



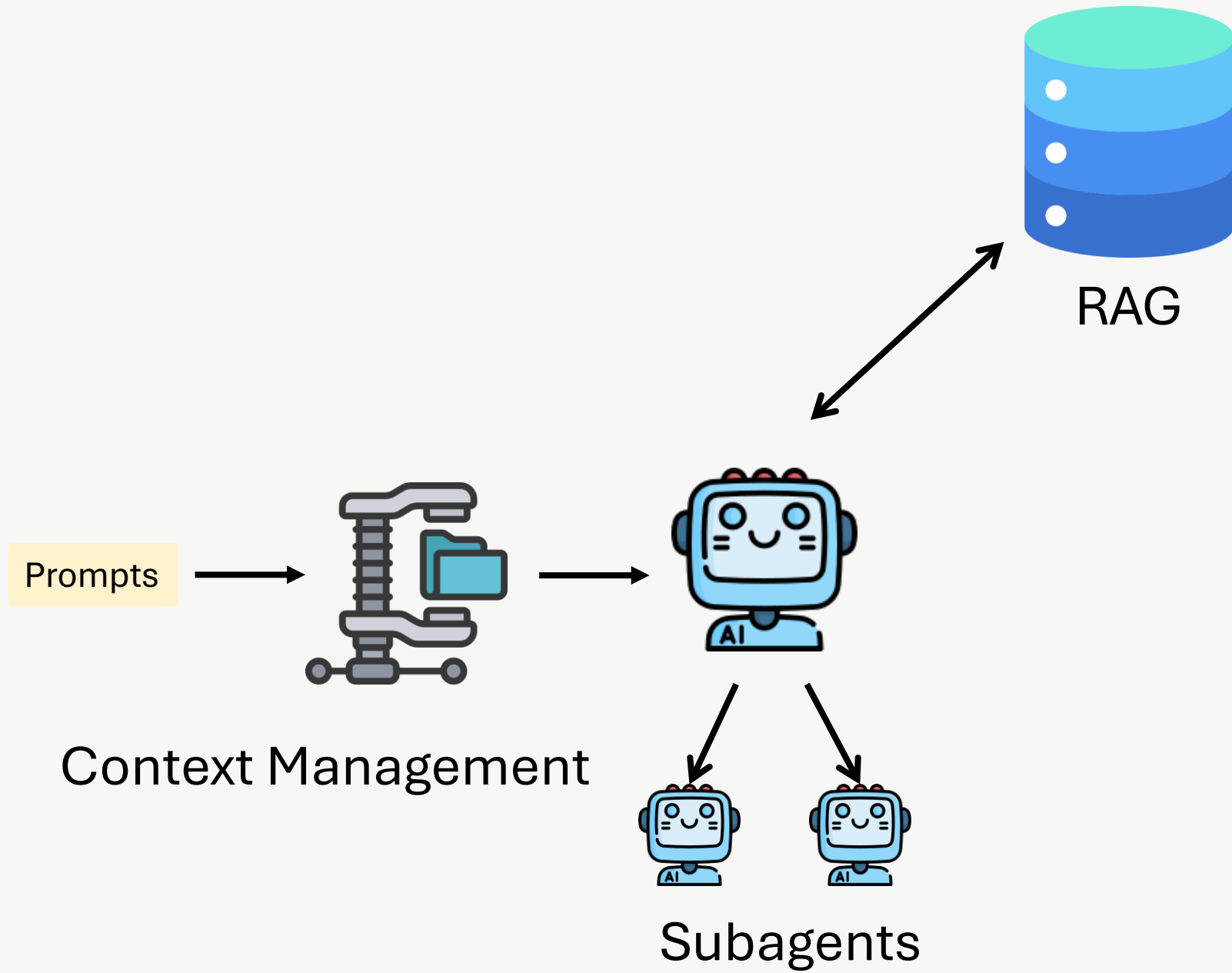
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

CSC4801

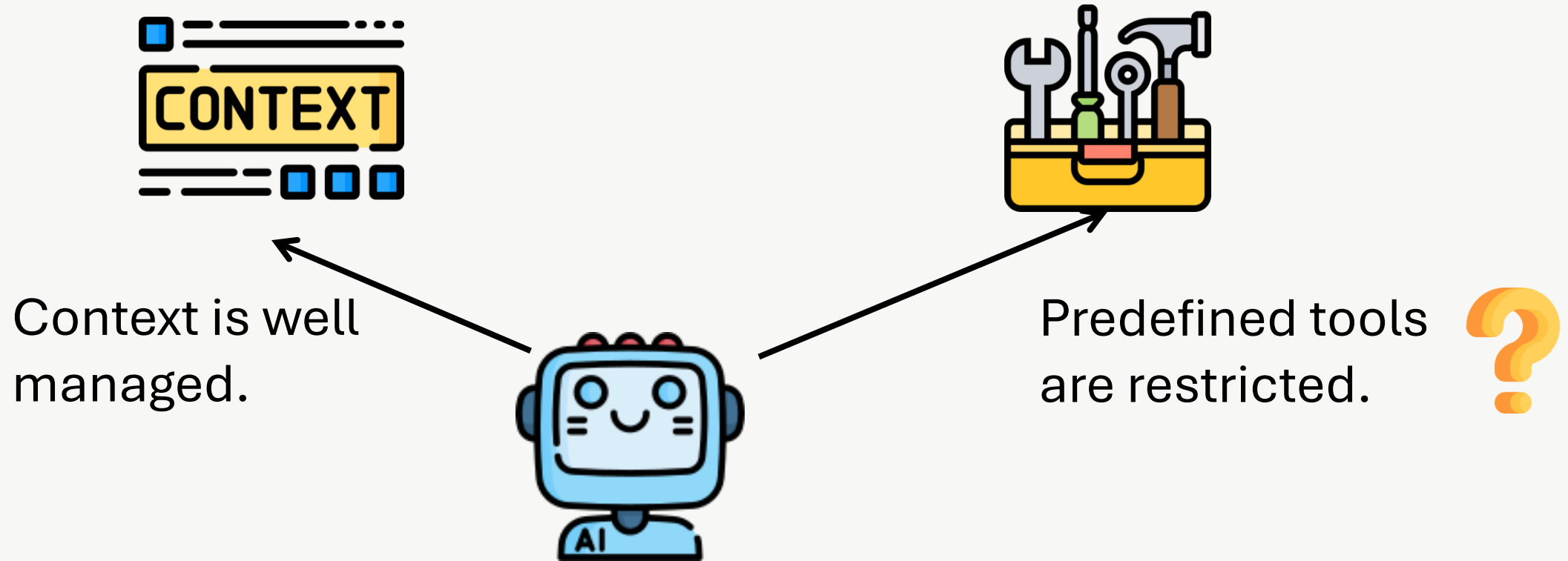
AI-assisted Software Engineering

Lecture 04: MCP & Skills

Instructor: Jinsheng Ba | 2026 Fall



Problem



Predefined Tools

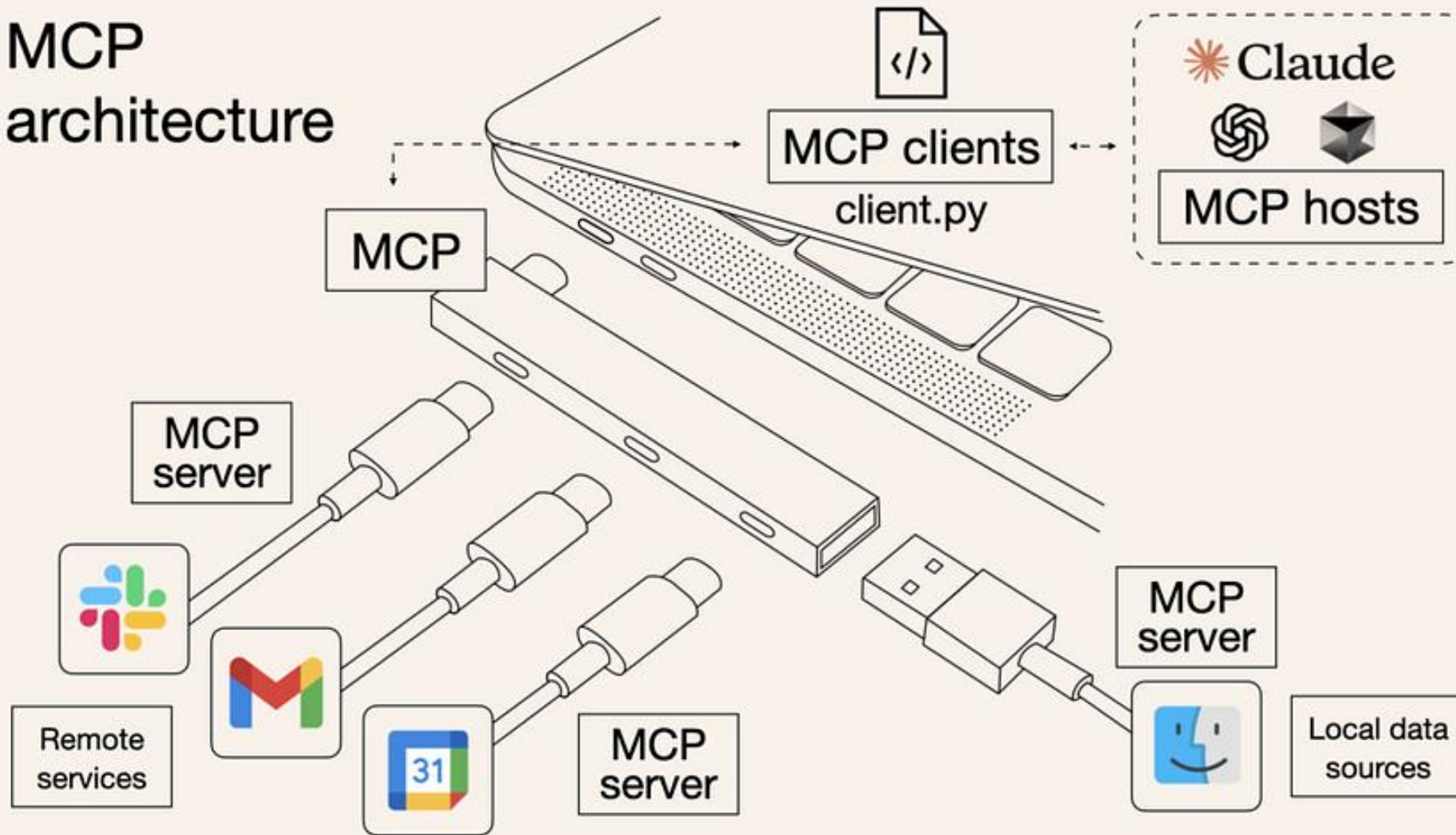
Tool	Description	Permission Required
Bash	Executes shell commands in your environment. See Bash tool behavior	Yes
Edit	Makes targeted edits to specific files. See Edit tool behavior	Yes
Glob	Finds files based on pattern matching. See Glob tool behavior	No
Grep	Searches for patterns in file contents. See Grep tool behavior	No
Read	Reads the contents of files. See Read tool behavior	No
Write	Creates or overwrites files. See Write tool behavior	Yes

| Problem

How to efficiently add a new tool?

