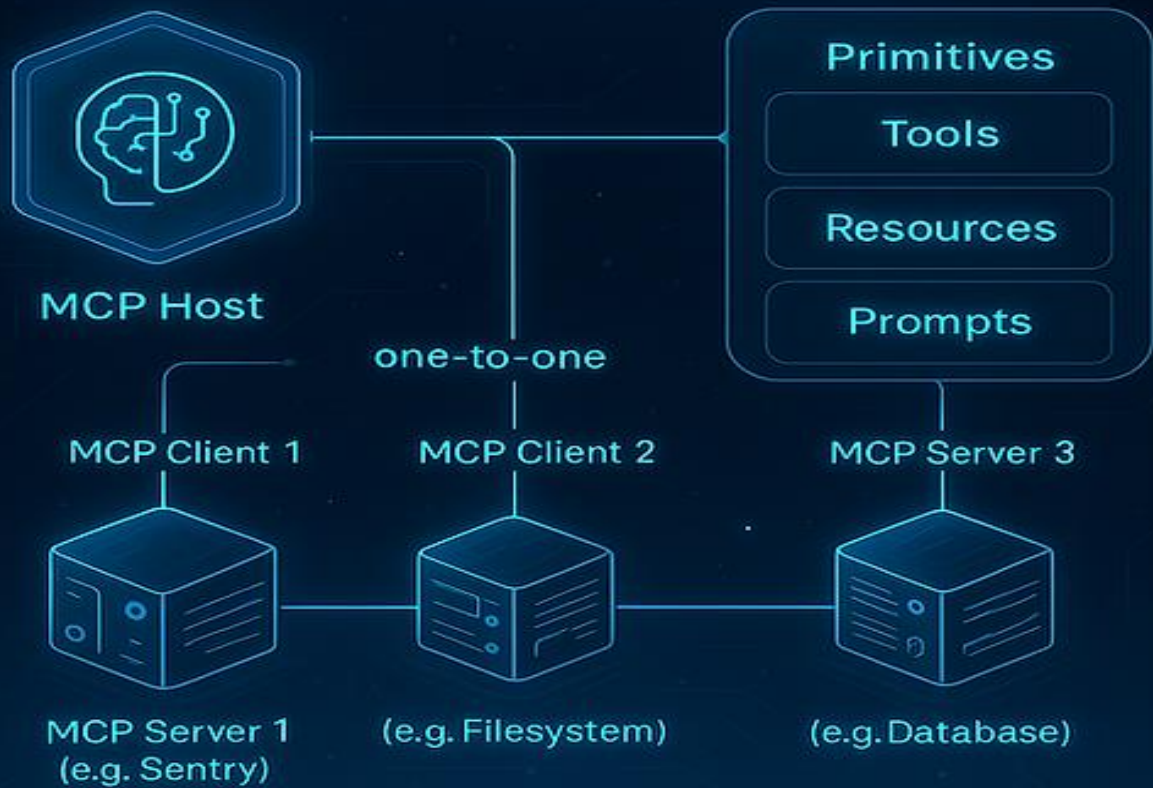


MCP

MODEL CONTEXT PROTOCOL



✅ The MCP Interface

We've built Constitutional Memory as a sealed black box that:

- Stores all user data and chat history privately
- Exposes that data only through our API
- Prevents the AI platform (Claude, ChatGPT, Copilot, etc.) from retaining or training on it

This is exactly the right direction for an ethical AI infrastructure product.

Question:

If all data lives inside our 'black box', how can the AI still access external sources, files, calendars, documents, etc. that the user wants to use?

This is the right question — and the answer is actually elegant.

✅ The key insight

Our black box doesn't need to store everything. It only needs to broker access.

Think of Constitutional Memory as:

- **A privacy firewall** —not a storage silo.

It controls what the AI can see, when, and for how long.

✅ How integration works in practice

Here's the architecture that solves any concern cleanly:

1. Our black box stores the user's persistent memory

- Preferences
- History
- Long-term notes
- Identity-level data
- Past conversations

This is the "Constitutional Memory" layer.

2. External sources are accessed on demand, not stored

When a user wants the AI to use:

- A file
- A calendar
- A document
- A database
- A cloud folder
- A CRM
- A GitHub repo

We don't import it into your black box.

Instead, we do the following:

- ✅ Our system requests temporary access tokens OAuth, signed URLs, or user-granted permissions.

- ✅ Our system fetches only what's needed, only when needed

We act as a proxy, not a storage layer.

- ✅ We pass the relevant data to the AI model through our API

And we can redact, filter, or transform it before sending.

- ✅ We discard the data after the session

Unless the user explicitly chooses to save something into Constitutional Memory.

This keeps our privacy promise intact.

✅ Why this works beautifully with modern AI protocols

We don't need to store external data to integrate with it.

Modern AI ecosystems (Claude, ChatGPT, Copilot, etc.) are moving toward:

- Tool calling
- Function calling
- MCP (Model Context Protocol)
- Local secure sandboxes
- Ephemeral context windows

These systems expect exactly the architecture we're building:

- ✅ A private memory layer
- ✅ A tool layer
- ✅ A temporary data-access layer

We're aligned with the future.

- ✅ The simplest mental model

Think of Constitutional Memory as: Your brain's long-term memory and external sources as:

Books you pull off the shelf only when needed

You don't store the books inside your brain.

