NAPAI AGENTIC AI · FINALS
2026.11.21

Chronicler

Research meetings, kept as a reviewed record you can ask.

TEAM
07 · 牛排

TRACK
Campus self-proposed

THEME
Library, research & knowledge services

MOTIVATION

Week 1: a decision is spoken.



Researcher A 00:13

U-Net 跑完了，但同一個 session 的影像同時出現在訓練集跟驗證集。

U-Net finished, but images from one acquisition session are in both training and validation.

Lab Lead 00:31 · decision

那驗證集改成以 session 切分，整個 session 一起保留。

Then split validation by session. Keep whole sessions together.

Researcher B 00:48 · task

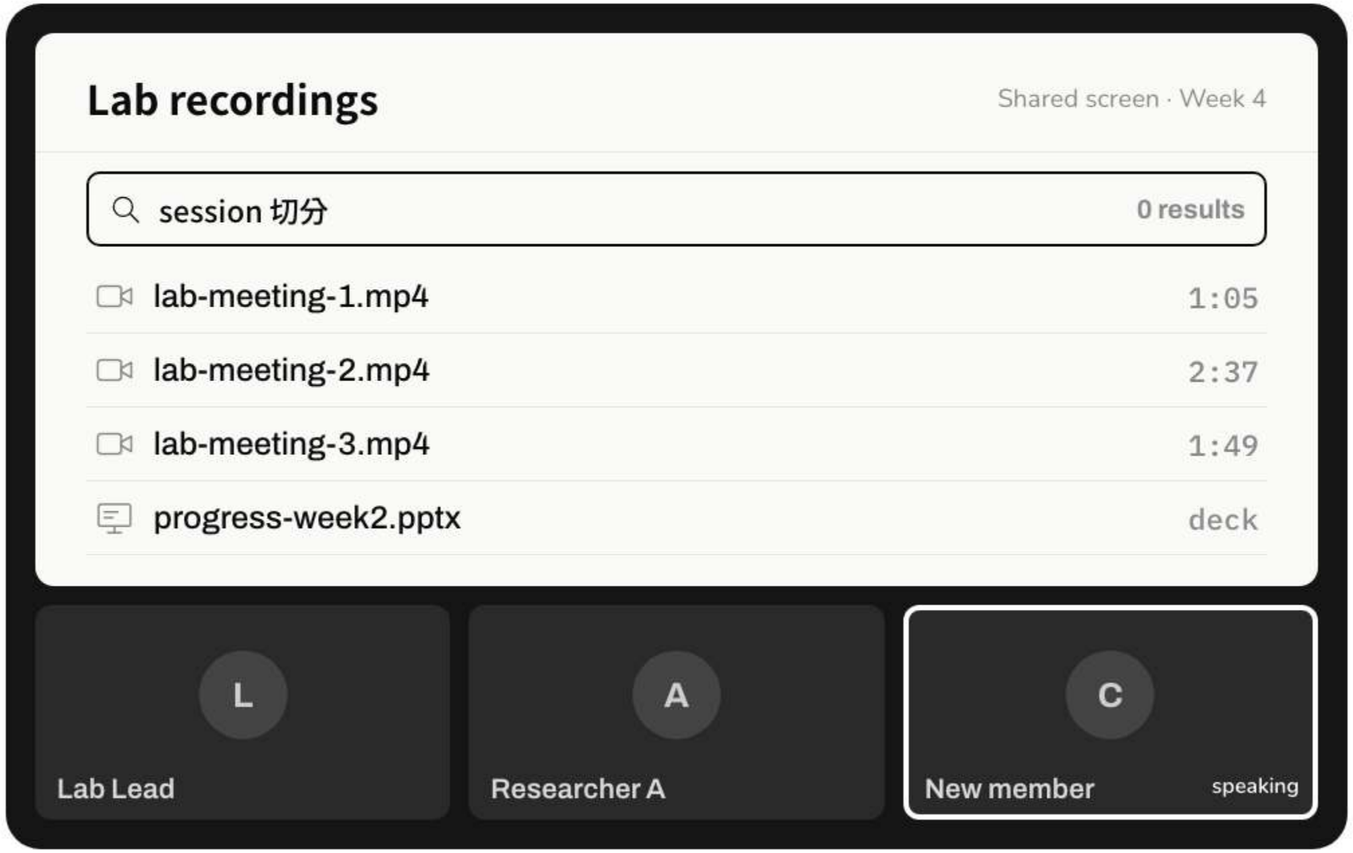
模糊的標註我這週再檢查一次。

I'll recheck the ambiguous annotations this week.

Illustrative scene from our demo lab. Speech is Mandarin; English gloss in grey.

MOTIVATION

Week 4: nobody can find it.



New member 00:05
驗證集為什麼要用 session 切？
Why do we split validation by session?

Researcher A 00:09
好像是第一週……還是第二週講的？
I think it was week one… or week two?

Lab Lead 00:15
誰有那次的筆記？
Does anyone have notes from that meeting?

Nobody does.

Recording lengths match our field site: 65, 157 and 109 minutes, 331 in total.

PROBLEM

Decisions are spoken, not recorded.