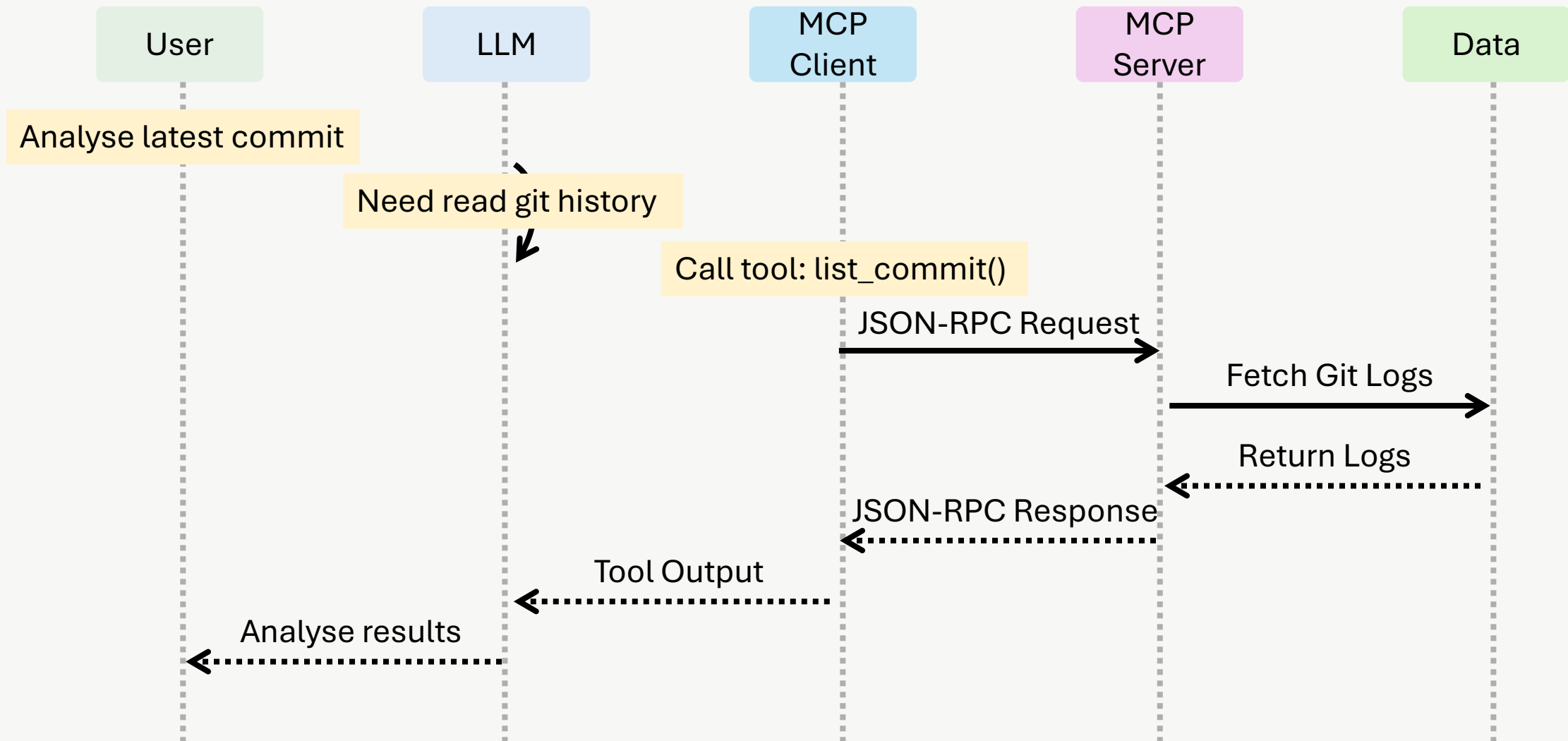
4 Model Context Protocol

MCP architecture

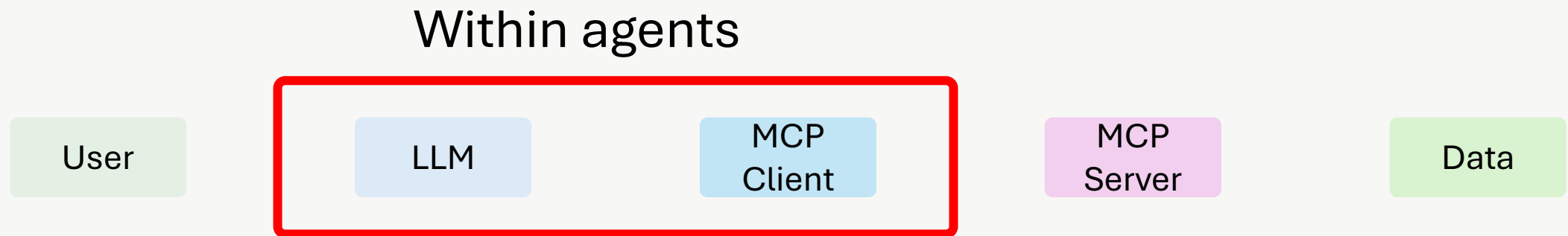


A unified interface to add tools for agents.

MCP Workflow



Architecture

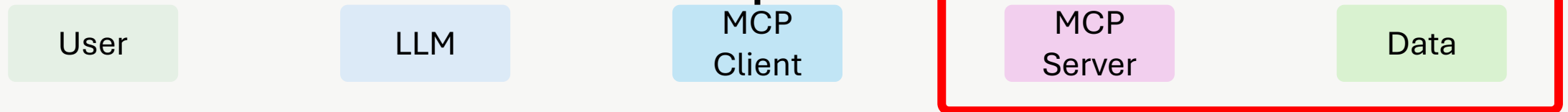


Architecture

Tool	Description	Permission Required
Bash	Executes shell commands in your environment. See Bash tool behavior	Yes
Edit	Makes targeted edits to specific files. See Edit tool behavior	Yes
Glob	Finds files based on pattern matching. See Glob tool behavior	No
Grep	Searches for patterns in file contents. See Grep tool behavior	No
Read	Reads the contents of files. See Read tool behavior	No
Write	Creates or overwrites files. See Write tool behavior	Yes

Register as a tool

External servers



Example: FastMCP

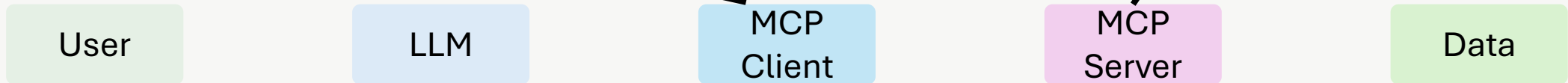
```
# Call the add tool
result = await client.call_tool(
    "add",
    {
        "a": 3,
        "b": 5
    }
)
```

```
from fastmcp import FastMCP

mcp = FastMCP("Demo 🚀")

@mcp.tool
def add(a: int, b: int) -> int:
    """Add two numbers"""
    return a + b

if __name__ == "__main__":
    mcp.run()
```



<https://gofastmcp.com/>

Let us build a simple MCP server

JSON-RPC Protocol

```
# Call the add tool
result = await client.call_tool(
    "add",
    {
        "a": 3,
        "b": 5
    }
)
```



```
{
  "jsonrpc": "2.0",
  "id": 1,
  "method": "tools/call",
  "params": {
    "name": "add",
    "arguments": {
      "a": 3,
      "b": 5,
    }
  }
}
```

```
from fastmcp import FastMCP

mcp = FastMCP("Demo 🚀")

@mcp.tool
def add(a: int, b: int) -> int:
    """Add two numbers"""
    return a + b

if __name__ == "__main__":
    mcp.run()
```



Protocol

```
{
  "jsonrpc": "2.0",
  "id": 1,
  "method": "tools/call",
  "params": {
    "name": "get_weather",
    "arguments": {
      "city": "Zurich"
    }
  }
}
```

<https://gofastmcp.com/servers/server>

<https://modelcontextprotocol.io/docs/learn/server-concepts>

Tools are functions that clients invoke to perform actions or access external systems.

```
@mcp.tool
def multiply(a: float, b: float) -> float:
    """Multiplies two numbers together."""
    return a * b
```

