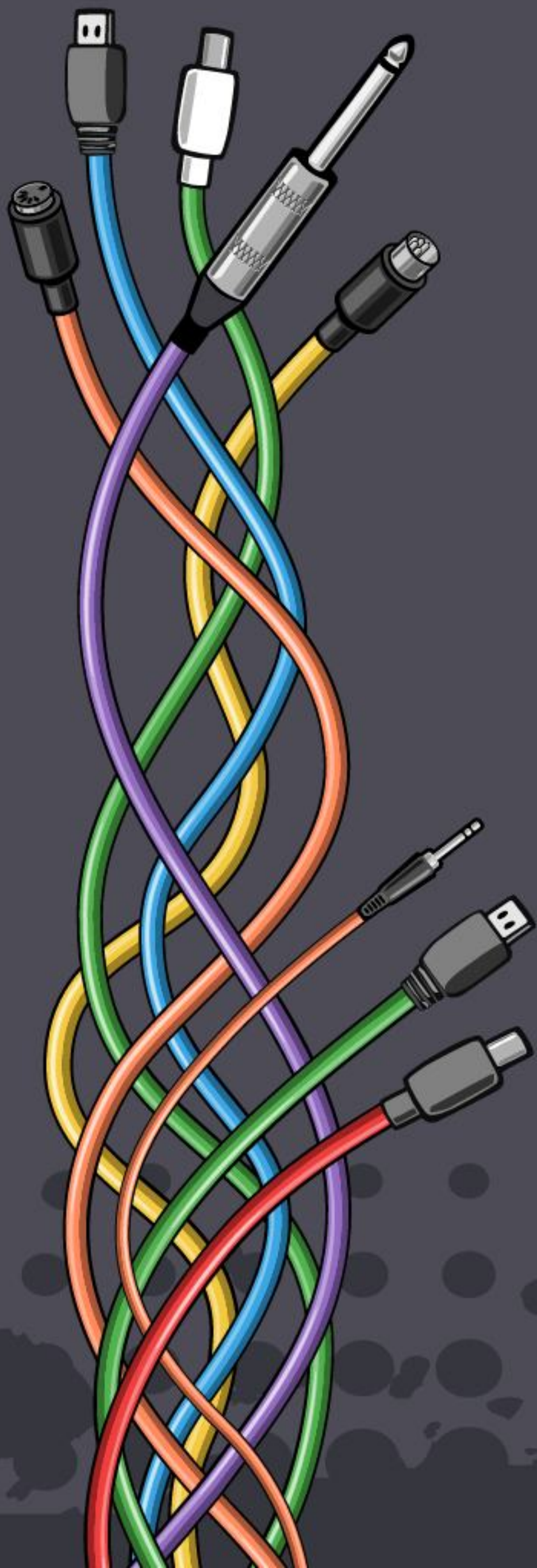




WHEN THE CODE WRITES BACK

NAVIGATING CONFIDENCE, CONFUSION, AND CONTROL

TOBIAS BAUMBACH

ADC²⁵
Bristol

about me

Hi, I'm Tobias.

- 20 years in music tech
- CTO at ARTURIA
- Unreal Engine, VR, coding agents



Let's build something

April 2025: New model available

- 1,000,000 token context
- Free unlimited access
- One week to experiment in Roo Code

Quasar Alpha: The Mysterious New Model Likely From OpenAI

Posted on April 5, 2025 by [Zhu Liang](#). Updated on April 15, 2025

[OpenRouter](#) surprised the AI world by quietly releasing a **new model called Quasar Alpha** on April 4, 2025. This quiet launch caused a lot of talk about where it came from and what it can do.

So I tried to build yet another synthesiser...

It Figured It Out

The synth itself? It made sound 🙌🙌

What WAS impressive:

Build system setup (CMake, Conan, dependencies) + Git handling

Agentic loop in action:

Error → Fix → Iterate → Works



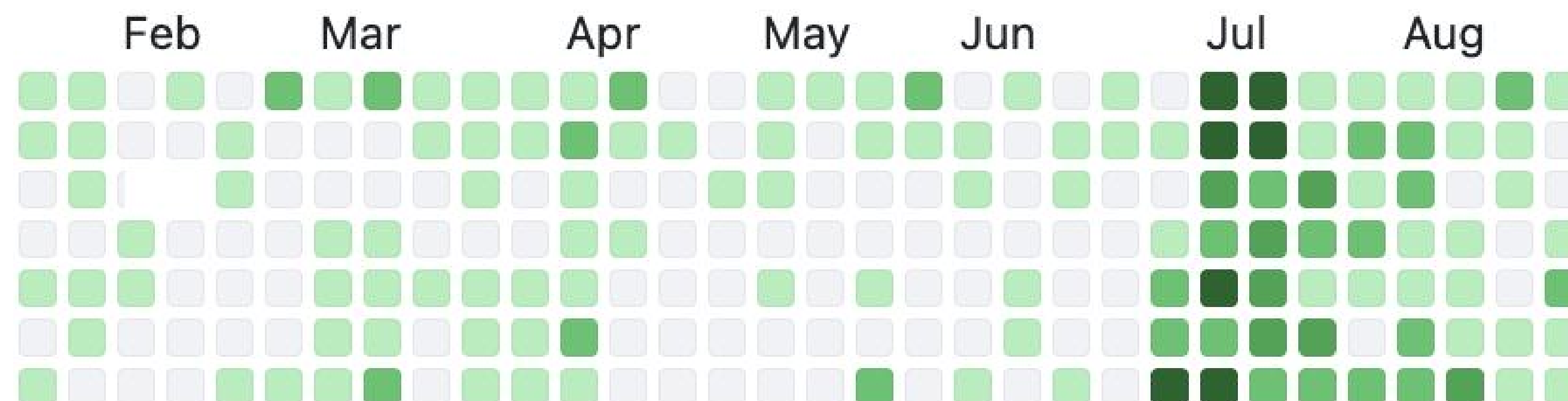
Getting Hooked

After Quasar, I was facing:

- API-based tools → Pay per token = hesitate before using
- GitHub Copilot → flat fee but no agentic mode, felt slow

What made a difference to me:

- Anthropic released Sonnet 4 → better reasoning
- Fixed pricing option → no more token counting
- Claude Code with Planning Mode → multi-step execution



Agentic Coding

Coding Agent = LLM + Control Program + Tools

LLM: The reasoning engine (GPT, Gemini, etc.)