Scattered

Decisions live in hours of Mandarin speech and slides, spread across weeks.

Week 1’s decision sits inside 331 minutes of video.



Unverified

When notes exist, they are informal, and nobody checks what is kept.

Nobody had notes from that meeting.



Untraceable

A decision can’t be traced back to the moment, or the slide, it came from.

“Week one… or week two?”

INTERVIEWS · PROFESSORS

“They write what to do, not why.”

Two professors who lead labs · P01: 12-question interview · P02: 6 written answers

> **What gets lost most is what the last meeting discussed, what it decided, and what each student should finish.**
>
> **P01** interview, Q2

> **A student listed the same two tasks three weeks running. It was settled long ago, but with no record we had to redo the plan.**
>
> **P01** interview, Q3

> **Students write down what to do, but not why they are doing it.**
>
> **P02** written answer, Q2

> **We gave detailed advice on a report. The student’s notes said only: “Optimize project report.”**
>
> **P02** written answer, Q3

P01 answered in Mandarin (edited translation); P02 answered in English. Shortened for the slide.

Answers became requirements.

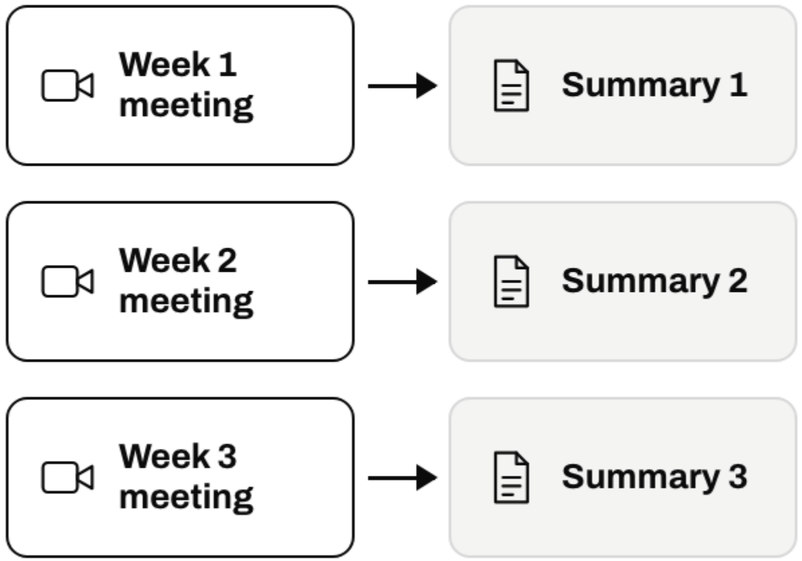
P01	“AI can draft an internal record, but a PM or senior student should check it before it is kept.”	→	**A person approves** R2
P02	“A decision made in the meeting is not always the final version.”	→	**Decisions keep history** R3
P01	“Experiments that didn’t work should be kept, not deleted, so the next student doesn’t repeat them.”	→	**A lasting record** R1 · R3
P01	“Notes need a recap of last week, progress on each item, and the TODOs that carry over.”	→	**Organized across weeks** R1 · R4

Can today's AI tools solve this?

Two widely used examples: Zoom AI Companion and NotebookLM.

Zoom AI Companion

An AI assistant built into the meeting platform



What it does well

- ✓ Summary and action items after each meeting
- ✓ Answers questions about your meetings
- ✓ Host reviews a summary before it is shared

Gap for a research lab

- ✗ **One summary per meeting, not a project record**
- ✗ **No history of how a decision changed**
- ✗ **Recordings processed in the vendor's cloud**

NotebookLM

A notebook that answers only from sources you upload

PDF · Slides · Audio · YouTube · Notes

Picked by hand, for each notebook

Notebook
All sources read as text, side by side

Answer
“The split was changed to…” 1 opens the passage

What it does well

- ✓ Answers grounded only in your sources
- ✓ Inline citations open the exact passage
- ✓ Reads documents, audio and YouTube

Gap for a research lab

- ✗ **Sources are hand-picked for each notebook**
- ✗ **No review step and no version history**
- ✗ **Runs in the vendor’s cloud**

THE GAP

What a research lab needs.

Requirement, from our interviews		Zoom AI Companion	NotebookLM
R1	**Knowledge organized by project, across weeks**	◐	◐
R2	**Only human-approved content becomes memory**	◐	○
R3	**Decisions keep their history**	○	○
R4	**Answers trace to the exact second or slide**	◐	◐
R5	**Unpublished data stays on campus**	○	○

● **met** ◐ **partly met** ○ **not found in public docs** Vendor documentation and press coverage, Nov 2026.

APPROACH

Why not one big agent?

Read

INPUT	Hours of audio, slides, frames
WORK	Long multimodal perception
WRITES	Drafts only
RISK	Speech and vision errors

Person approves

Organize

INPUT	Approved revisions
WORK	One consistent structure
WRITES	Memory, via checked commit
RISK	Corrupting the record

Person asks

Answer

INPUT	A member's question
WORK	Search, read, decide, cite
WRITES	Nothing
RISK	Hallucination, injected text

One agent with every tool would mix noisy perception, write access and untrusted questions in one context. Splitting gives each job only the access it needs, and puts people where judgment is required.