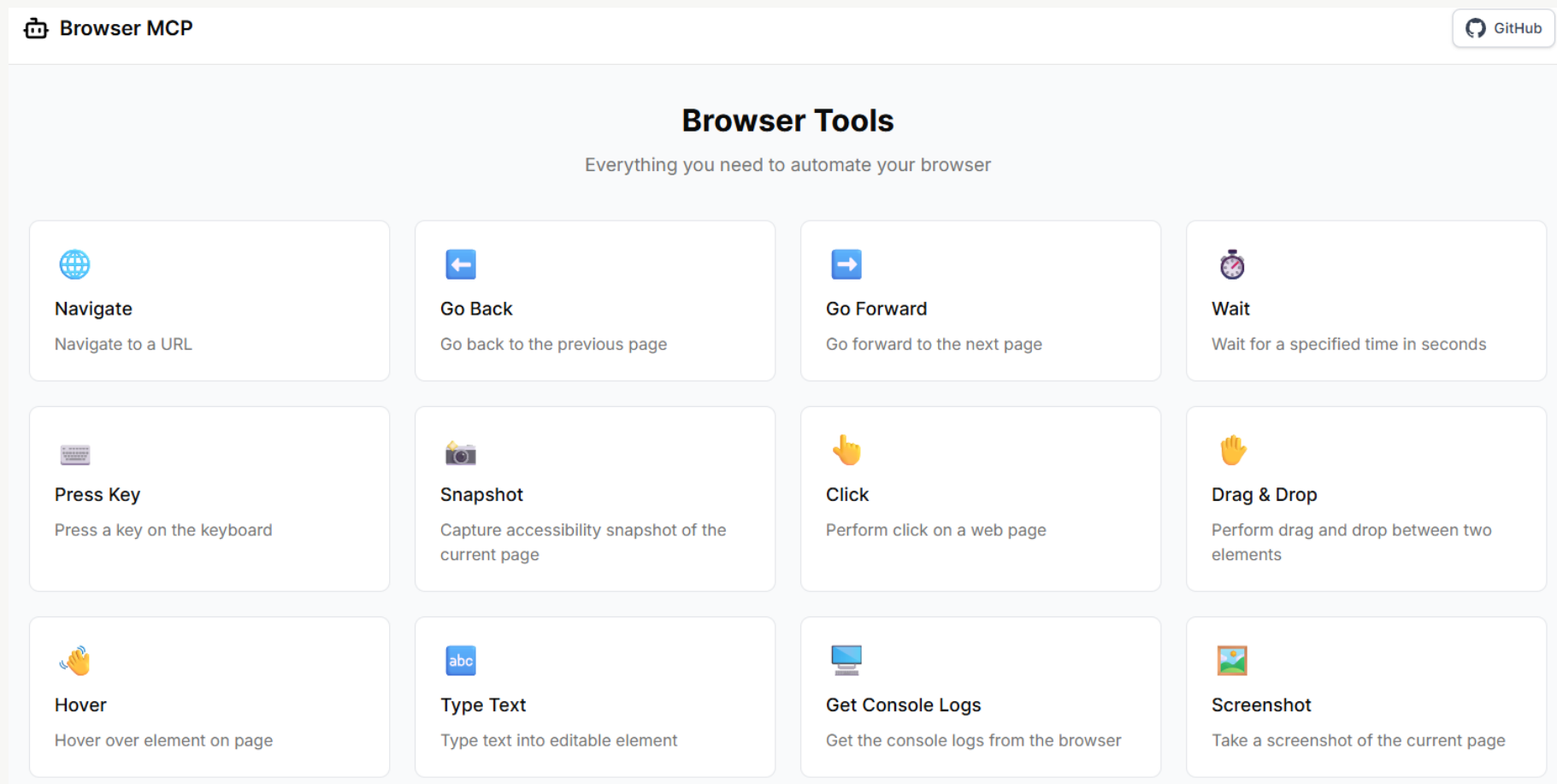
Resources expose data that clients can read — passive data sources rather than invocable functions.

```
@mcp.resource("data://config")
def get_config() -> dict:
    return {"theme": "dark", "version": "1.0"}
```

Prompts are reusable message templates that guide LLM interactions.

```
@mcp.prompt
def analyze_data(data_points: list[float]) -> str:
    formatted_data = ", ".join(str(point) for point in data_points)
    return f"Please analyze these data points: {formatted_data}"
```

Example: Web Browser



<https://browsermcp.io/>

Agent VS Human

Bot vs. Human worldwide

Bot (automated) vs. human HTTP requests distribution to HTML content



 **Cloudflare Radar**

Last 7 days | Sep 21, 2026, 07:45 UTC

MCP Market

<https://mcpmarket.com/>

36,974 Servers | Updated 1 hour ago

Find The Best MCP Servers

Directory of awesome MCP servers and clients to connect AI agents with your favorite tools.

All

Developer Tools

API Development

Data Science & ML

Productivity & Workflow

Analytics & Monitoring

>

Official MCP Servers

View all official servers →

**Firecrawl**

↗

Empowers LLMs with advanced web scraping capabilities for content extraction, crawling, and...

WEB SCRAPING & DATA COLLECTION ☆ 4.2k

**ElevenLabs**

↗

Enables interaction with Text to Speech and audio processing APIs for MCP clients.

API DEVELOPMENT ☆ 1.4k

**Magic**

↗

Generate modern UI components instantly from natural language descriptions within your IDE.

DEVELOPER TOOLS ☆ 5.2k

**Browserbase**

↗

Enables LLMs to control cloud browsers for web interaction, data extraction, and task automation...

BROWSER AUTOMATION ☆ 3.4k

**Exa**

↗

Enables AI assistants like Claude to perform web searches using the Exa AI Search API.

API DEVELOPMENT ☆ 4.6k

**Tavily**

↗

Integrates Tavily's search and data extraction capabilities with AI assistants via the Model Context...

WEB SCRAPING & DATA COLLECTION ☆ 2.1k

17

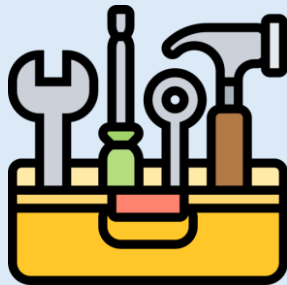
| Next Problem

Are these tools safe to execute?

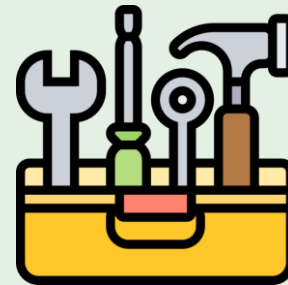
5 Sandbox

Tool Sandbox

- Environment isolation
- Trackable execution
- Recoverable to snapshot



Sandbox 1 (Isolated env)



Sandbox 2 (Isolated env)

Multi-Layer Sandbox

- 1. Running bash commands in the sandbox via /sandbox
- 2. Running the MCP server in a sandbox
 - {
 - "mcpServers": {
 - "risky-server": {
 - "command": "npx",
 - "args": ["@anthropic-ai/sandbox-runtime", "npx", "-y", "risky-mcp-server"]
 - }
 - }
 - }
- 3. Running the entire coding agent in a sandbox
 - `npx @anthropic-ai/sandbox-runtime claude`
- 4. Running OS in sandbox

Sandbox Discussion

- Every day coding on my own laptop
 - Bash sandbox via /sandbox
- Coding on shared servers
 - container/VM or sandbox runtime
- Trying out a random MCP server from the internet
 - wrap it with the sandbox runtime, or run in a VM

| Next Problem

How to reuse the workflow of using agents?

6 Skills

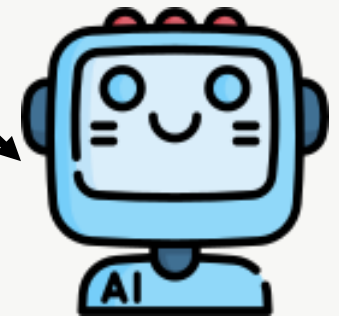
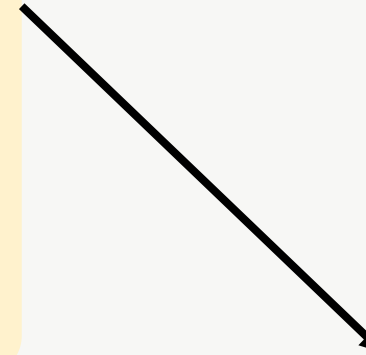


Debug this issue <issue description>:

1. Reproduce the failure locally
2. Collect stack traces
3. Identify the smallest failing component
4. Write a regression test
5. Implement the fix
6. Run the test check script

Rules:

- Never modify tests to hide failures
- Every bug fix requires a regression test



How do you
debug this issue?



Here is my prompt.

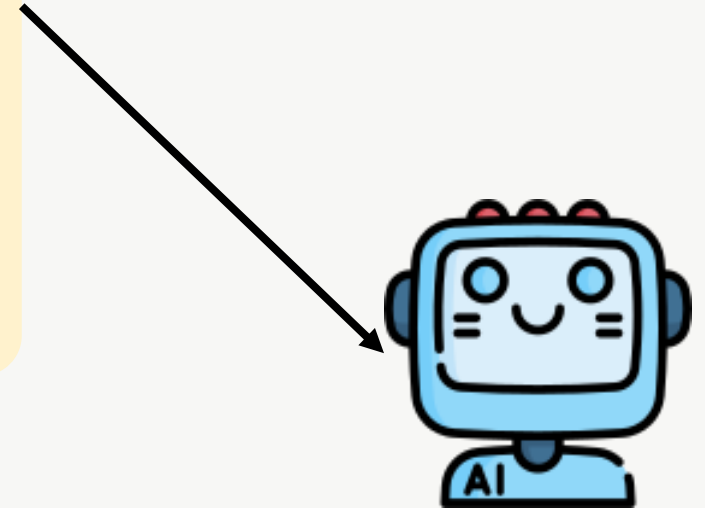


Debug this issue <issue description>:

1. Reproduce the failure locally
2. Collect stack traces
3. Identify the smallest failing component
4. Write a regression test
5. Implement the fix
6. Run the test check script

Rules:

- Never modify tests to hide failures
- Every bug fix requires a regression test





Debug this issue <issue description>:

1. Reproduce the failure locally
2. Collect stack traces
3. Identify the smallest failing component
4. Write a regression test
5. Implement the fix
6. Run the test check script

Rules:

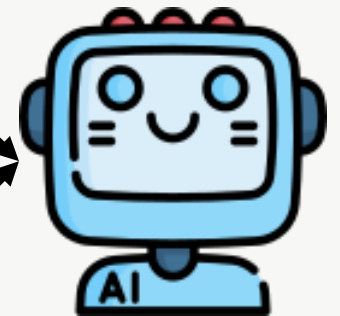
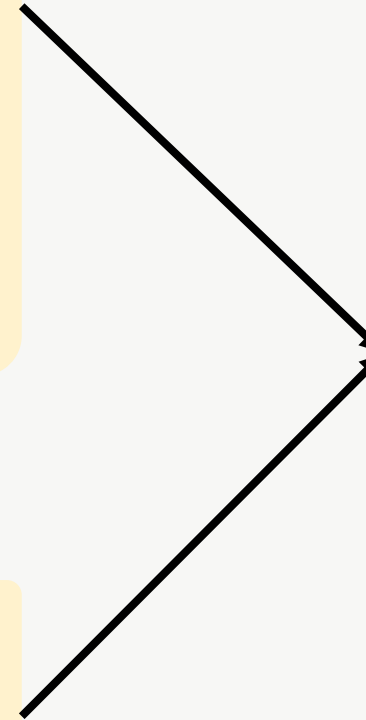
- Never modify tests to hide failures
- Every bug fix requires a regression test

Copy it



Debug this issue <issue description>:

...