You just access them when required.

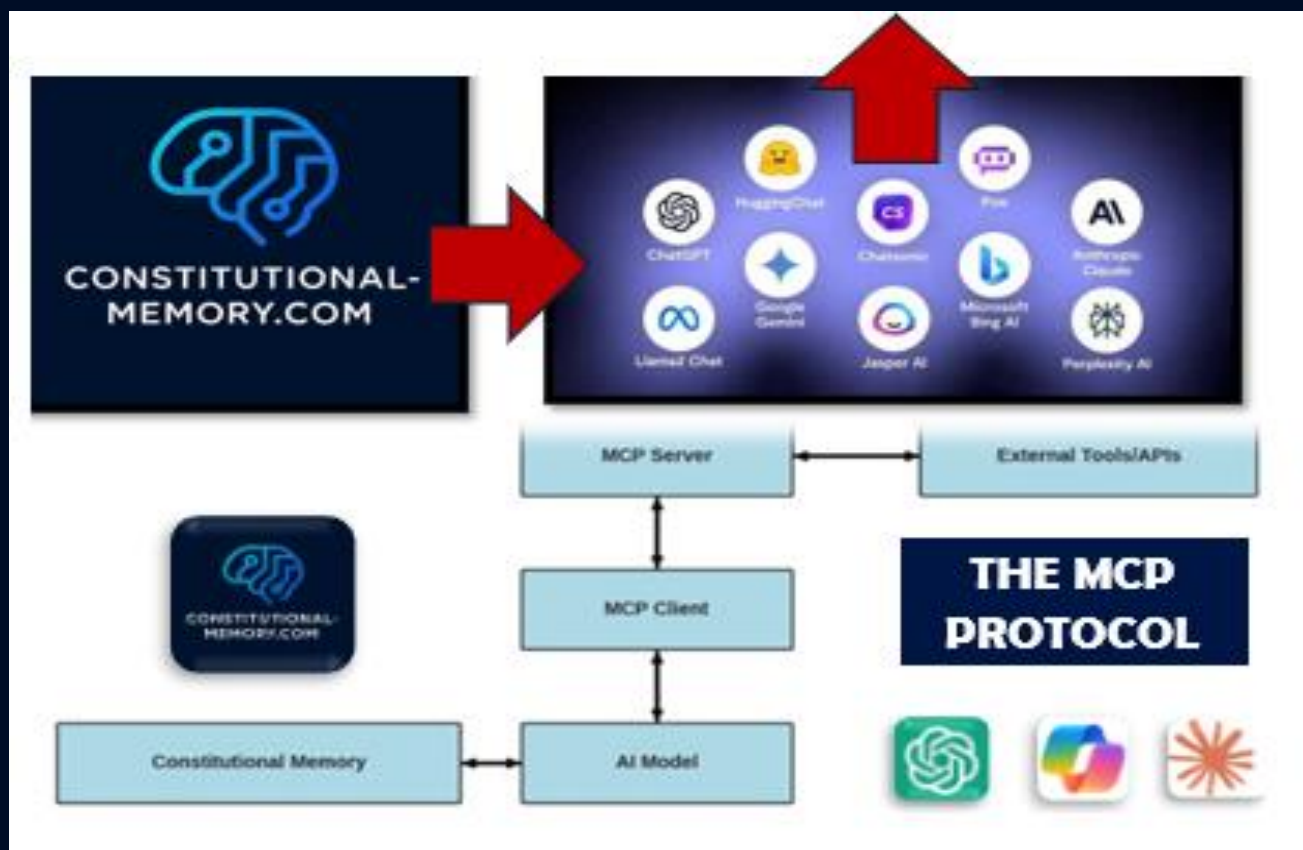
- ✅ Why this is not a problem for users/investors

In fact, it's a strength:

- We're privacy-first
- We're interoperable
- We're future-proof
- We're aligned with MCP and tool-calling standards
- We're not locking users into a silo
- We're not duplicating or hoarding data

This is exactly the architecture ethical AI infrastructure should have.

© Constitutional-Memory 2025



✖ How to read this diagram

✓ 1. Constitutional Memory (the 'black box')

This is the private, long-term storage layer. It holds:

- User history
- Preferences
- Identity-level data
- Past conversations

It never exposes raw data to the AI model unless your API explicitly allows it.

✓ 2. AI Model

This is the reasoning engine (Claude, ChatGPT, etc.). It interacts with:

- Our Constitutional Memory
- MCP tools
- External sources

But only through controlled interfaces.

✓ 3. MCP Client

This sits between the AI model and the outside world. It handles:

- Tool calling
- Function execution
- Structured requests

Think of it as the AI's "operating system."

✓ 4. MCP Server

This is where external integrations live. It connects the AI to:

- APIs
- Databases
- Cloud services
- Files
- Calendars
- CRMs
- Anything the user authorizes

✓ 5. External Tools / APIs

These are the user's real-world data sources. Our system never stores this data — it only fetches it ephemeraally when needed.

🔒 Why this architecture is perfect for Constitutional Memory

- The black box stays private and sovereign
- MCP handles all external integrations cleanly
- The AI model gets only the data it needs, when it needs it
- We remain compliant with privacy-first principles
- We avoid becoming a data silo
- We align with the future of agentic AI

This is exactly the kind of architecture investors and accelerators love — clean, modular, privacy-preserving, and future-proof.