AI and Research Practice

Lab 1: Using agents and the terminal

Derek Harmon

November 10, 2026

BROAD COLLEGE *of* BUSINESS

Opening session was Oct 13

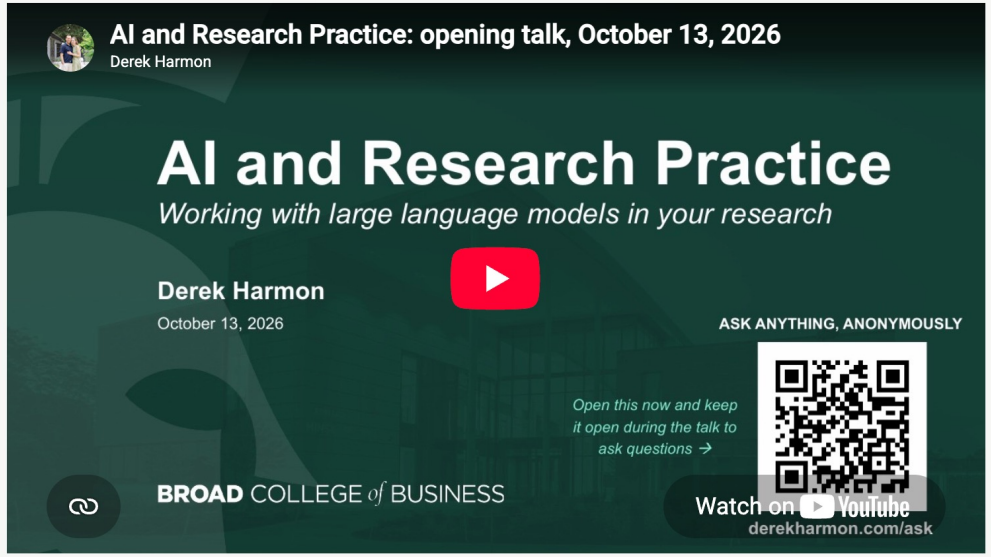
Available online @
https://www.derekharmon.com/ai/lab/oct-13

Lab sessions

AI and Research Practice

Recorded

The opening talk. How large language models actually work, and what that means for getting output you can trust in research.



Slides (PDF)

Reminder of my goals

1. Most conversations are about permissibility or the future of AI. My aim is to discuss how people can use these tools to get high-quality output.
2. Create a space for talking about this, where people can develop their intuitions, learn practical skills, and talk with others that are interested.

A caveat

I use these tools heavily, and part of my research is on the effective uses of these tools.
But I am always still learning, and will not claim to always be right.

Three labs

LAB 1 · TODAY

Using agents and the terminal

Developing agent architecture and working at scale in the terminal.

LAB 2 · NOV 17 · TYPE 1

Context and verification when a right answer exists

Attaching sources and building checks that improve output accuracy.

LAB 3 · DEC 1 · TYPE 2

Context and verification when no right answer exists

Curating context and running robustness checks against what you intended.

Today

1. **Motivation**
   A task we want to accomplish.
2. **Agents**
   What an agent is, the payoff. Why it's hard to use agents in a chat.
3. **The terminal**
   What it is, and why researchers use it.
4. **Agents in the terminal**
   The whole task, run live.
5. **Closing thoughts**
   Chat or terminal, which plan, and how to start.

QUESTIONS ANYTIME
Put them in the Zoom chat.

1 Motivation

Walking through concrete examples is the best way (in my opinion) to learn these tools.

Today’s task: who replaced the CEO?

For every CEO change at an S&P 500 firm in 2025, build an event sample that includes:

FIRM	NEW CEO	DATE	INSIDER?
—	—	—	—
—	—	—	—
—	—	—	—

THE RESEARCH STEP	Building the event sample for a CEO succession study.
WHY IT’S HARD	The filings mix real CEO changes with CFOs, board seats, interim CEOs, and the same change filed more than once.