Debug this issue <issue description>:

1. Reproduce the failure locally
2. Collect stack traces
3. Identify the smallest failing component
4. Write a regression test
5. Implement the fix
6. Run the test check script

Rules:

- Never modify tests to hide failures
- Every bug fix requires a regression test

Reusable workflow description: Skills

Memory?

Debug this issue <issue description>:

1. Reproduce the failure locally
2. Collect stack traces
3. Identify the smallest failing component
4. Write a regression test
5. Implement the fix
6. Run the test check script

Rules:

- Never modify tests to hide failures
- Every bug fix requires a regression test



MEMORY.md

- [user description]
- [feedback/preference]
- [project status]
- [reference doc]

It is not a global workflow polluting context.

Skills

- A record of experience (Standard Operating Procedure) of a certain type of task
- For example:
 - How to type log fields
 - How to query slow SQLs
 - How to write commit message



Skill Distillation

<https://titanwings.github.io/colleague-skill-site/gallery/>

<https://github.com/titanwings/colleague-skill>

COMMUNITY SKILLS

Skill Gallery

社区贡献的 dot-skill。他们离开了，但智慧还在。

+ 提交我的 Skill

查看提交模板

215

已收录 Skill

55

meta-skill

165

社区贡献者

搜索 Skill 名称、描述、标签...

全部

skill

meta-skill

共 215 个 Skill

d

dot-skill

meta-skill

把任何人蒸馏成一个 AI Skill —— 不止同事，伴侣、家人、老友、偶像、名人、小说角色，甚至...

other

蒸馏任何人

六层人格结构

titanwings

★ 18k

2026-04-13

求

求是 Skill

置顶

从实事求是的思想中整理出调查研究、战略判断等方法论工具集。用于问题分析、政策研判、...

other

务实

辩证

HughYau

★ 3.1k

2026-04-08

毛

毛选.skill

置顶

提炼《毛选》中的矛盾分析、根据地思维与战略判断框架。用于组织问题、竞争策略、逆境突...

other

战略思维

辩证法

leezythu

★ 787

2026-04-08

女

女娲.skill

meta-skill

从个人的心智模型、决策发式与表达模式中提炼可复用技能。把你独特的思维方式结构化，变...

other

创造性

结构化

alchaincyf

★ 21k

2026-04-08

K

Khazix Skills

"数字生命卡兹克"的 AI 工具箱，包含他长期打磨的 Prompts 与 Skills。当前开源两个 Skill：hv-...

other

深度研究

长文写作

KKKKhazix

★ 12k

2026-04-20

张

张雪峰.skill

蒸馏张雪峰的升学规划认知操作系统，基于 5 本著作、15+ 篇深度采访和 30+ 条一手语录，提...

other

就业倒推

普通家庭视角

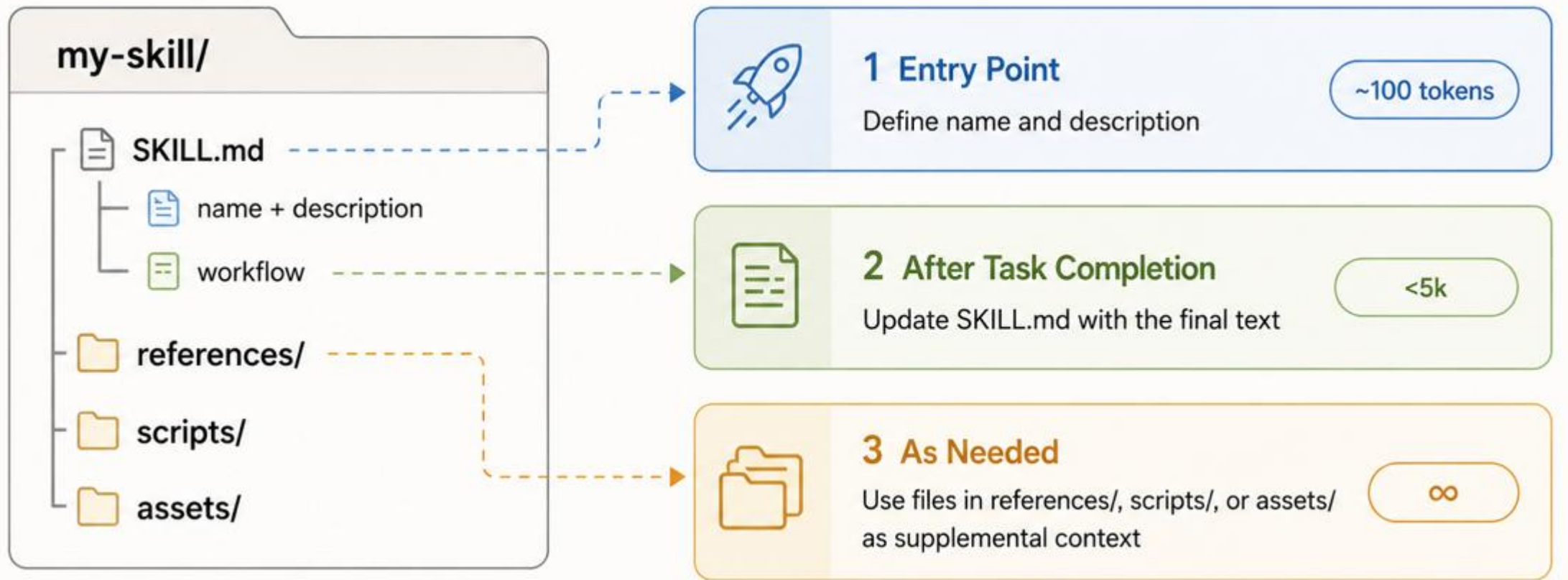
alchaincyf

★ 7.3k

2026-04-20

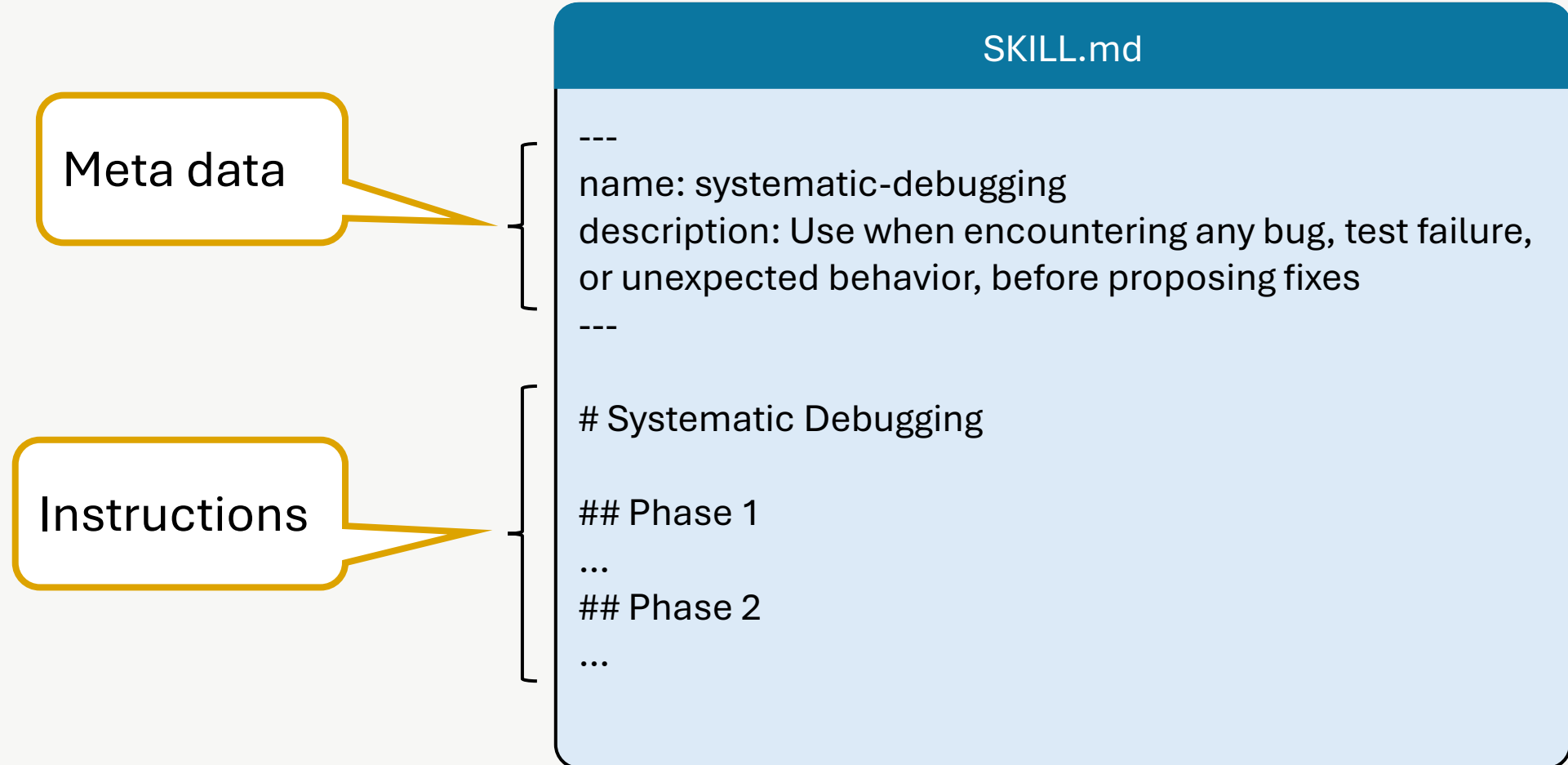
31

Skill.md



A skill = Prompts + scripts + other reference files

How to Trigger a Skill?



How to Trigger a Skill?

Debug this issue



SKILL.md

name: systematic-debugging

description: Use when encountering any bug, test failure, or unexpected behavior, before proposing fixes

Systematic Debugging

Phase 1

...