Notetakers summarize meetings.
Chronicler keeps a lab's record.

A research-memory system that turns a lab's meetings and slides into a reviewed, versioned record that members can question, with citations.

Recordings in. A record out.



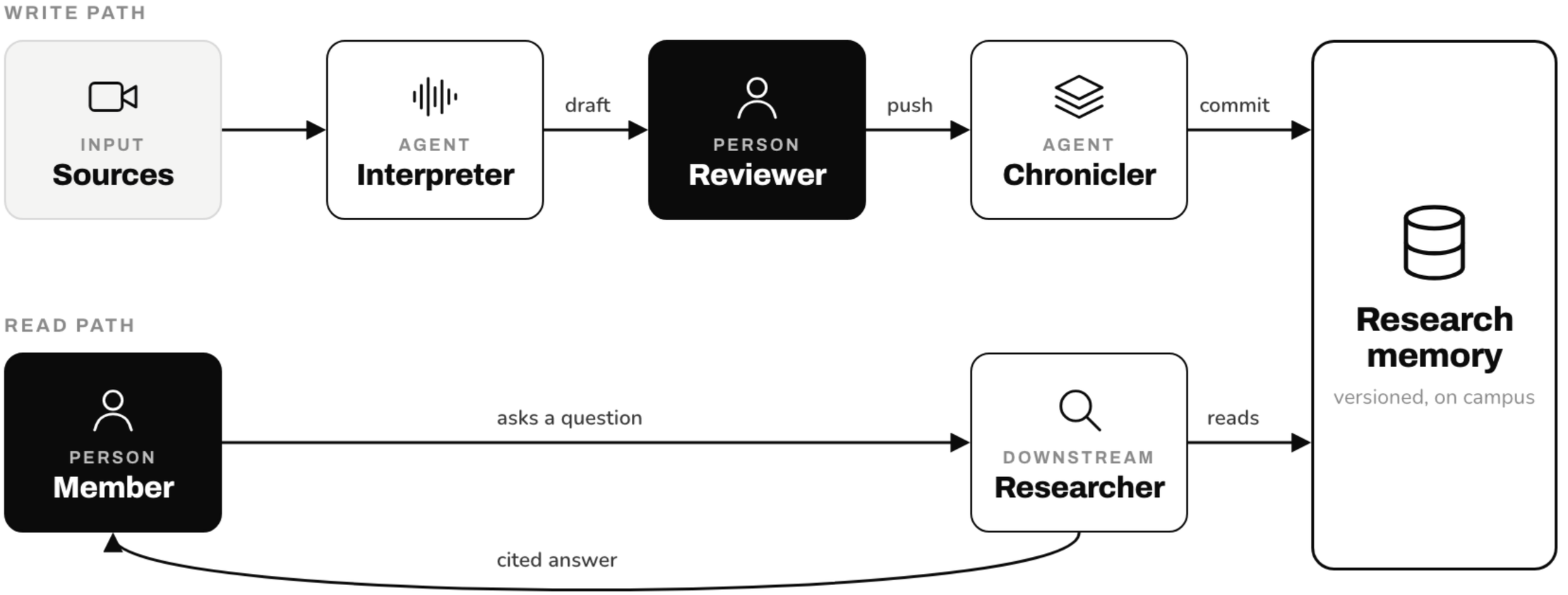
DESIGN PRINCIPLES

Each requirement became a rule.

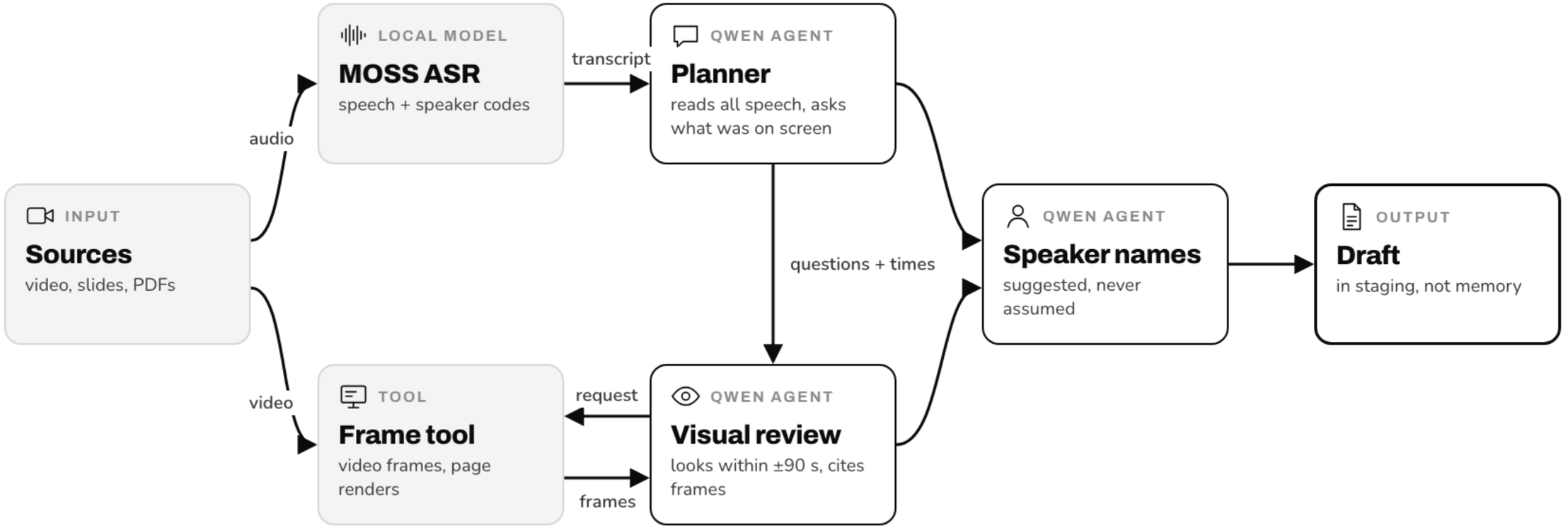
R1	Organized by project, across weeks	→	**Project → topic → segment memory**
R2	Only approved content becomes memory	→	**A person pushes every version**
R3	Decisions keep their history	→	**Immutable versions, never overwritten**
R4	Answers trace to the exact moment	→	**Citations to the second, slide or page**
R5	Unpublished data stays on campus	→	**Local models on one machine**