Two ways most people would go about it

1

BY HAND

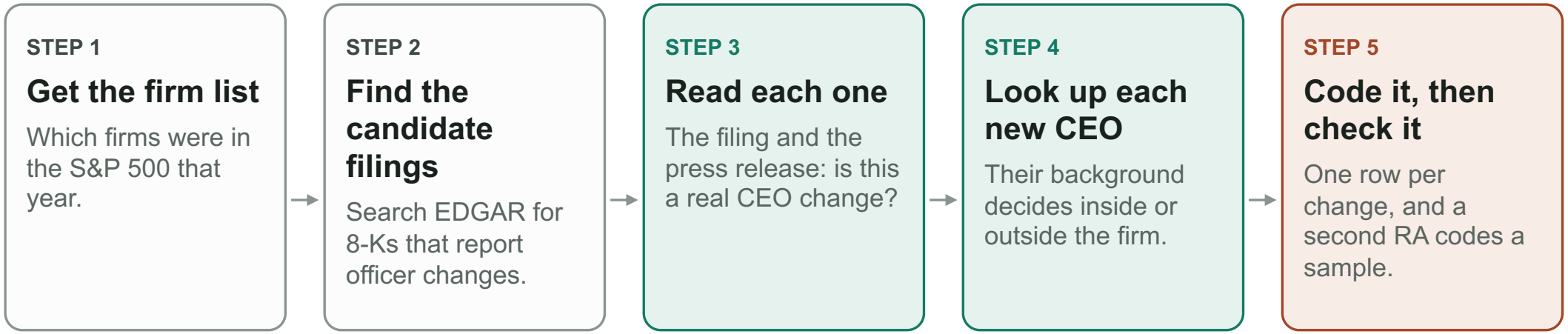
An RA team: the firm list, each filing read one by one, a spreadsheet, and a second coder.

2

IN A CHAT

Ask ChatGPT, Claude, or Gemini to do it for you.

Approach 1: by hand

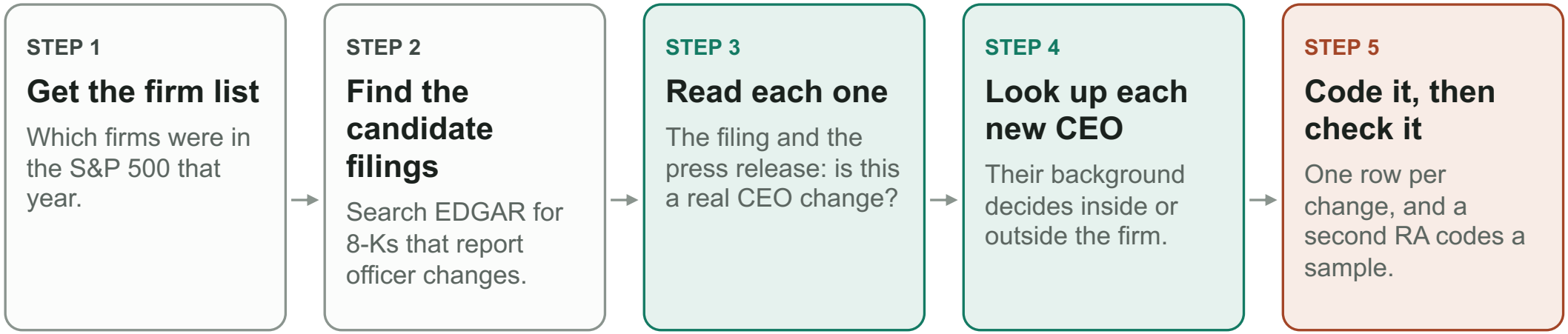


Approach 2: in a chat

ChatGPT	Claude	Gemini	Copilot
CAN RUN	CAN RUN	CAN RUN	CAN RUN
GPT-6 Astra	Opus	Pro	OpenAI models
Sol	Sonnet	Flash	Anthropic models
Luna	Haiku	Flash-Lite	