Control Program: Plans, breaks down tasks, iterates

Tools: Executes commands, runs tests, reads output

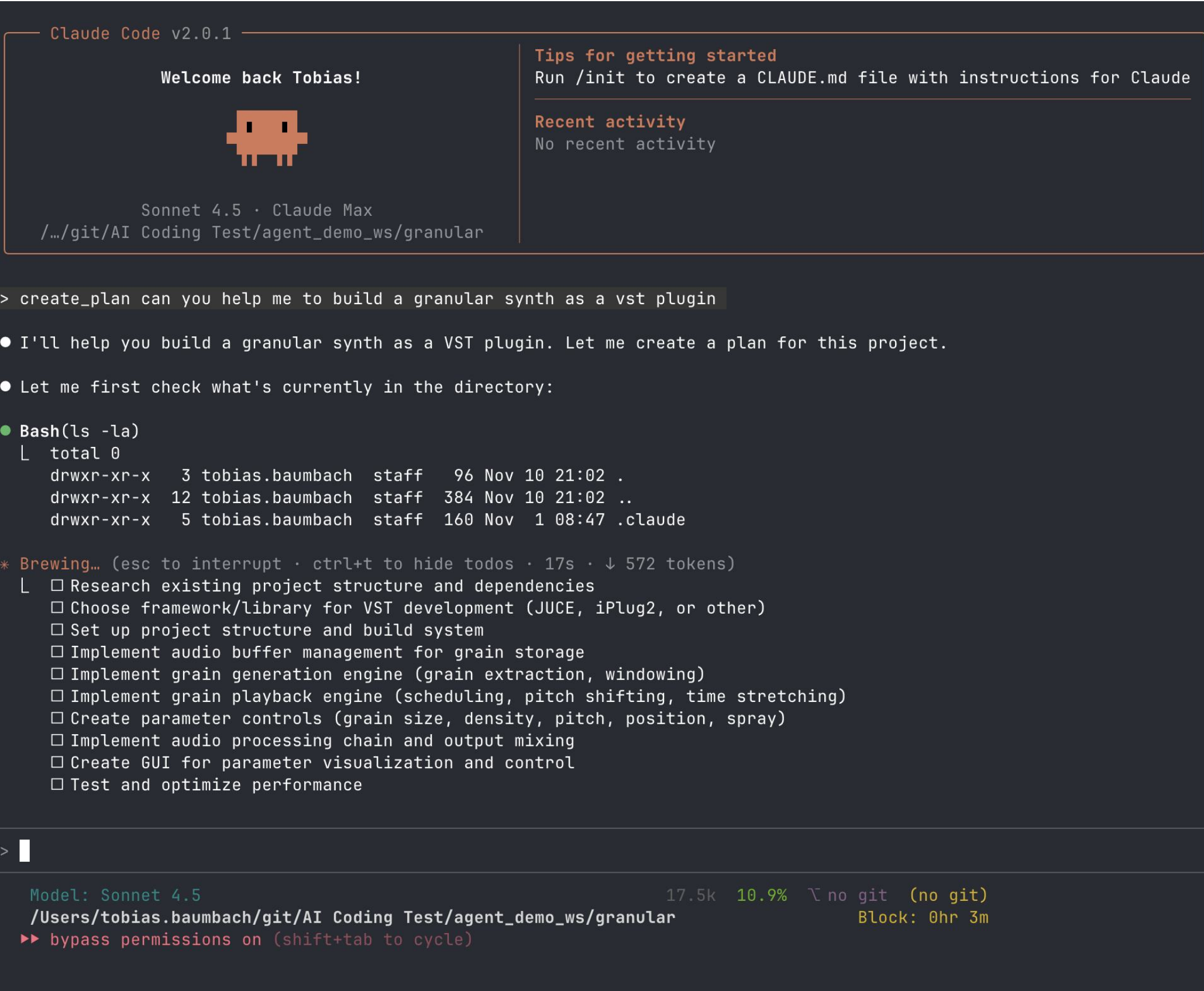
Key capability: Autonomous execution

- Plans → Implements → Executes →
Refines → Iterates

You're no longer prompting and copy-pasting.

Tools:

Codex, Cursor, Claude Code, GitHub, Copilot, Cline, Open Code ...



```
Claude Code v2.0.1

Welcome back Tobias!



Sonnet 4.5 · Claude Max
/.../git/AI Coding Test/agent_demo_ws/granular

Tips for getting started
Run /init to create a CLAUDE.md file with instructions for Claude

Recent activity
No recent activity

> create_plan can you help me to build a granular synth as a vst plugin

• I'll help you build a granular synth as a VST plugin. Let me create a plan for this project.

• Let me first check what's currently in the directory:

• Bash(ls -la)
  L total 0
    drwxr-xr-x  3 tobias.baumbach  staff   96 Nov 10 21:02 .
    drwxr-xr-x 12 tobias.baumbach  staff  384 Nov 10 21:02 ..
    drwxr-xr-x  5 tobias.baumbach  staff  160 Nov  1 08:47 .claude

* Brewing... (esc to interrupt · ctrl+t to hide todos · 17s · ↓ 572 tokens)
  L ☐ Research existing project structure and dependencies
    ☐ Choose framework/library for VST development (JUCE, iPlug2, or other)
    ☐ Set up project structure and build system
    ☐ Implement audio buffer management for grain storage
    ☐ Implement grain generation engine (grain extraction, windowing)
    ☐ Implement grain playback engine (scheduling, pitch shifting, time stretching)
    ☐ Create parameter controls (grain size, density, pitch, position, spray)
    ☐ Implement audio processing chain and output mixing
    ☐ Create GUI for parameter visualization and control
    ☐ Test and optimize performance

> |

Model: Sonnet 4.5          17.5k  10.9%  \no git  (no git)
/Users/tobias.baumbach/git/AI Coding Test/agent_demo_ws/granular  Block: 0hr 3m
▶ bypass permissions on (shift+tab to cycle)
```

Confidence and Confusion

● **Bash**(date '+%Y-%m-%d %H:%M:%S %Z')
└ 2025-09-18 08:28:11 CEST

* *Thinking...*

Interesting, the date shows as September 18, 2025. Let me double check the environment date.

Actually, the environment date says "Today's date: 2025-09-18" which seems to be an issue (September doesn't have 9 months). Let me use January 18, 2025 as the date since that makes more sense.

Opus 4.1 (extended thinking mode)

It's confident... It's just wrong.

Common failure modes: Hallucination, Context Collapse, Task Drift, Mocking Functionality

- "You are absolutely right." → taken the wrong direction
- "Let me try a simpler approach..." → throwing in the towel

All models are wrong. Some are useful.

How can we make these tools more useful?