UNDER THE HOOD

Three agents, two people.



From a recording to a cited draft.



Agents only read and write drafts. Grey boxes are deterministic tools; the slide decks follow the same path through LibreOffice, Poppler and OCR.

Same clip, two transcripts.

Zoom transcription

Simplified Chinese

01:00:30–01:00:53

这样才能match到pb的，就是在sd去match到pb的，然后第二个就是我们让那个开始动起来，那我们这边假设是和之间它没有任何交互作用，也不会碰撞也不会有反应对，所以

01:00:54–01:01:17

那这边的sd很简单，可以写出来就是等于跟号二DW就是那个。然后这边跟他三弟就是…

~~Reference transcript~~ MOSS · ours

Traditional Chinese

01:00:30–01:00:53 · Speaker A

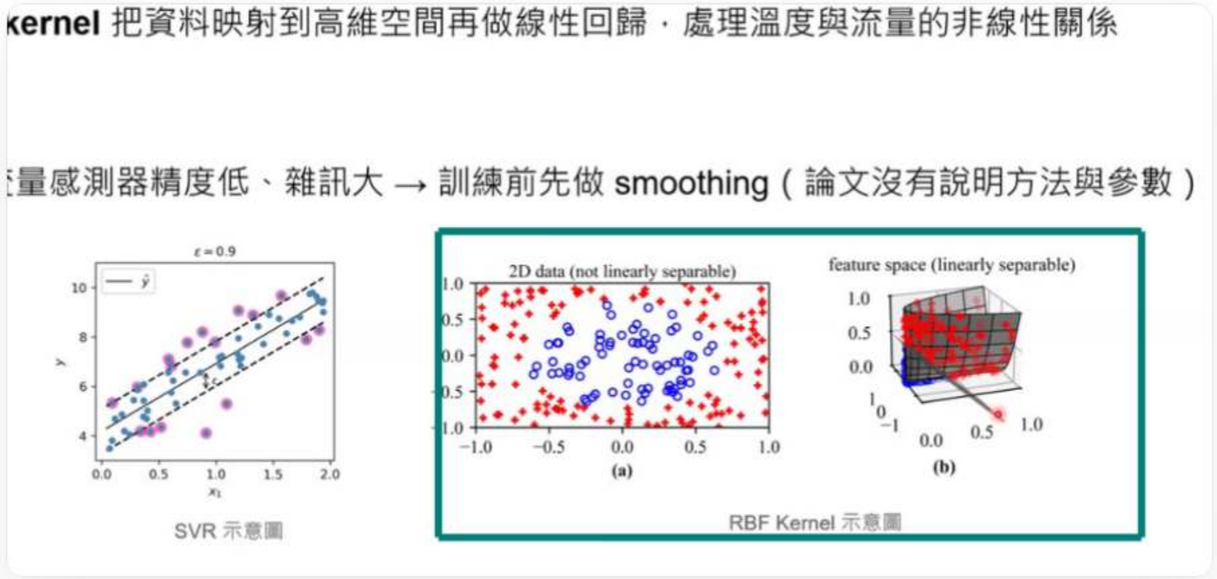
這樣才能 match 到 PDE 的 initial condition，就是在 SDE 去 match 到 PDE 的 initial condition。然後這個就是我們讓那個 particle 開始動起來，那我們這邊假設是 pure diffusion 啊，就是 particle 和 particle 之間它沒有任何交互作用…

01:00:53–01:01:17 · Speaker A

對，所以那這邊的 SDE 很簡單，可以寫出來，就是 dx 等於根號二 D，D，W，W 就是那個呃，Brownian motion 嘛…

Real lab-meeting clip, no edits. Zoom wrote “跟他二弟” (“his second brother”) for 根號二 D (“square root of 2D”).

Speech alone is ambiguous.



