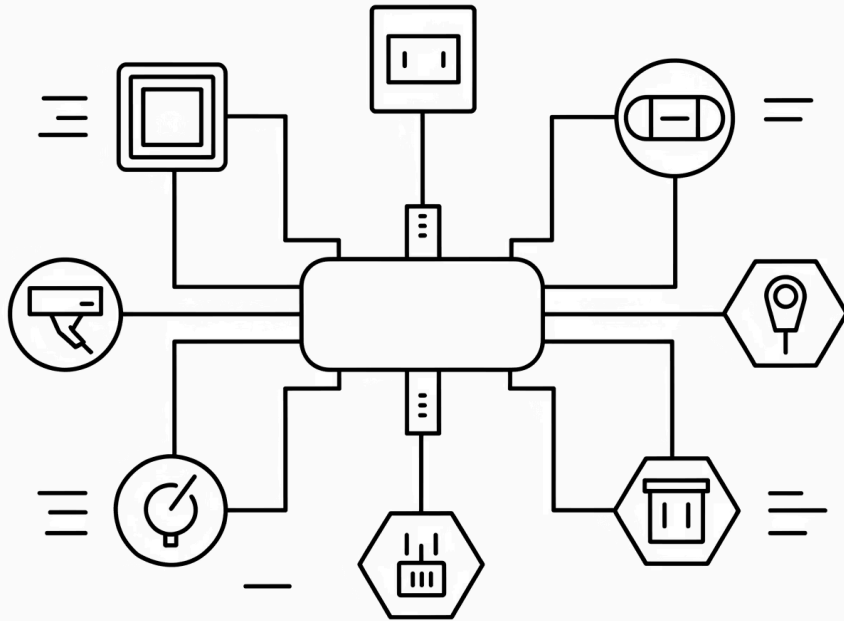


VS Code & MCP: Your Editor, Supercharged

How the Model Context Protocol transforms GitHub Copilot into a fully agentic development partner — connected to databases, APIs, repos, and more.

What Is the Model Context Protocol?



MCP is an open standard developed by Anthropic that defines how applications share context with large language models. Think of it as a **universal adapter** — instead of requiring custom integrations between every AI model and every tool, MCP standardises the interface.

Read Files

Query DBs

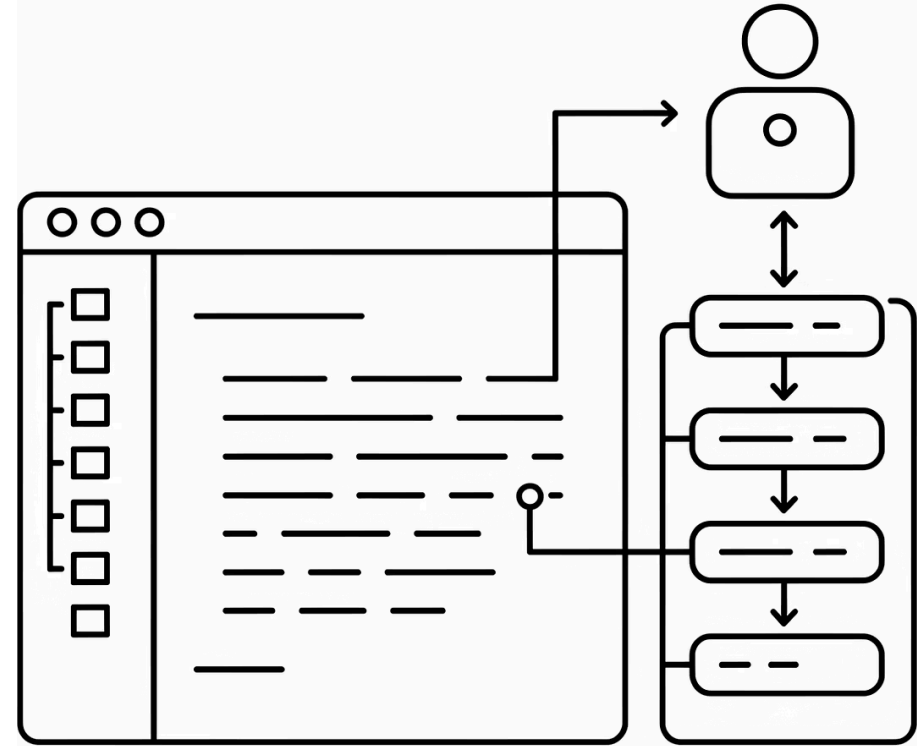
Call APIs

Manage Repos

Agent Mode: The Gateway to MCP

MCP tools come alive through **Agent Mode** in Copilot Chat. Unlike Ask or Edit modes, Agent Mode lets Copilot independently plan and execute multi-step tasks — running terminal commands, editing files, calling MCP tools, and self-correcting errors.