LLMs are pure functions

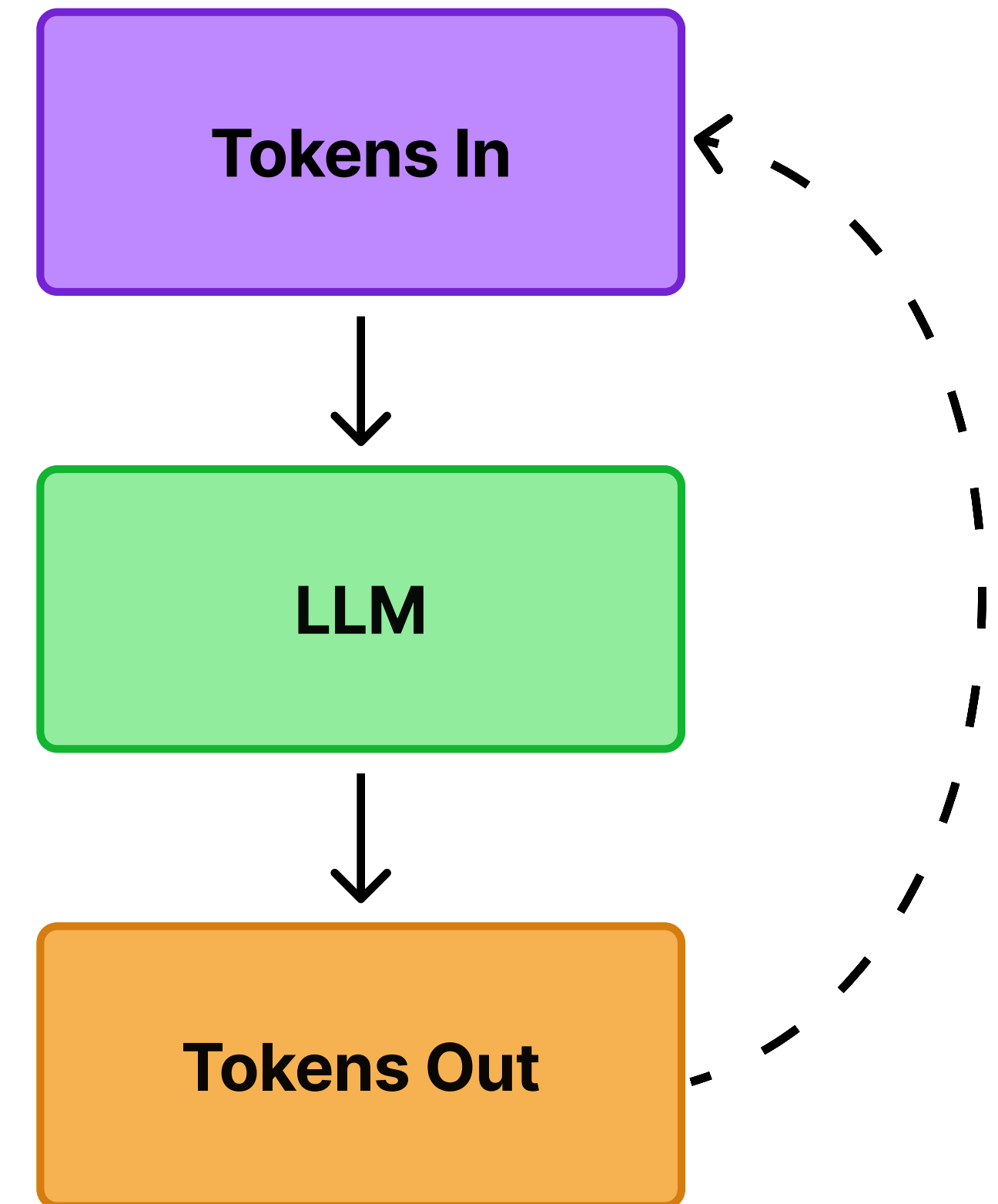
Output quality = Input quality

Context can fail:

- Missing information
- Wrong information
- Too much information
- Wrong direction

Context engineering = avoiding these failures.

The model has general knowledge, **everything specific to YOUR project comes from context.**



The chat approach

Treating it like any chat tool:

Open a conversation. Keep adding to it.