Frame retrieved near the speech

SPEECH

「轉到三維…再轉到 2D…非線性的關係」

"...to three dimensions... then back to 2D... a nonlinear relationship"

VISUAL CONTEXT

RBF kernel maps data into a higher-dimensional space.

The draft carries both.

Final Markdown transcript

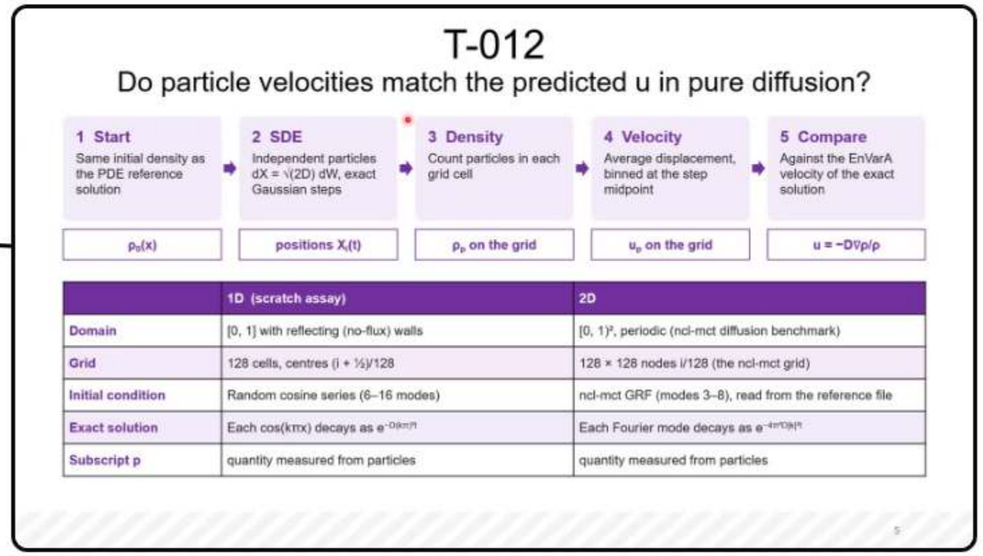
01:00:00–01:00:29 · Speaker A
對，然後因為 SDE 它只有 particle，所以我們要去算這個每個 particle 它對應到的
呃，mass 是多少，就質量是多少，然後我們再把它根據那個 initial condition 去，
呃，去分布，把去 match 到那個 PDE 的 initial condition，就是
s 是多少，那我們就要塞幾個 particle 進去，這樣

Understood visually by the agent

視覺脈絡 [01:00:26]：畫面顯示標題為「T-012 Do particle velocities match the predicted u in pure diffusion?」的簡報頁面。頁面包含一個五步驟的流程圖（1 Start, 2 SDE, 3 Density, 4 Velocity, 5 Compare）以及一個比較 1D (scratch assay) 與 2D 設定參數（Domain, Grid, Initial condition, Exact solution, Subscript p）的表格。在流程圖中，步驟 2 SDE 的方塊上方有一個紅色圓點標記，該標記在後續幀中移動至步驟 1 Start 及標題下方。
與發言的關係：發言者 Speaker A 正在解釋如何將 SDE 的 particle 對應到 PDE 的 initial condition（對應畫面步驟 1 Start 與 2 SDE），並提到假設是 pure diffusion 且 particle 之間無交互作用（對應畫面標題與步驟 2 SDE 中的 Independent particles）。

01:00:30–01:00:53 · Speaker A
這樣才能 match 到 PDE 的 initial condition，就是在 SDE 去 match 到 PDE 的 initial condition。
然後這個就是我們讓那個 particle 開始動起來，
那我們這邊假設是 pure diffusion 啊，就是 particle 和 particle 之間它沒有任何交互作用，也不會碰撞，也不會有反應。

01:00:53–01:01:17 · Speaker A
對，所以那這邊的 SDE 很簡單，可以寫出來，就是 dx 等於根號二 D，D，W，W 就是那個
呃，Brownian motion 嘛，然後這邊根號二 D 就是，
就是它根據 diffusion 的大小，它會動，它的動的 scale 是多少。
對，然後 dx 就是 particle 每步會動多少，這樣。



Frame at 01:00:26, cited by the note

Real output. 視覺脈絡 = visual context; 與發言的關係 = how it relates to the speech.

Better context, better answers.

QUESTION

“According to the speaker, what does W represent in the motion equation?”

ANSWER FROM ZOOM’S TRANSCRIPT

“The context does not explicitly state what W represents in the motion equation.”

Zoom heard: 「跟号二DW就是那个」

ANSWER FROM OUR TRANSCRIPT

“According to the speaker, W represents Brownian motion.”