Phase 2

...

- If the prompt matches the metadata:
 - Load full instructions and follow them.

| How to Trigger a Skill?

Debug this issue



- If the prompt matches the metadata:

- Load full instructions and follow them.
- Execute scripts only when necessary.

Lazy load for saving context window

SKILL.md

```
---  
name: systematic-debugging  
description: Use when encountering any bug, test failure,  
or unexpected behavior, before proposing fixes  
---  
  
# Systematic Debugging  
  
## Phase 1  
...  
## Phase 2  
...
```

Example

[< Back to List](#)

From collection "superpowers"

systematic-debugging

by obra ☆ 128.7k 🏆 0 📄 GitHub

A structured methodology for debugging software issues that enforces root cause analysis before proposing fixes. It provides a four-phase process with concrete steps, common pitfalls to avoid, and integration with other skills like test-driven development. The skill prevents random guesswork and ensures systematic problem resolution.

LIVE LOGIC CIRCUIT

```
graph TD; START[START INPUT NODE  
Encounter bug, test failure, or unexpected behavior] --> CHECK[CHECK ERROR MESSAGES PROCESSING NODE  
Read error messages, stack traces, line numbers, file paths];
```

Grade S

9.2

AI SCORE

✓ Best Practices

Best Practices: A

Download Complete Skill

Copy SKILL.md only

> One-Click Install

▶ Try this Skill

♡ 21

USER RATING

Login to rate this Skill

USER RATING

👍 UP | 👎 DOWN

FILE EXPLORER

11

condition-based-waiting-example...

condition-based-waiting.md

CREATION-LOG.md

defense-in-depth.md

find-polluter.sh

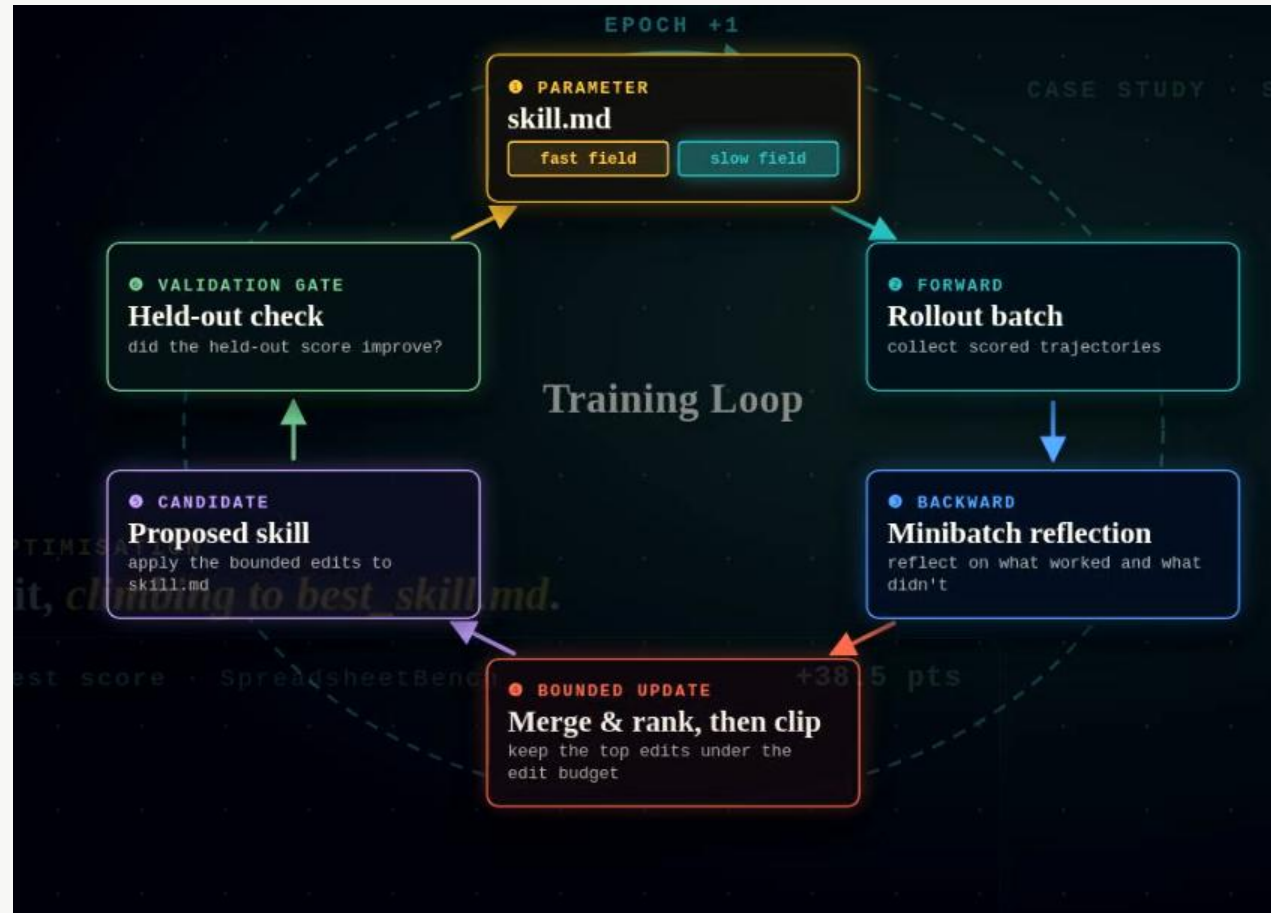
You can share your skills at:
<https://www.skillhub.club>

<https://www.skillhub.club/skills/obra-superpowers-systematic-debugging>

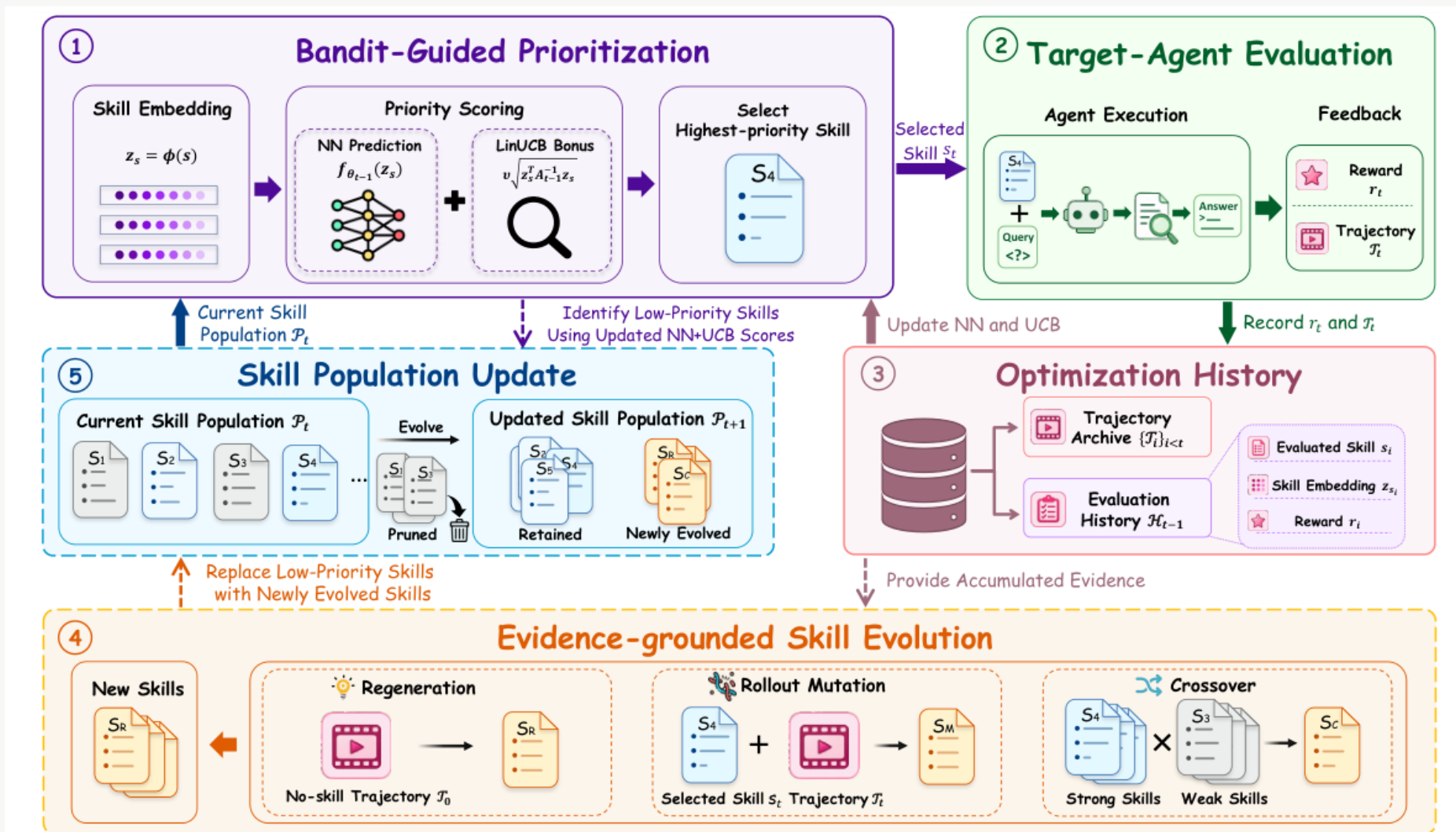
Let us build a simple Skill

Are your skills good enough?