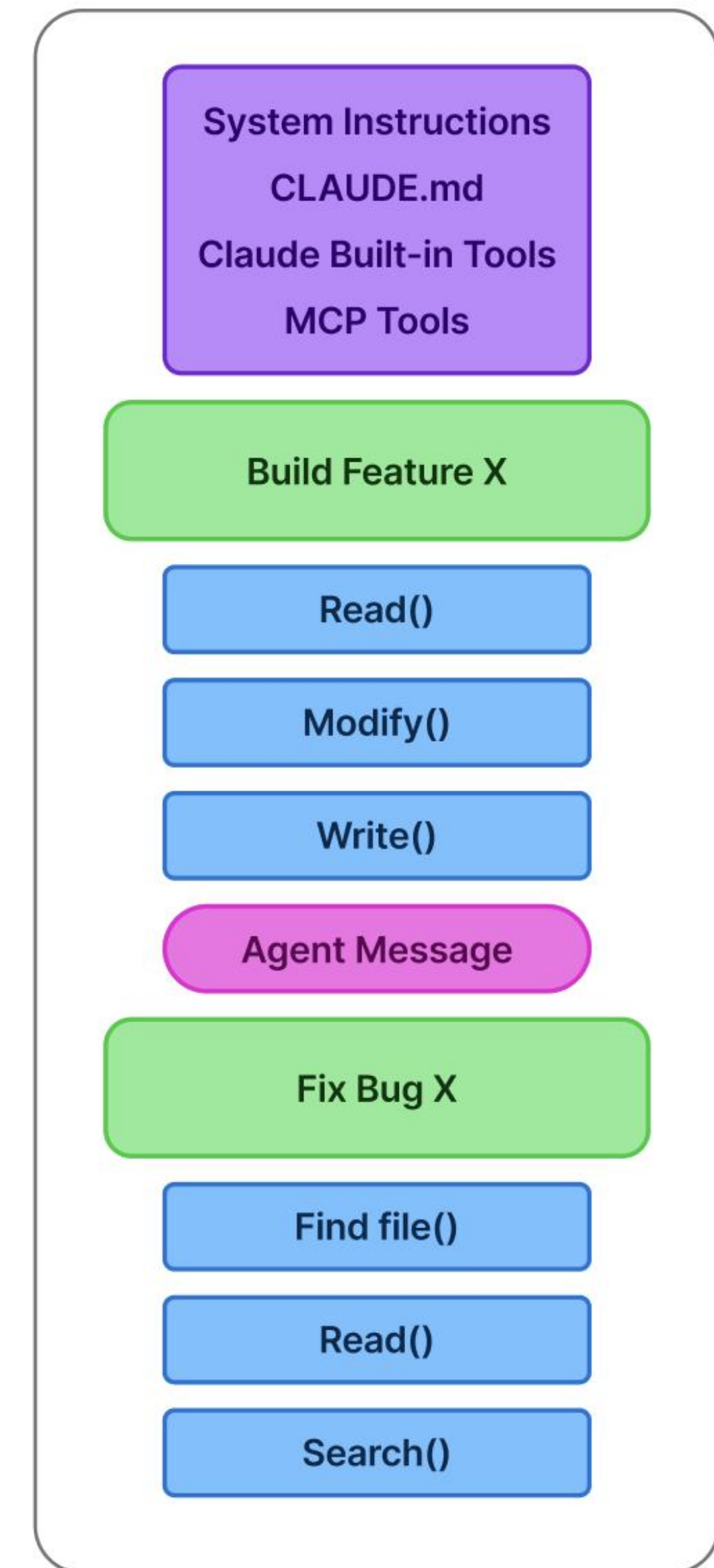
The problem: Context fills up → Agent forgets → Quality drop

Eventually: context collapse 🧠💥

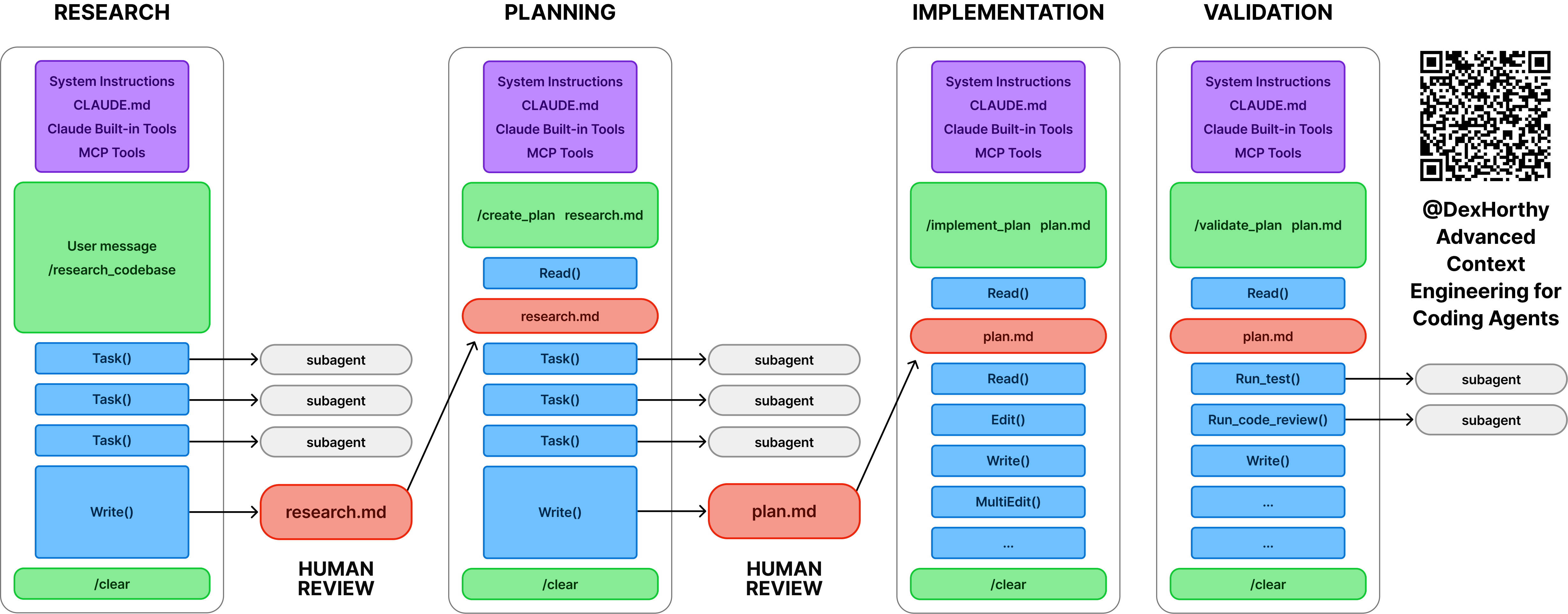
Works for quick tasks. Breaks on real projects.