DEMO 1 ONE CHAT, ONE REQUEST · LIVE

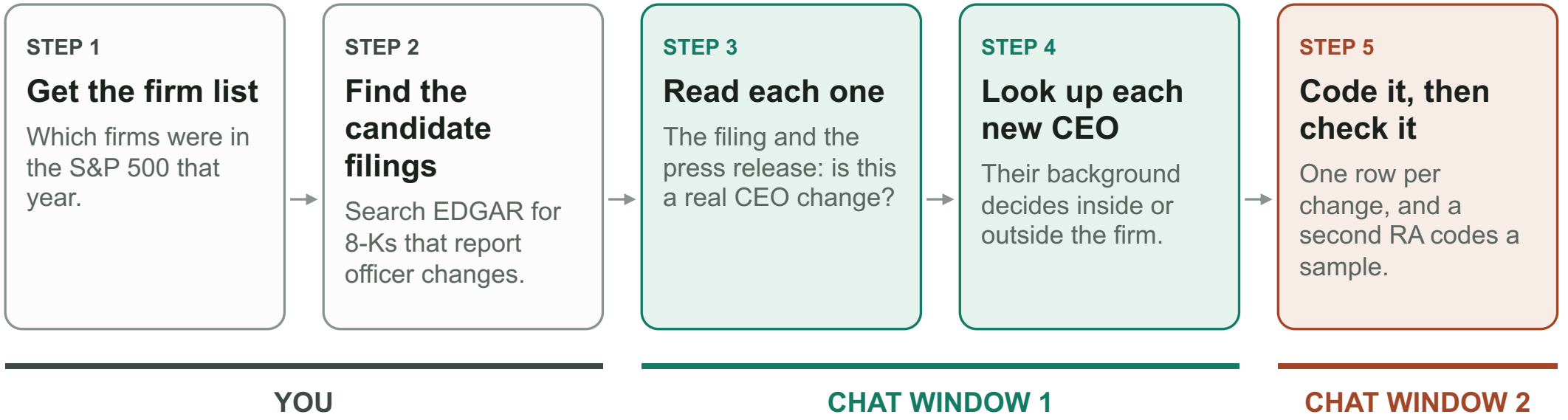
"For every firm in the S&P 500 as of January 1, 2025, find out whether it named a new permanent CEO in 2025 …" + AGENTS.md's rules



Ask the chat to do it all. One chat does every step, including checking its own work.

DEMO 2 YOU AND THE CHAT · TWO FIRMS, CODED TWICE · LIVE

“Here are the rules and two firms’ 2025 8-Ks, with their press releases. Was there a real CEO change? Who, when, and was the new CEO an insider?”



Split the work. You gather the filings. The chat does the reading, the part that takes the most time.

What about all 500 firms?

2

FIRMS IN DEMO 2

8 PDFs I can paste in, and two tables I can check by eye.

500

FIRMS IN 2025

Too many officer-change 8-Ks to paste in, and too many to check.

And this is just 1 year

An agent can do the entire process itself, at scale.

So what's an agent?

2 Agents

What an agent is

A model that works toward a goal by taking actions, seeing the results, and deciding what to do next.

A GOAL

One job, stated up front.

ITS OWN CONTEXT

What it can see: its instructions, plus whatever it reads.

RUNS IN A LOOP

It acts, sees what happened, and goes again until the job is done.



The common shorthand: *“a model that uses tools in a loop to reach a goal.”*

Agent vs. chat, where you are the loop

AGENT

It keeps going until the job is done.