SkillOpt: Optimize Your Skills



COBRA-Skills: Spend Evaluations Selectively



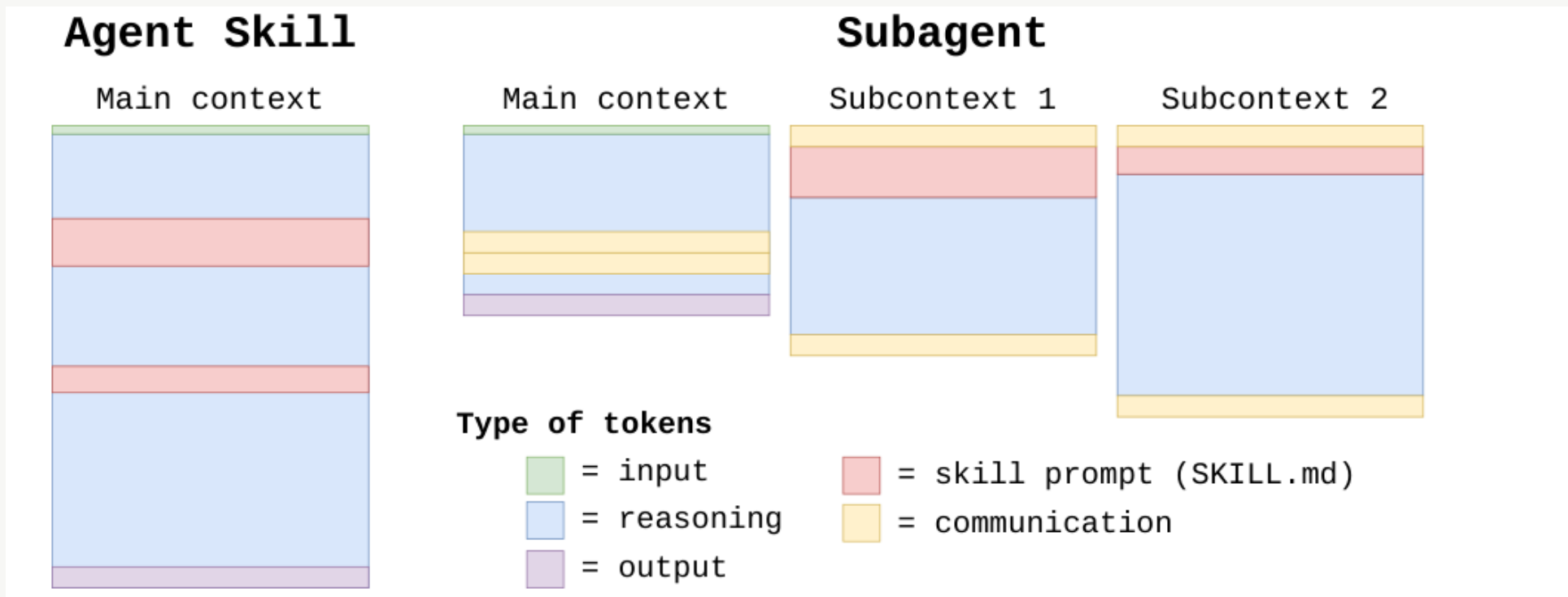
Skill Issue: Do optimized skills really help?

Table 1: Final test-set score and optimization cost per repository. *Spend* is the total agent API cost of the run.

repository	optimizer	Δ test-set	paired	spend	wall clock
JetBrains/koog	GEPA	+1	0.525	\$368.82	4.7 h
	SkillOpt	+2	0.546	\$401.44	8.14 h
kotest/kotest	GEPA	+3	0.554	\$182.48	4.9 h
	SkillOpt	+1	0.519	\$263.56	13.3 h
ktorio/ktor	GEPA	+4	0.567	\$312.87	13.0 h
	SkillOpt	-2	0.437	\$484.81	25.2 h
total				\$2,013.98	69.2 h

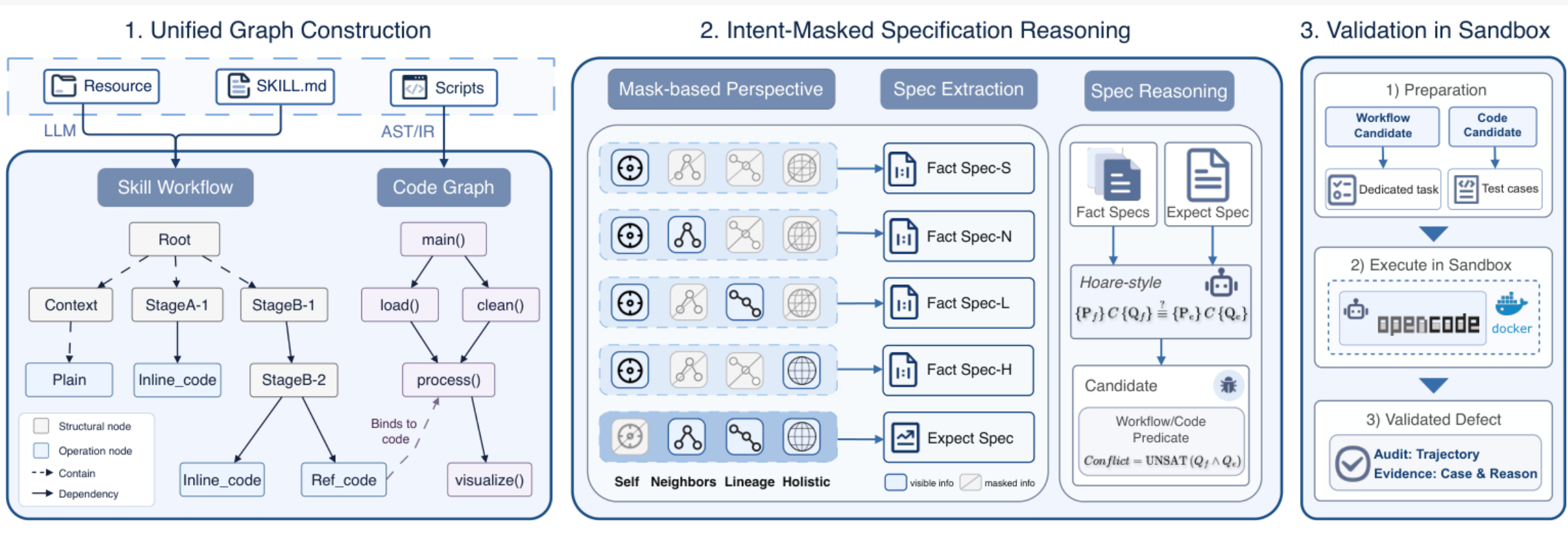
On three Kotlin repositories, observed skill-optimization gains remain difficult to distinguish from agent run-to-run variance.

Subagents vs Agent Skills: Invocation Matters



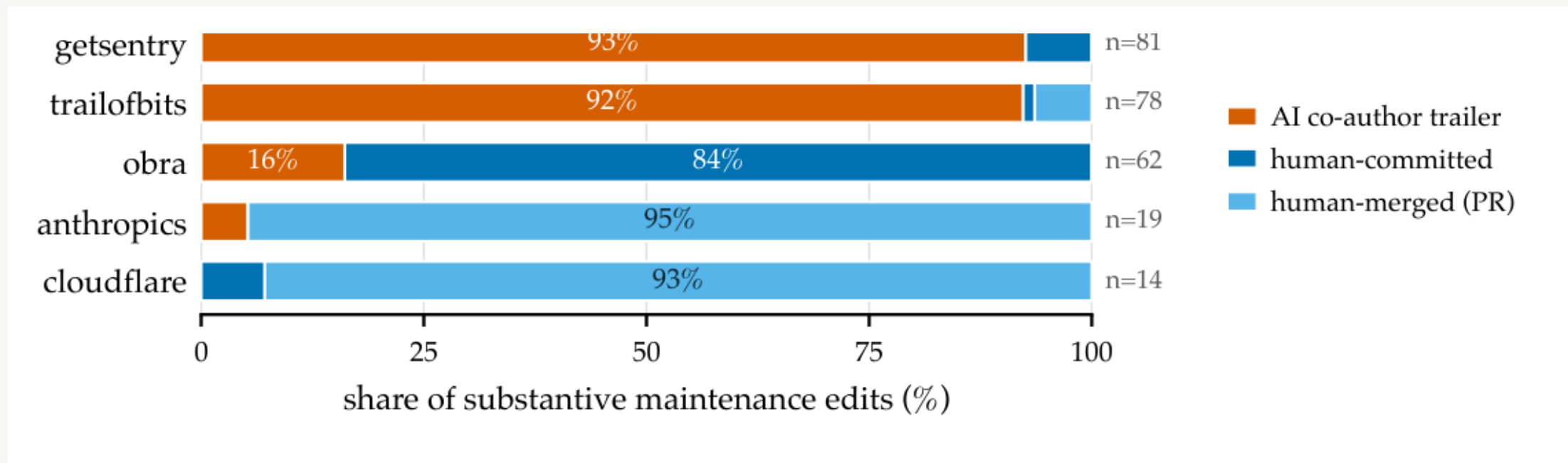
With clear I/O contracts, subagent execution can improve results, at the cost of extra communication tokens.

SkillSpec: Check Intent Against Behavior



SkillSpec compares declared intent with encoded behavior and validates suspected skill defects in an isolated sandbox.

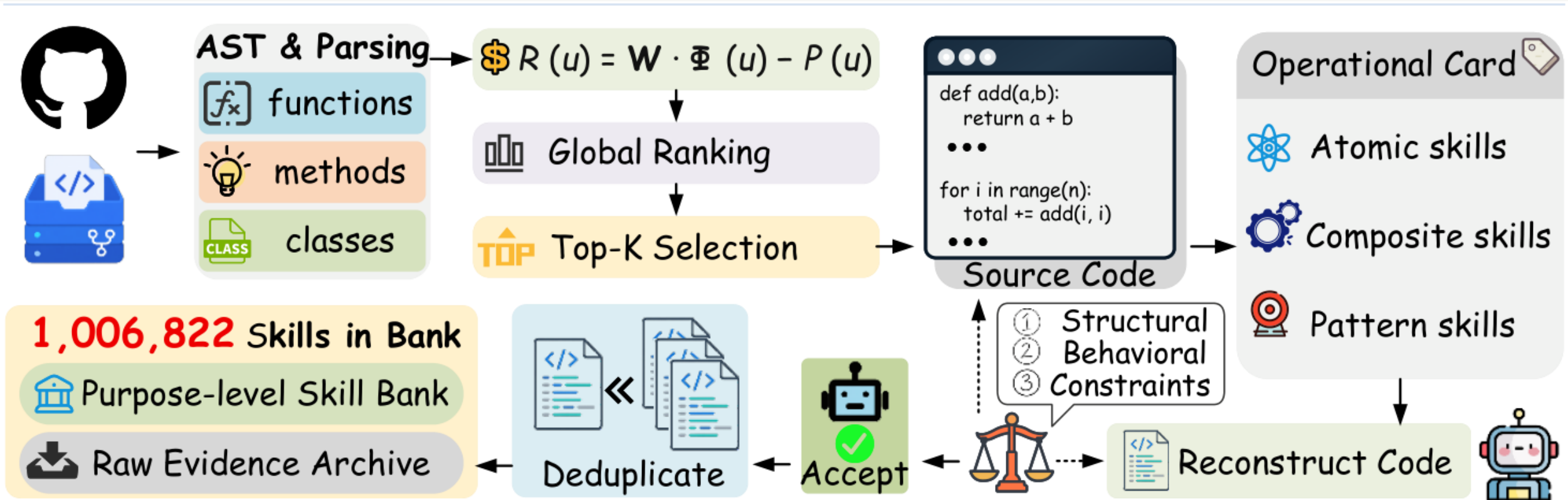
Who Maintains Agent Skills?