- Select "Agent" in the Copilot Chat mode dropdown
- Copilot automatically invokes the right MCP tool based on context
- Reference tools explicitly with `#toolname` in your prompt



Configuring Local MCP Servers

How It Works

Local servers run as processes on your machine, communicating via **stdio**. Configuration lives in `.vscode/mcp.json` — committable to source control to share setup with your team.

Use `${input:...}` variables to keep credentials out of the config file. VS Code prompts for values at server start — essential for DB passwords, API keys, and tokens.

```
{
  "servers": {
    "myDatabaseServer": {
      "type": "stdio",
      "command": "node",
      "args": ["/path/to/server/index.js"],
      "env": {
        "DB_CONNECTION_STRING":
          "${input:dbConnection}"
      }
    }
  }
}
```

The MCP Server Gallery & Remote Servers

Server Gallery

Type @mcp in the Extensions search bar to browse a curated list from the GitHub MCP Registry. Install per workspace or user profile, manage via Extensions panel or Command Palette.



Remote Servers (SSE)

Host servers on cloud, team infrastructure, or vendor services using Server-Sent Events. Ideal when central access management is needed. Copilot Business/Enterprise admins can enforce policies on permitted servers.



Debug Mode

VS Code offers a dedicated debug mode for Node.js and Python MCP servers — develop, test, and iterate on custom servers in the same environment where you'll ultimately use them.

Practical MCP Use Cases

Repository & Project Management

The GitHub MCP server lets Copilot search repos, create issues, open PRs, and inspect CI/CD runs — all via natural language. *"Create a GitHub issue for the bug we just found"* triggers the full flow.

Database Queries

Expose read-only query tools so Copilot understands your actual schema when generating SQL. Use `.github/copilot-instructions.md` to always prefer the MCP tool over raw SQL generation.

Infrastructure Management

Connect Kubernetes, Terraform state, or cloud provider APIs to give Copilot live infrastructure context — enabling diagnosis, change suggestions, and IaC generation with real configuration details.

Custom Internal Tools

Build private MCP servers in any language exposing internal APIs, documentation, or datasets. Share via your enterprise private MCP registry — discoverable only within the organisation.

AI Models Available in VS Code

Copilot's Agent Mode filters the model list to those with **strong tool-calling support**. Each model brings distinct strengths to MCP-heavy workflows.



OpenAI GPT-4.1 / GPT-5

GPT-4.1 is the reliable everyday default. GPT-5 handles complex multi-step agentic tasks; GPT-5 mini trades depth for speed and lower request consumption.



Claude Sonnet 4 & Opus 4

Anthropic's flagships excel at long-context reasoning, careful planning, and orchestrating multi-tool sequences. **Auto mode** defaults to Sonnet 4 and GPT-5.



Google Gemini 2.5 Pro / 2.0 Flash

Gemini 2.5 Pro leads on data transformation and multimodal inputs — ideal for attaching logs or CSVs. Flash prioritises low latency for quick iterative tasks.

Auto Selection & Bring Your Own Key

Auto Model Selection

The **Auto** option routes requests to the best available model based on load, availability, and task context. It transparently switches between Claude Sonnet 4, GPT-5, and GPT-5 mini — helping avoid rate limits without manual switching.

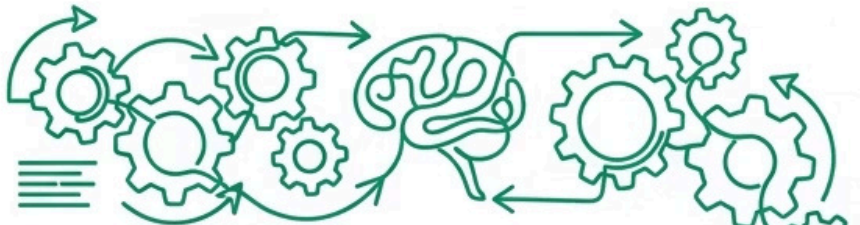
- 📌 Auto is the recommended default for most developers who don't want to manage model selection manually.