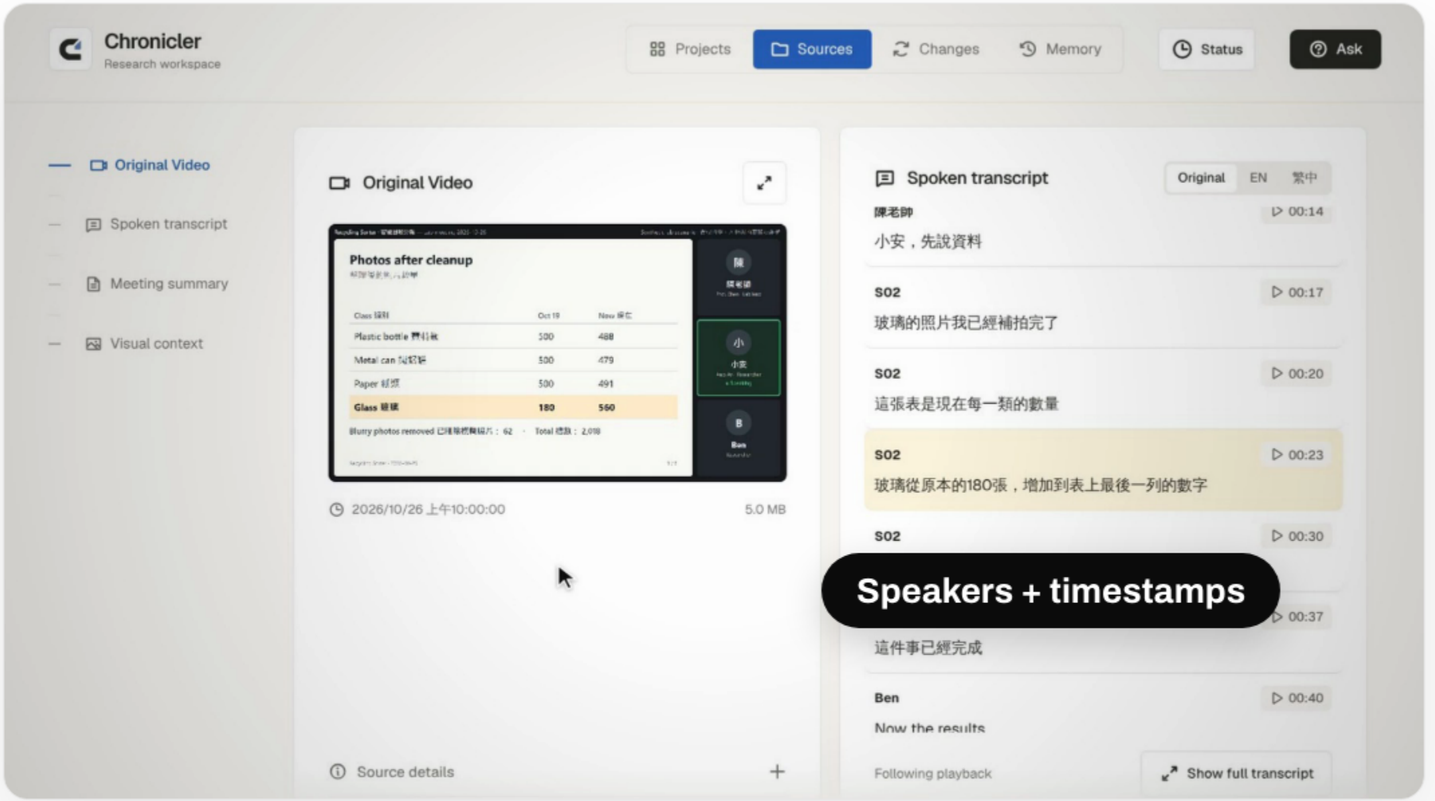
We kept: 「W 就是那個呃，Brownian motion 嘛」

Same question, answered from each transcript of the same clip.

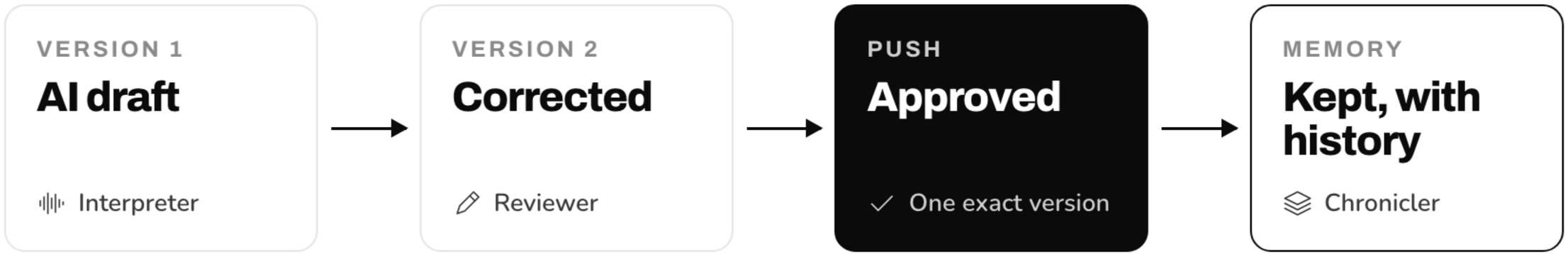
01

It reads the meeting.

Speech, slides and video frames become a draft.