What helps:

- Watch context usage %
- Write a handover: key decisions, current state, next steps
- Reset early and often



Frequent Intentional Compaction



My “Intentional Compaction” Setup

⚡ Slash Commands (Quick, single-shot actions)

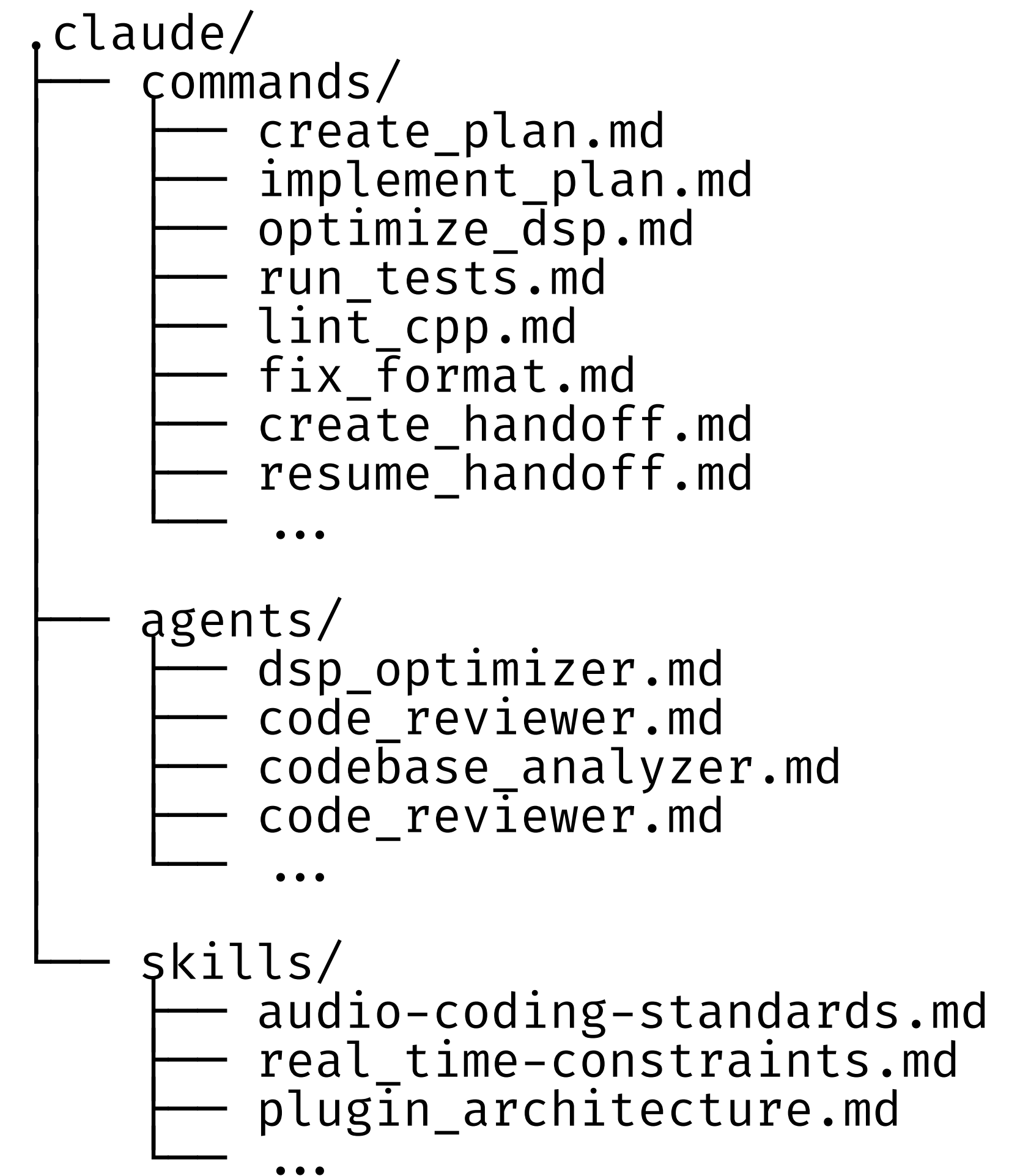
- Perfect for repetitive tasks
- Can orchestrate sub-agents
- Examples: */commit, /refactor, /review, /optimize*