Bring Your Own Key (BYOK)

Connect directly to Anthropic, OpenAI, Azure OpenAI, Mistral, or **Ollama** for fully local models like Phi-4 or Llama. BYOK is available on individual Copilot plans — ideal for:

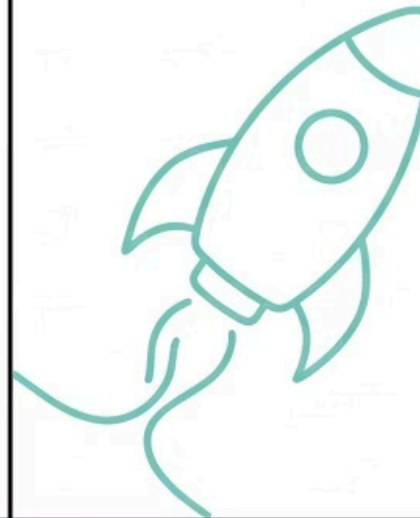
- Fine-tuned internal models
- Newer releases not yet in Copilot
- Fully local inference with no data egress

Choosing the Right Model for MCP Workflows



MULTI-STEP TOOL ORCHESTRATION

Claude Opus 4/Sonnet 4: long-context reasoning, careful planning.



RAPID ITERATIVE DEVELOPMENT

GPT-4.1/GPT-5 mini: speed, code generation.

ANALYTICAL TASKS, DATA ATTACHMENTS

Gemini 2.5 Pro: logs, CSVs, structured data.