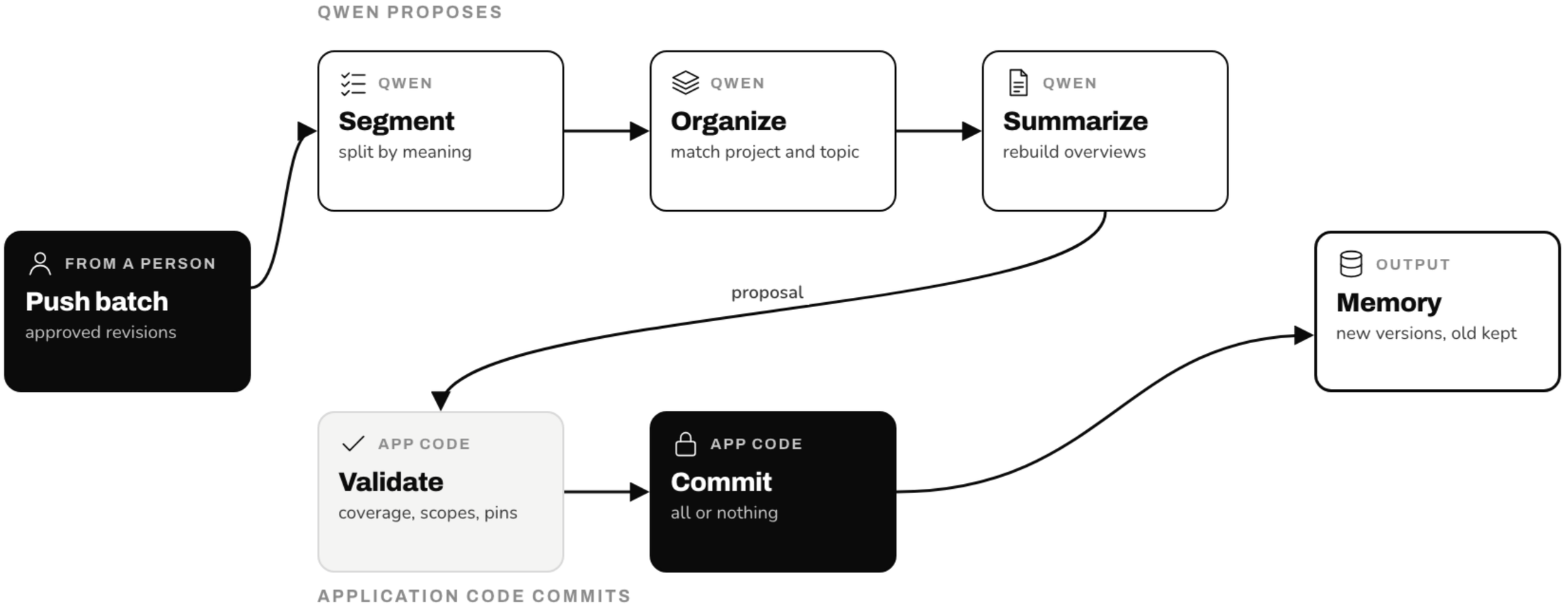


A person approves what is kept.



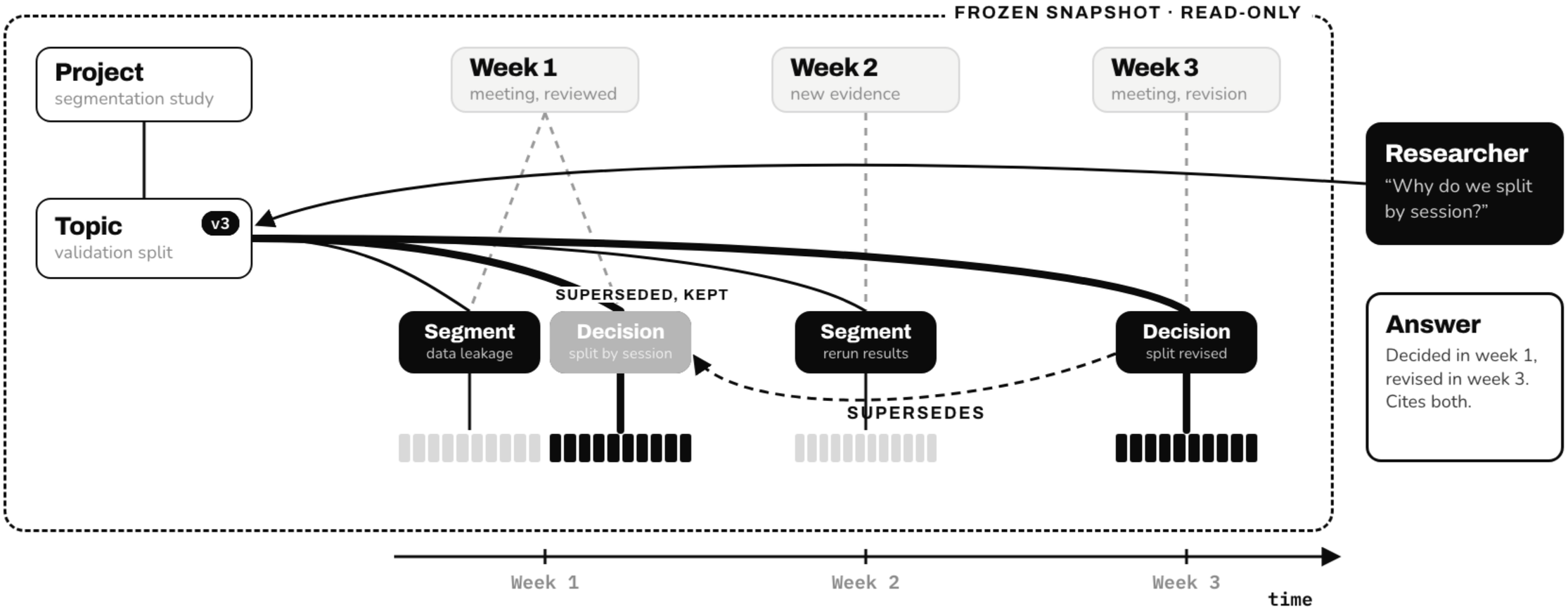
An AI draft can never enter memory on its own.