- You set the goal
- ↓
- It acts and sees the result
- ↓
- It decides the next step

CHAT

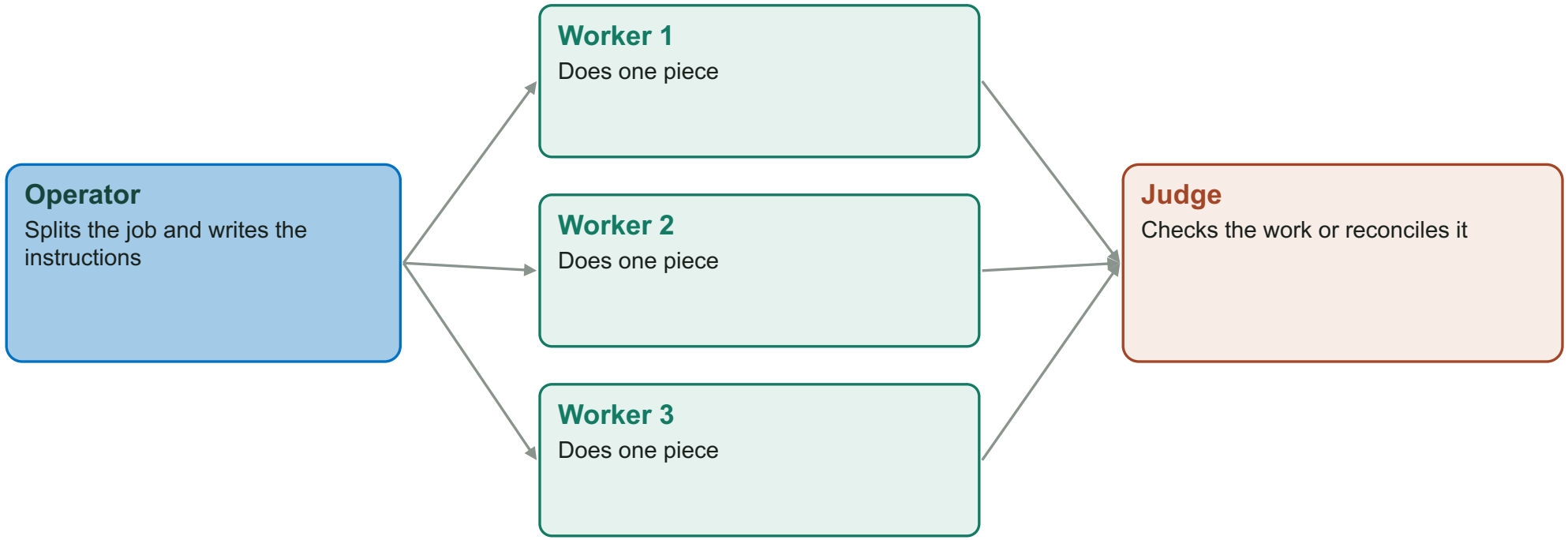
One answer per message, then it waits for you.

- You ask
- ↓
- It answers
- ↓
- You read, decide, and ask again

An agent does the steps itself instead of waiting on you after every reply.

Several agents, wired together

AGENTIC ARCHITECTURE



Each agent has its own goal and its own context. Deciding what each one sees is the design work.

Why are agents a superpower for research?

Focus

Each agent sees only its piece and the material for it. Limits distractions and errors that arise when you provide context that isn’t relevant to the task.