Original paper crop: Figure 2, PDF p. 7 | arXiv posted 04 Sep 2026

Across five repositories, all 254 substantive skill edits were human-authored or human-merged, while 62% carried AI co-author trailers.

Mine Reusable Procedures From Code



Code2Skill builds a retrievable skill bank from source code, using reconstruction and source-aware checks.

Evening, Marina

How can I help you today?



Chat

Cowork



Add files or photos



Record a skill



Skills



Connectors



Plugins



Record a skill

Your screen, clicks, typing, and voice are recorded, then sent to Claude and turned into a repeatable skill.

Cancel

Start recording

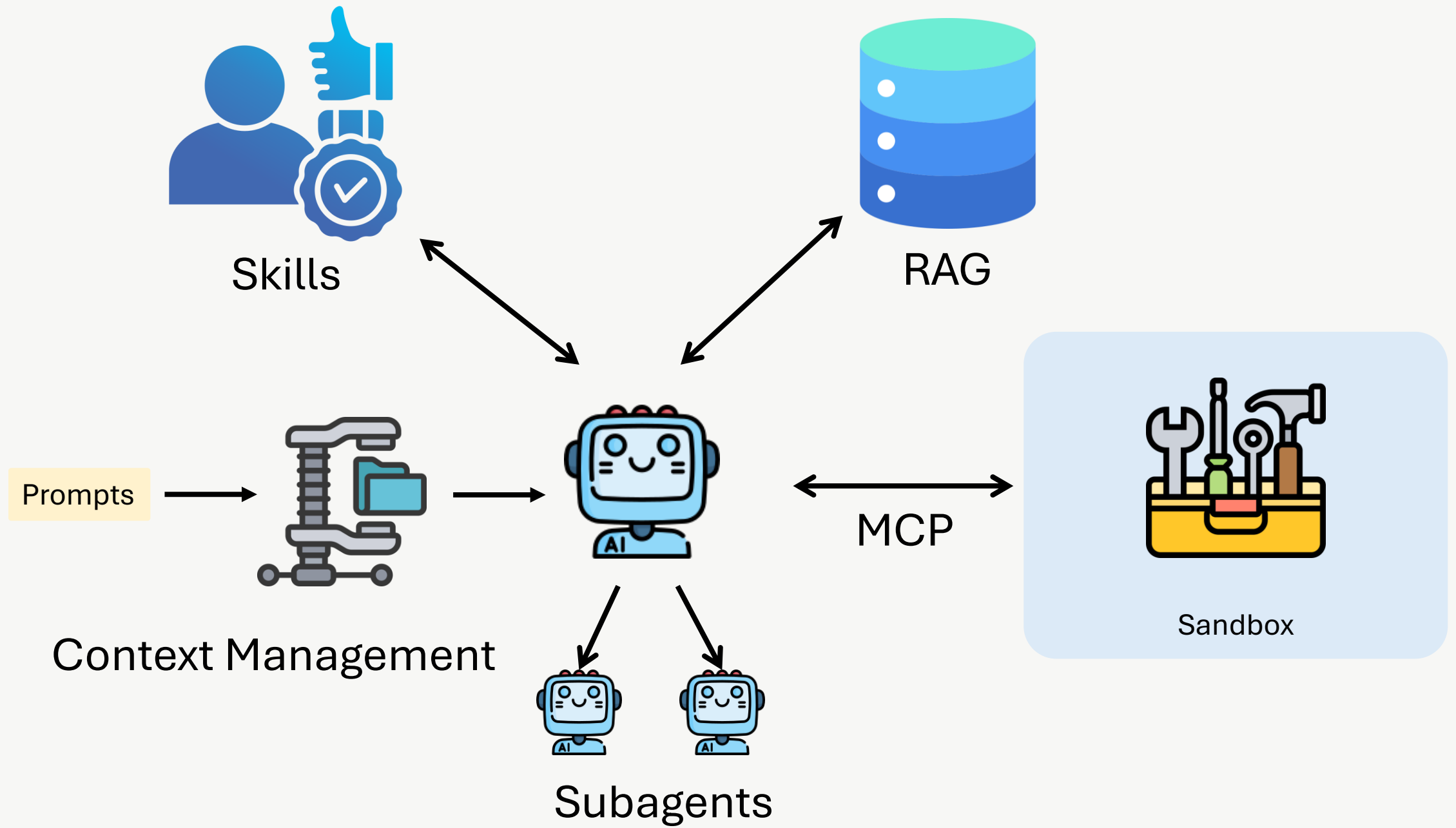


Recorded demonstration · 23.4s



Saved as /file-expenses





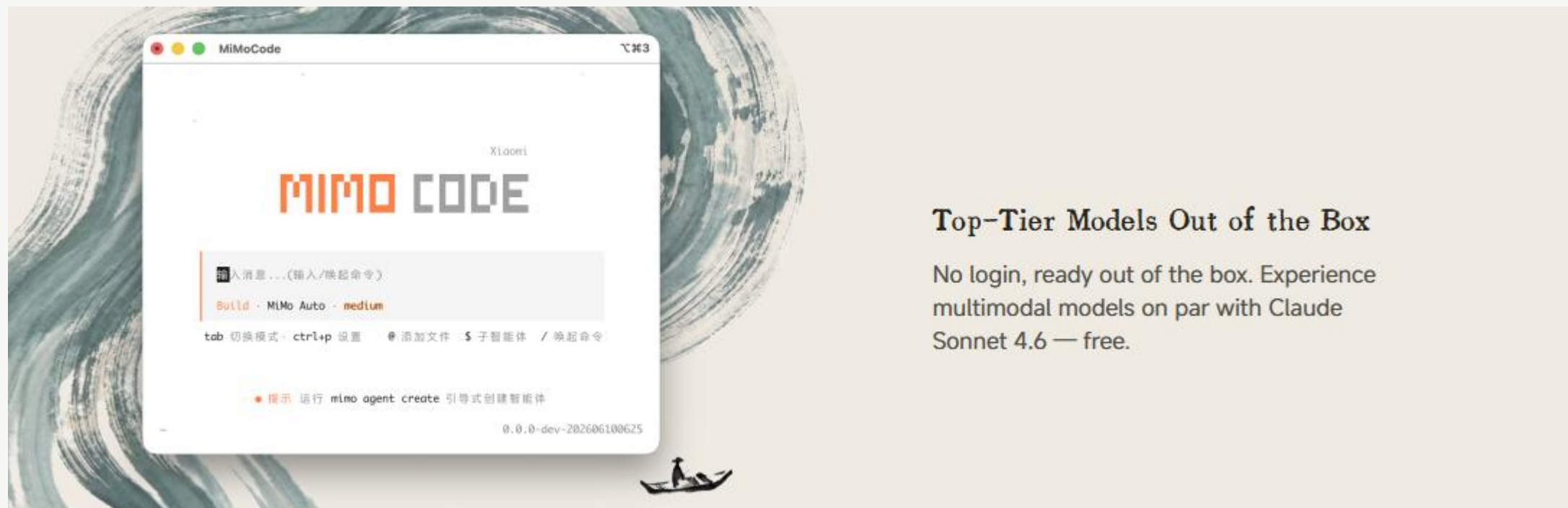
| All Agents Do is Prompt Organization

- Model system prompts
 - Not visible
- Project system prompts
 - Claude.md in each project folder
- Memory/Skills
 - Extended capabilities
- User prompts
 - Your inputs

| Other Agents

- Coding agents are major agents
- Share similar architectures with other agents

MiMo Code



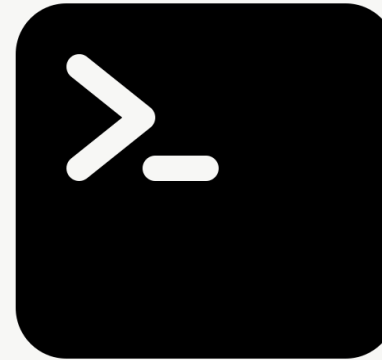
Top-Tier Models Out of the Box

No login, ready out of the box. Experience multimodal models on par with Claude Sonnet 4.6 — free.