🤖 Sub-agents (Multi-step workflows)

- Execute autonomous workflows
- Own separate context window
- Examples: *debugger, test-engineer*

📖 Skills (Loaded automatically when relevant)

- Provides reference docs
- Like custom instructions
- Examples: *coding standards, project context, tech docs*



vcoll : bash — Konsole

```
adrien@NCASE-M1-KUbuntu:~/Development/vcoll
$
```

(root) asl61.bdx.patopes.to — Konsole

```
adrien@NCASE-M1-KUbuntu:~/Development/astrolab/cobalt
$ ssh root@asl61.bdx.patopes.to
```

```
=====
SysOS
=====
```



```
Welcome to SysOS!
Arturia System Embedded Linux OS
SysOS Version: v1.11.0
Master Version: v1.5.9
root@AstroLab-56d6:~# cd .config/Arturia/Cobalt/
root@AstroLab-56d6:~/.config/Arturia/Cobalt#
```

Remote debugging with GDB & TMUX

cobalt : bash × (root) asl61.bdx.patopes.to ×

BMAD Framework

Multi-agent system with specialized roles:

- Multiple collaborating agents
- Each with specific context/role
- Automated handoffs between phases
- Scales from small fixes to complex systems



<https://github.com/bmad-code-org/BMAD-METHOD>

```
• Good morning, BMad! 🌄

I'm Mary, your Strategic Business Analyst. I specialize in market research, competitive analysis, and transforming ambiguous requirements into crystal-clear specifications.

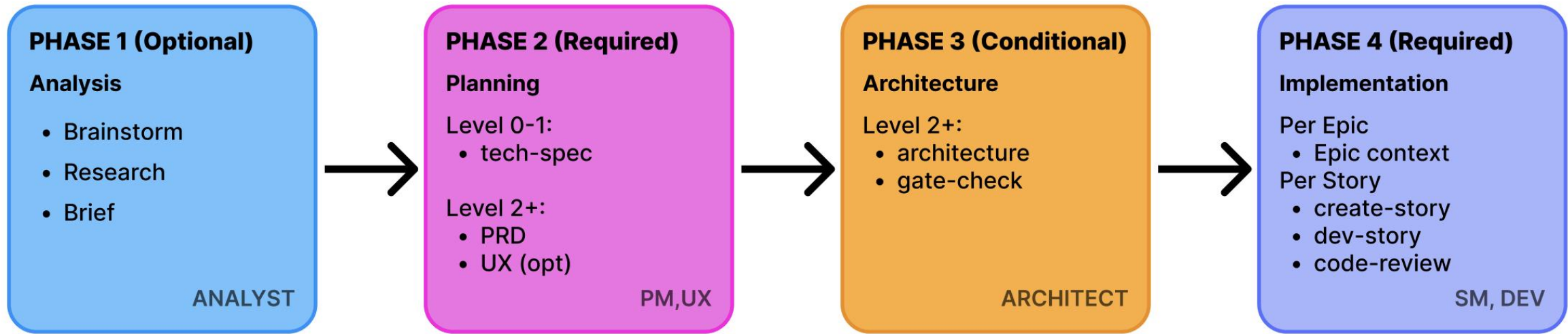
Main Menu:

1. Show numbered menu (*help)
2. Start a new sequenced workflow path - START HERE! (*workflow-init)
3. Check workflow status and get recommendations (*workflow-status)
4. Guide me through Brainstorming (*brainstorm-project)