Independence

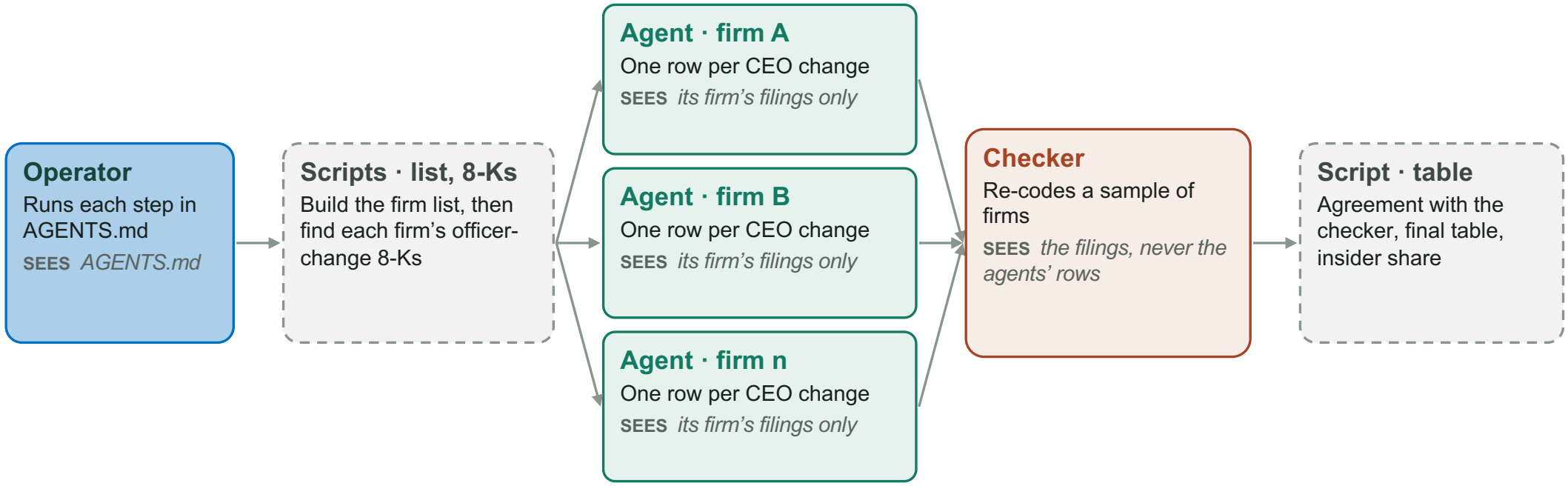
Agents that can’t see each other don’t anchor on each other. A disagreement thus tells you something useful, and a separate agent can reliably do a check.

Scale

The work can be bigger than one ChatGPT conversation can hold.

Today’s task, as an architecture

SCRIPTS IN GREY · AGENTS IN COLOR



Focus

Each firm agent sees only its firm’s filings.

Independence

The checker never sees the agents’ rows.

Scale

The firms run at the same time.

A script or an agent?

SCRIPT

Follows steps written in advance.

Same result every run.

Can’t handle what its author didn’t plan for.

IN TODAY’S TASK

Find every 8-K that reports an officer change.

AGENT

Gets a goal and works out the steps.

Decides what to do, sees the result, goes again.

Its steps depend on what it finds.