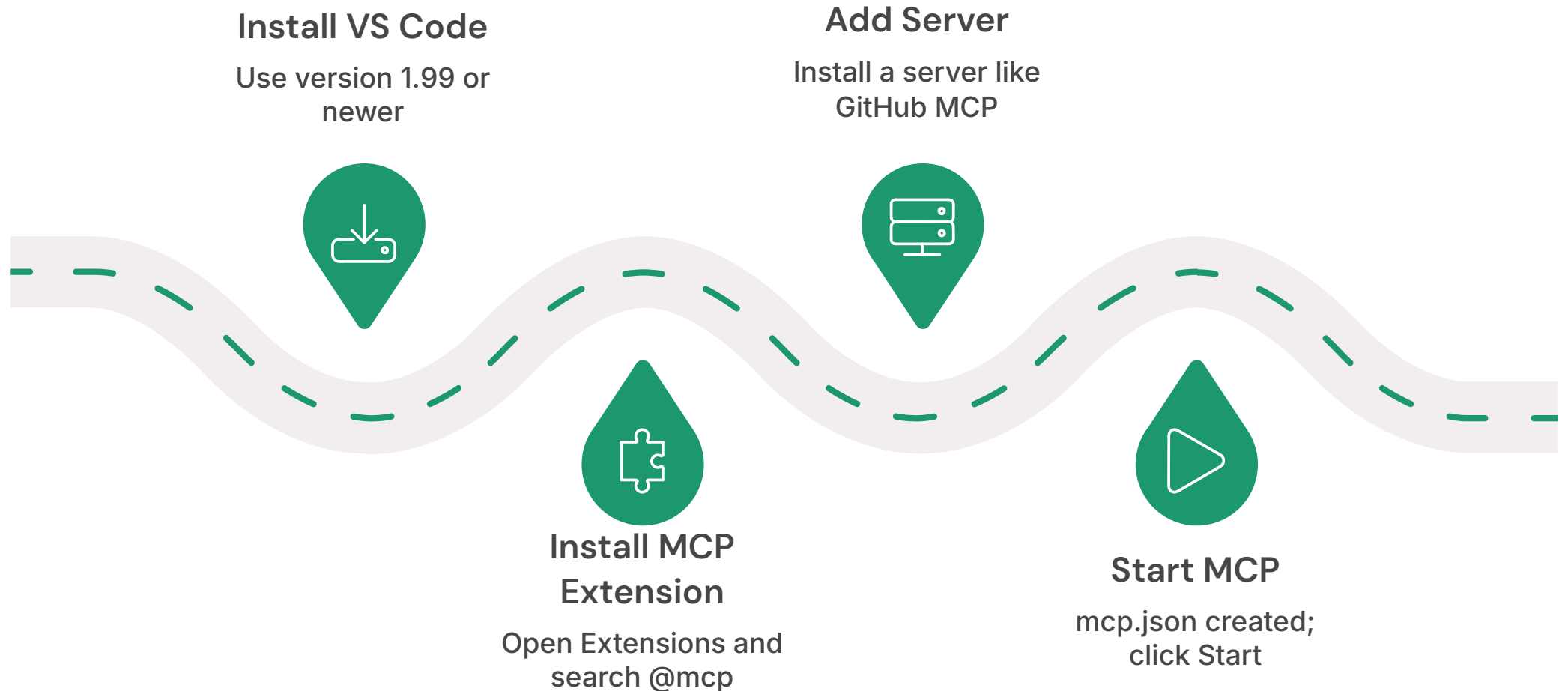
UNCERTAIN OR MIXED WORKLOADS

Auto mode: transparent optimization.



When orchestrating multiple MCP tool calls in sequence, **Claude models lead on careful multi-step planning**. For rapid single-tool invocations, favour speed with GPT-4.1 or GPT-5 mini. For data-heavy analytical work, Gemini 2.5 Pro is the standout choice.

Getting Started in 7 Steps



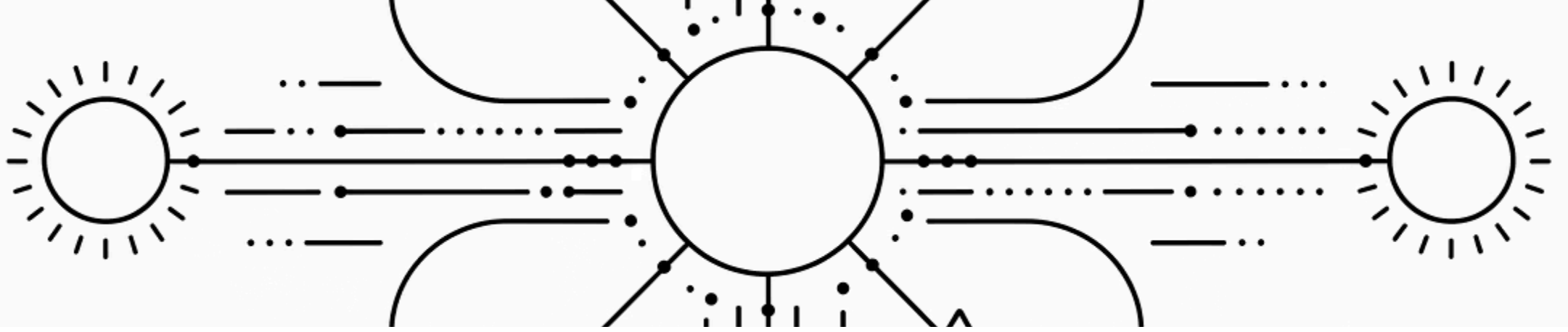
The full setup takes under five minutes. For a custom local server, create `mcp.json` manually with a `stdio` configuration pointing to your executable, then use **MCP: List Servers** from the Command Palette to restart.



From Code Suggestions to True Development Partner

VS Code's MCP integration represents a meaningful shift in how developers interact with AI in their daily work. By connecting AI models to real tools and data through a standardised protocol, Copilot moves from a code-suggestion engine to a true development partner.

Agent Mode + MCP servers + multi-model choice = **the most capable AI-augmented coding environment available today.**



Shaping the Future of AI-Augmented Development

At Sitemule, we are dedicated to transforming the developer experience. Our innovative products and services integrate seamlessly with platforms like VS Code and leverage the Model Context Protocol, empowering engineers with intelligent, collaborative AI tools.

From advanced code generation and intelligent debugging to efficient project management and secure infrastructure interaction, we provide the cutting-edge solutions that turn every developer into a true development partner.

[Explore Our Solutions](#)

[Learn About Sitemule](#)