| Other Coding Agents

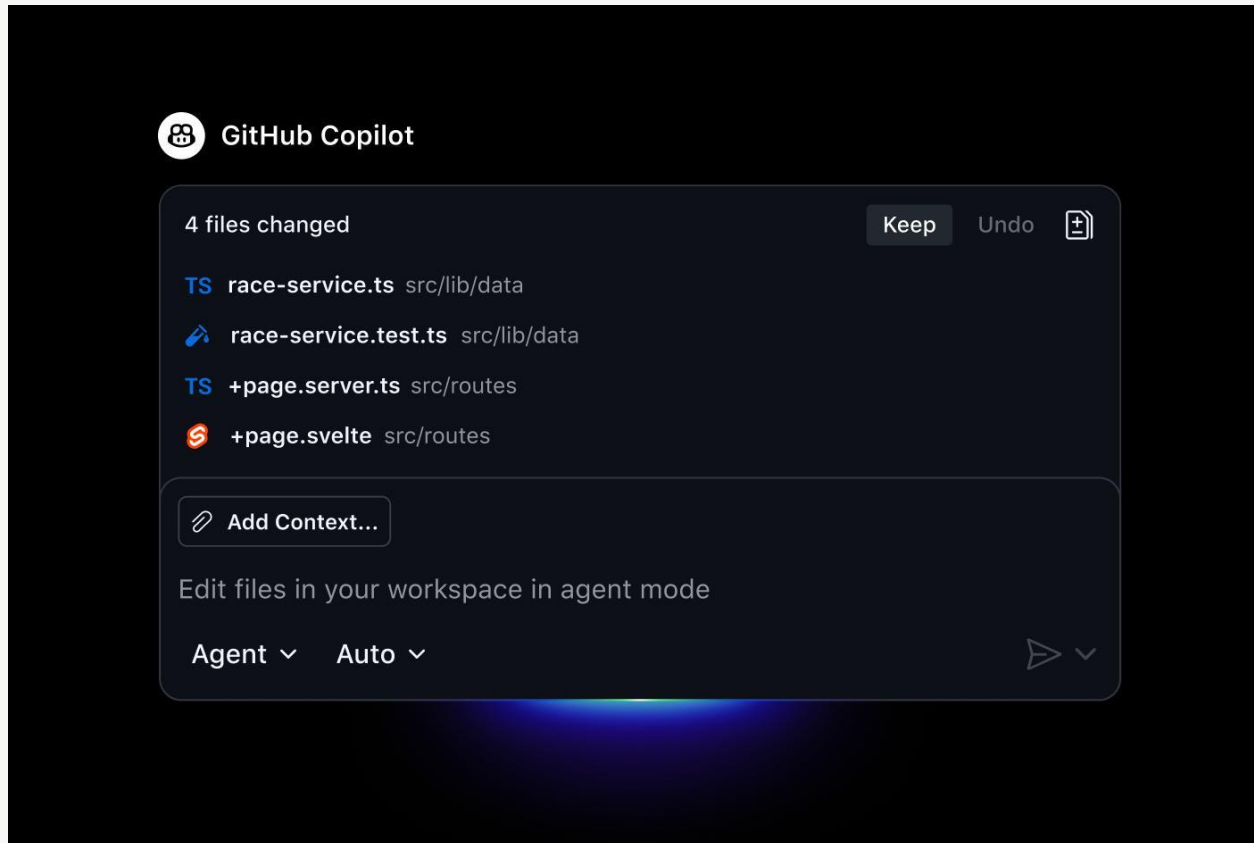


Visual Studio Code
extension



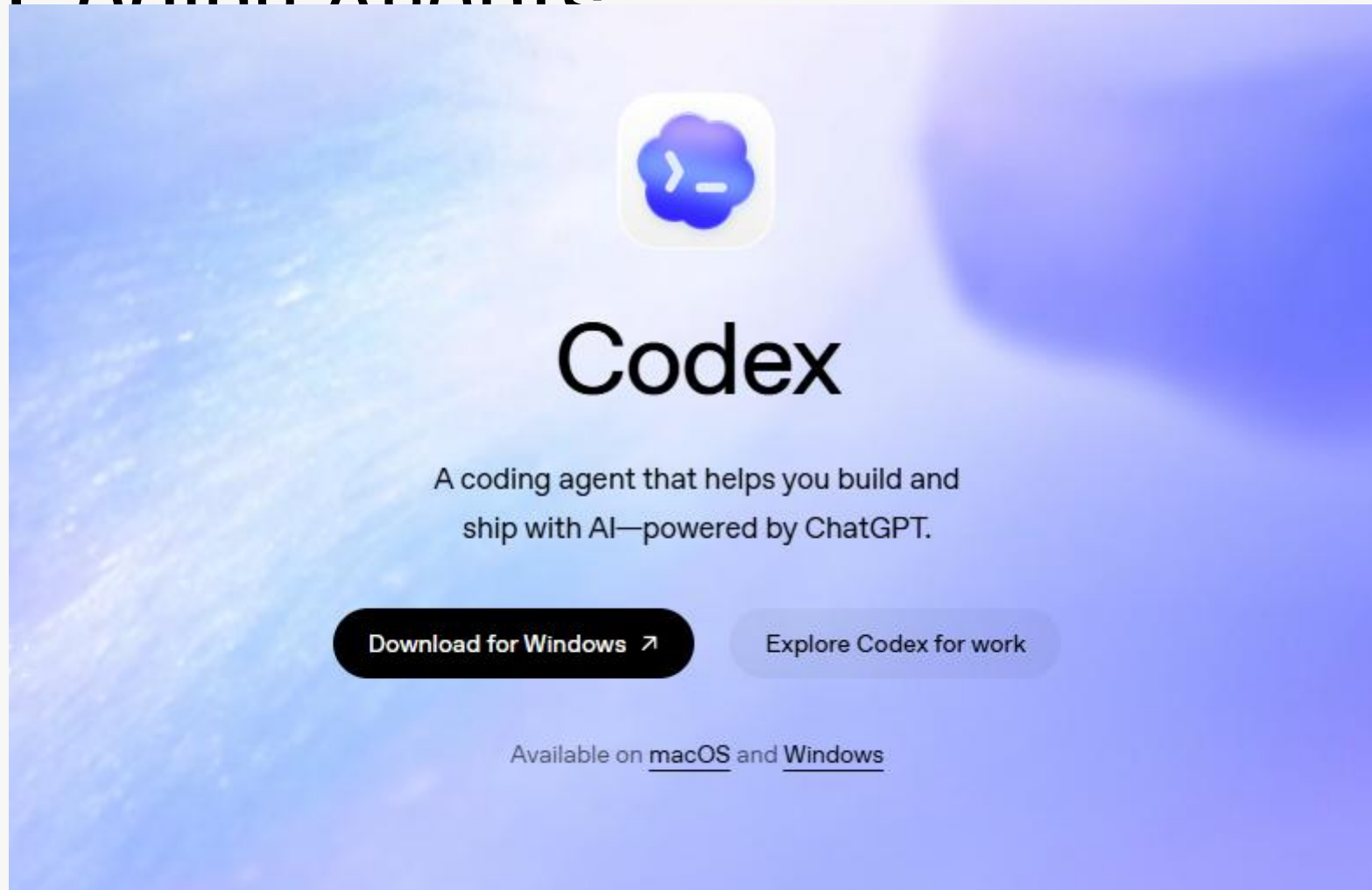
Terminal

Other Coding Agents



First-generation coding agent by GitHub.
<https://github.com/copilot>

| Other Coding Agents



Coding agent by OpenAI.
<https://openai.com/codex/>

Other Coding Agents



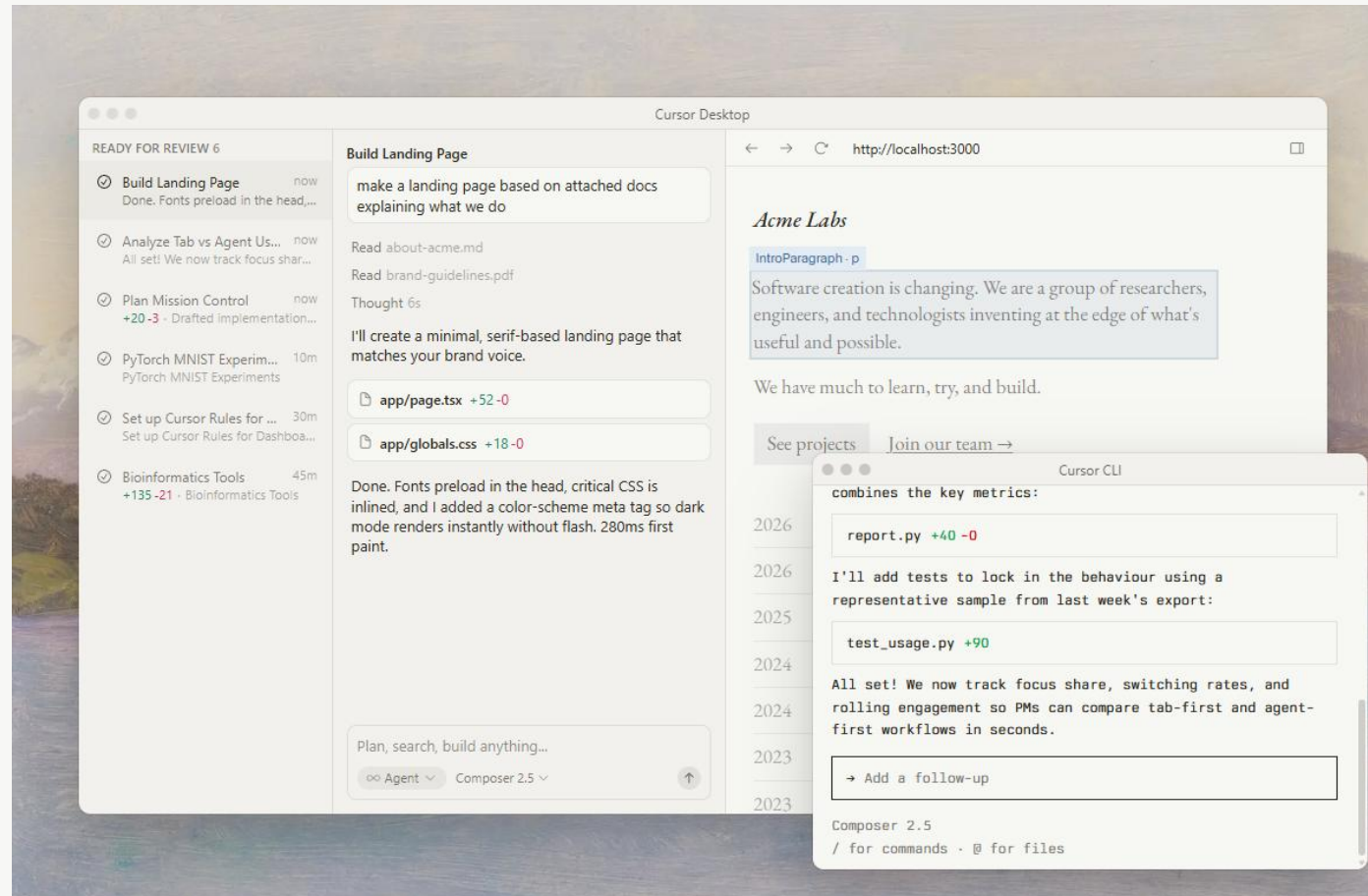
Coding agent by Google.
<https://geminikli.com/>

Other Coding Agents



Popular open-source coding agents.
<https://opencode.ai/>

Other Coding Agents



Cursor, a fork of VSCode with AI coding support.
<https://cursor.com/>

Project Announcement



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Questions

Next class: How to write a prompt