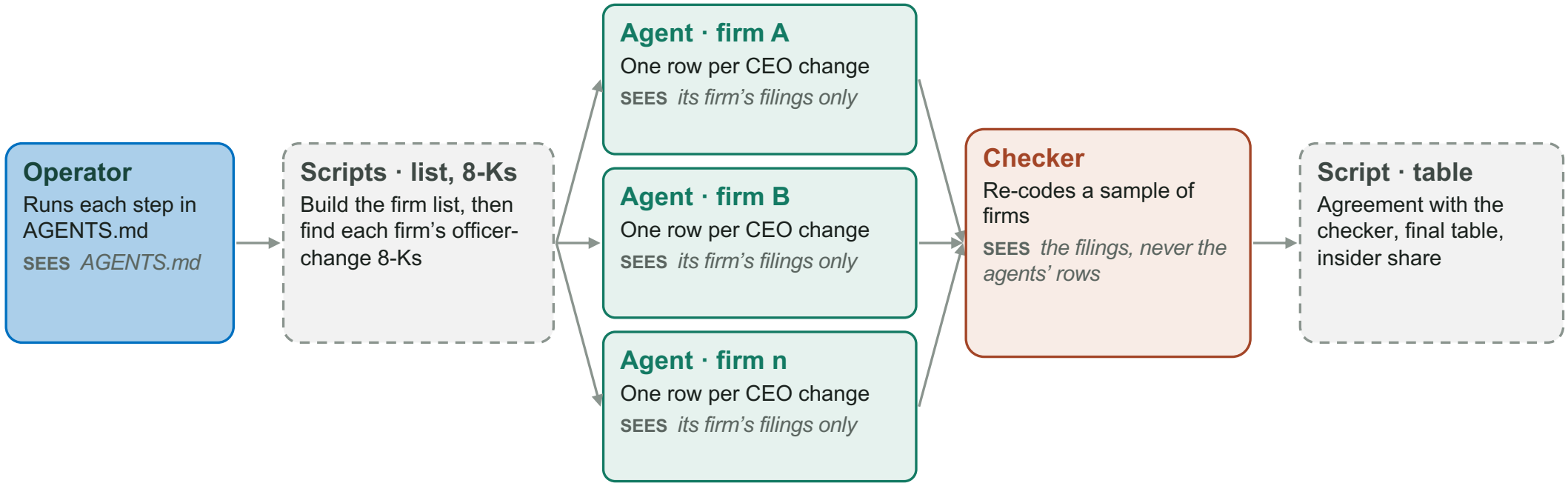
IN TODAY’S TASK

Is this a real CEO change? If the filing doesn’t say where the new CEO came from, go find out.

A script fits when you can write the steps ahead of time. An agent fits when the steps depend on what it finds.

Today's task, as an architecture

SCRIPTS IN GREY · AGENTS IN COLOR



An architecture like this needs something to run it (i.e., start each agent with only its own files, run them side by side, and hand their answers to the checker).

In Demo 2, I did that. But Codex or Claude Code can do it using the terminal.

But first, **what is the terminal**?

3 The terminal

What the terminal is

THE TERMINAL ON ITS OWN

It takes commands, not English.

ILLUSTRATION

```
~ $ cd ceo-succession
~/ceo-succession $ ls
AGENTS.md   firms.csv   output
~/ceo-succession $ open firms.csv
```

`cd` moves into a folder, `ls` lists what's in it, and `open` opens a file, as if you'd double-clicked it.

WITH AN AGENT INSIDE

Type `codex` or `claude`, and it takes plain English.

ILLUSTRATION

```
~/ceo-succession $ codex
› Follow AGENTS.md for every firm.
• Reading AGENTS.md
• Writing a script to get the 8-Ks
```

The agent reads and writes the files in that folder:

- Collect data from online and deposit it here
- Clean data and run analyses here
- Summarize or code files that are saved here
- Work on literature reviews or writing drafts here

A nicer terminal: Warp

What I'll use today.

A modern terminal where you can click to move the cursor, select text like a document, and read output in blocks. Free, and an account is optional.

Warp has its own AI features too. Today it's only the window. Codex does the work.

OTHER GOOD ONES

iTerm2 Mac

Ghostty Mac

Windows Terminal Windows

Why the terminal? Four answers

Agents are just one answer of four.

1

It runs agents for real

Separate contexts, run in parallel, and the checker waits for every worker. No copy and paste.

2

Your project is the context

Point it at a folder of papers, data, or drafts. It reads what it needs.

3

It acts, not just answers

It runs code on your data, writes and edits files, and checks its own results.

4

It’s repeatable

The instructions live in a file you can open and edit. Next month, rerun the same folder.

What to get used to in the terminal

A .MD FILE

```
## Rules for coding a firm
**Read only SEC filings.**
- A co-CEO counts.
- Don’t count an interim CEO.
```

Plain text. The agent reads and writes it directly.

THE SAME LINES IN A WORD FILE

```
<w:p><w:pPr><w:pStyle w:val="Heading2"/></w:pPr>
<w:r><w:t>Rules for coding a firm</w:t></w:r></w:p>
<w:p><w:pPr><w:pStyle w:val="FirstParagraph"/></w:pPr>
<w:r><w:rPr><w:b/><w:bCs/></w:rPr>
<w:t>Read only SEC filings.</w:t></w:r><w:r>
…
```

A zipped bundle of files. Hard for the agent to edit.

TWO .MD FILES THAT RUN THE WORK

THE INSTRUCTIONS

`AGENTS.md` (Codex) **`CLAUDE.md`** (Claude)

Your rules, written once. The agent reads them every time it starts, so you never paste them in again.