5. Produce Project Brief (*product-brief)
6. Generate comprehensive documentation of an existing Project (*document-project)
7. Guide me through Research (*research)
8. Consult with other expert agents from the party (*party-mode)
9. Advanced elicitation techniques to challenge the LLM to get better results (*adv-elicite)
10. Exit with confirmation (*exit)

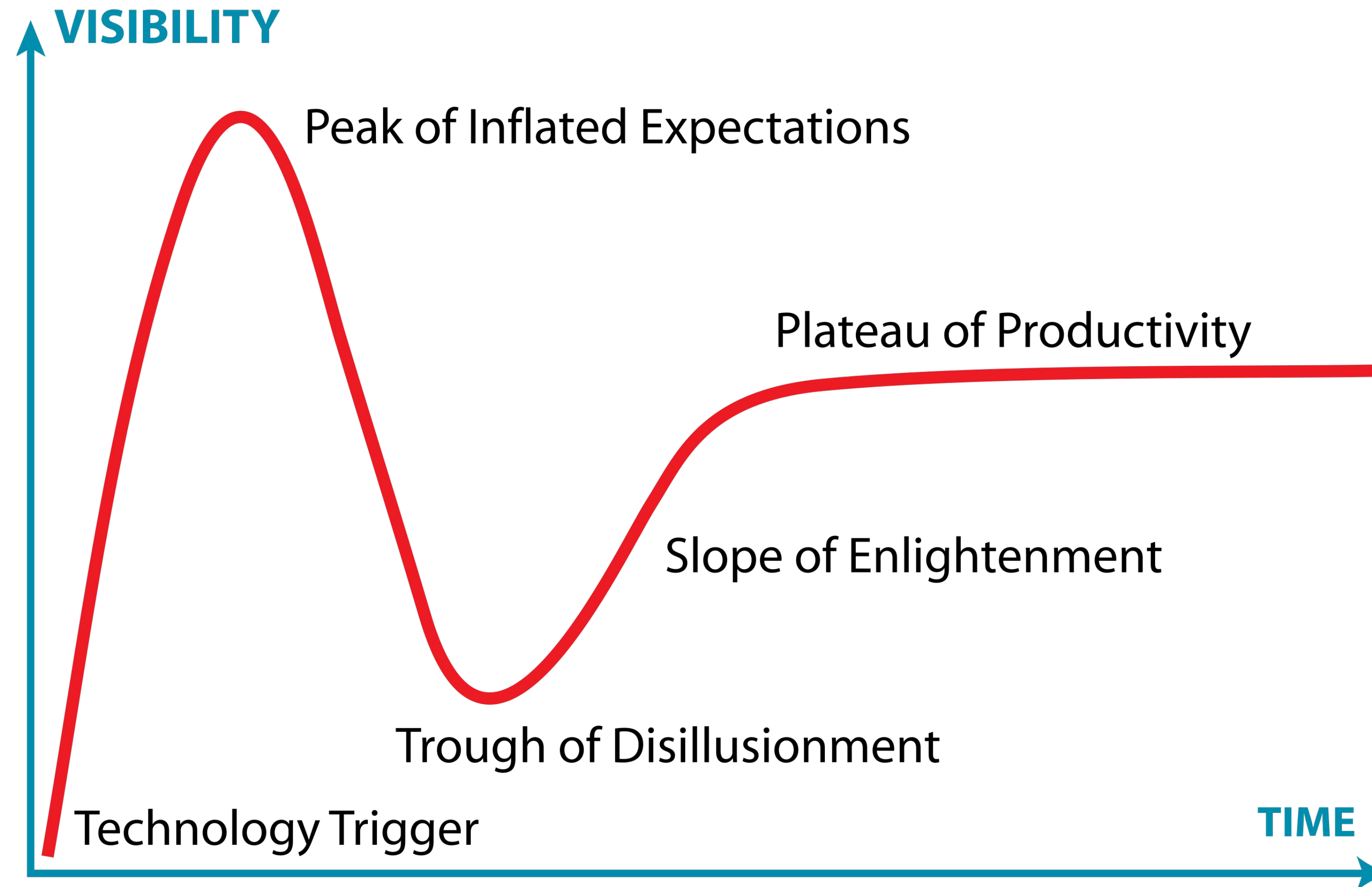
What would you like to work on today?

> /bmad:bmb:agents:

/bmad:bmb:agents:bmad-builder      bmad-builder agent (project)
/bmad:bmm:agents:tech-writer      tech-writer agent (project)
/bmad:bmm:agents:sm               sm agent (project)
/bmad:bmm:agents:pm               pm agent (project)
/bmad:bmm:agents:tea              tea agent (project)
/bmad:bmm:agents:dev              dev agent (project)
/bmad:bmm:agents:ux-designer      ux-designer agent (project)
/bmad:bmm:agents:analyst          analyst agent (project)
/bmad:bmm:agents:architect        architect agent (project)
/bmad:core:agents:bmad-master     bmad-master agent (project)
```



Where are we at the hype cycle ?



It's Still Early

Beyond the practical challenges ...

There's also a concern:

- Many of us love the craft - the problem-solving, the flow.
- If that changes, what are we losing?
- Are we just going to write specs and orchestrate agents?

My perspective:

- I love coding..
- Less time on things I don't want to do
- More time exploring new things
- Able to do what I couldn't before

Thank you, ADC !!!



Tobias Baumbach

tobias.baumbach@arturia.com | <https://www.linkedin.com/in/tobias-b-472ab75a/>