THE MEMORY

`PROGRESS.md`

Each session starts blank. This file says where things stand, so the next session picks up where you left off.

4 Agents in the terminal

Using agents to complete the whole task, end to end.

DEMO 3 THE SAME TASK IN CODEX · LIVE, ON 10 FIRMS

`› Build the list.  › Get the 8-Ks.  › Draw it.  › Now follow AGENTS.md.`

FOUR REQUESTS, LIVE

1 · The list Codex builds the S&P 500 from the SEC and WRDS.

2 · The 8-Ks it finds each firm's officer-change 8-Ks.

3 · The architecture it draws agent workflow as a picture.

4 · The agents one per firm, a checker, and a table script.

```
ceo-succession/
  AGENTS.md
  firms.csv
    500 firms (step 1)
  filings/
    10 firms' 8-Ks (step 2)
  architecture.html (step 3)
  output/
    ceo_changes.csv (step 4)
```

What Demo 3 showed

Each firm got its own agent, in its own context. The checker never saw their rows.

1

It ran agents for real

One agent per firm, separate contexts, no copy and paste.

2

The folder was the context

The filings landed in filings/, and the checker read them there.

3

It acted

It built the list, pulled the filings, and wrote the table to output/.

4

It’s repeatable

Its scripts and AGENTS.md stay in the folder. Rerun them next year.

The same design, run at full scale

All firms, 2010 – 2023

729

CEO CHANGES COLLECTED

2,091 officer-change 8-Ks and 1,397 press releases.

~60

MINUTES, START TO FINISH

Pull S&P 500 lists, download filings, run agents, and build the table.

~30%

OF A WEEK'S ALLOWANCE ON CHATGPT PLUS

All demos before this full-scale run took 2–3% of week's allowance.

Analysis. Outcome: the stock's CAR over the 3 days around the announcement.
Controls: forced exit, firm size, ROA, leverage, industry and year fixed effects.

5 Closing thoughts

Chat products have projects too

	CHAT / PROJECT FOLDERS	TERMINAL FOLDER
Your files	Uploaded copies	**The folder itself**
What it does with them	Reads them, runs some code	**Reads, writes, edits, and runs code on them**
Data analysis	In the vendor’s sandbox, which resets	**On your machine, so the data can stay put**
Instructions	Project instructions	**AGENTS.md, a file you can open**
What you can see	The answer	**Each step it took**

The difference is how much of the context you can see and control.

Choosing a tier and a tool

1) **Start with your data.** It decides which tools you may use.
2) **Then how much work you'll run.**

TOOL	COST TO YOU	TERMINAL AGENT	YOUR DATA
Copilot or Gemini, MSU login	Free	No	**MSU-licensed**
ChatGPT Edu, from your department	Department pays	Codex	**MSU-licensed**
ChatGPT Plus, personal	$20 a month	Codex	**Public data only**
Claude Pro, personal	$20 a month	Claude Code	**Public data only**
ChatGPT Pro, Claude Max	About $100–$200	Same, more usage	**Public data only**

What to take with you

1. **An agent is a model that works toward a goal on its own.**

   It acts, sees what happened, and decides the next step.

2. **Move to agents in the terminal when your work outgrows the chat.**

   It requires adjustment. But stick with it…it's worth the investment!

Conclusion

Three labs

LAB 1 · TODAY

Using agents and the terminal

Developing agent architecture and working at scale in the terminal.

LAB 2 · NOV 17 · TYPE 1

Context and verification when a right answer exists

Attaching sources and building checks that improve output accuracy.

LAB 3 · DEC 1 · TYPE 2

Context and verification when no right answer exists

Curating context and running robustness checks against what you intended.

Thank you!



TODAY'S RECORDING AND SLIDES

derekharmon.com/ai/lab/1

Posted after the session.

GUIDE, DEMOS, SLIDES, AI POLICIES, AND EVERYTHING ELSE

derekharmon.com/ai

The setup guide and the tools table are here too.