The model proposes. Code commits.



Model outputs are checkpointed by content hash, so a retry or double-click never commits twice.

One memory graph, over time.

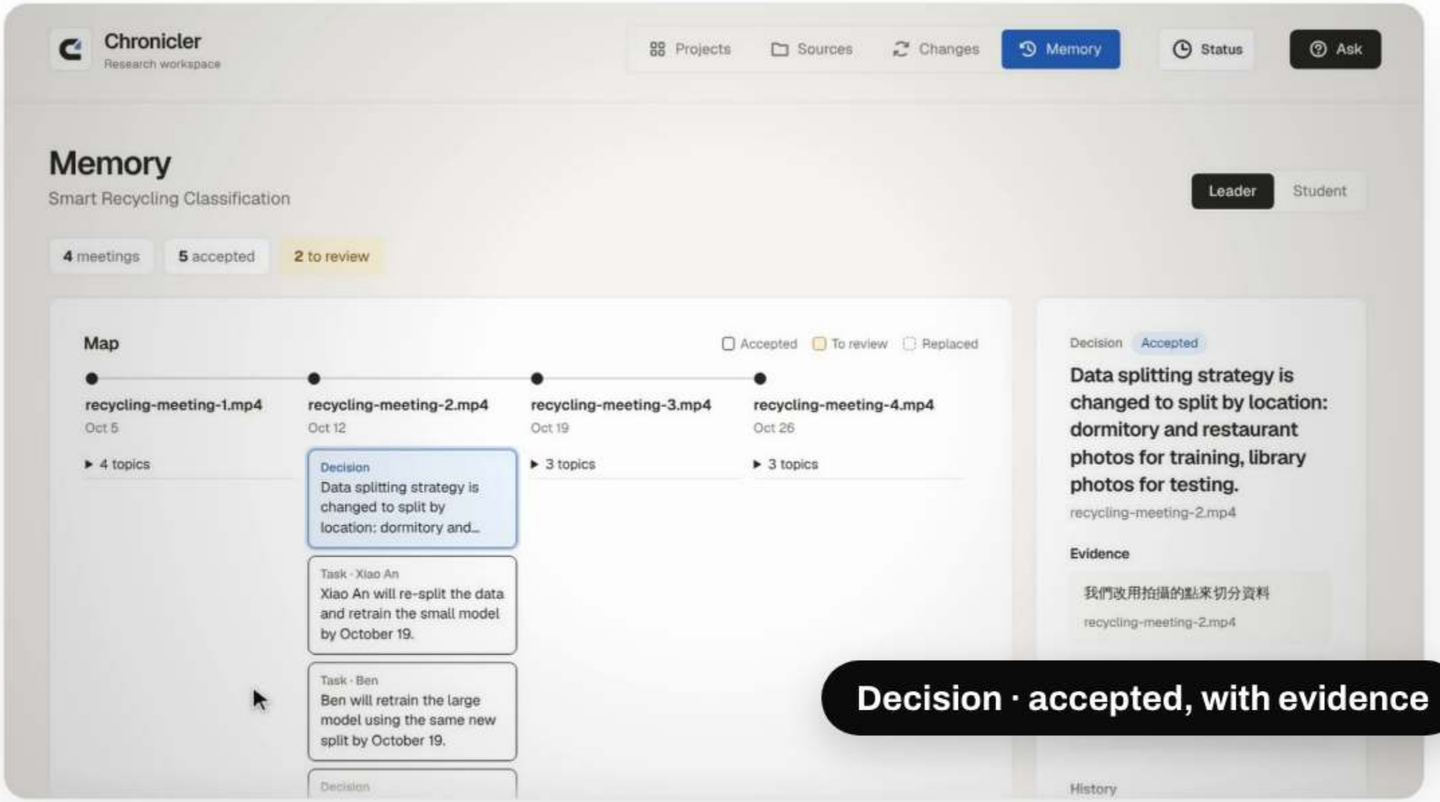


A downstream agent asks. It reads a frozen snapshot: topic → segments → exact units, in time order.

03

It files by project and topic.

Across weeks, with every version kept.



One memory, many uses.

Research memory

reviewed, versioned, read-only to every downstream agent

Researcher — BUILT

Answers questions, citing the exact second or slide

What changed? — BUILT

Proposes decisions and commitments; a person approves

Summaries — BUILT

Per meeting, topic and project

Meeting notes — NEXT

Recap, progress and TODO, as the professors asked

Related papers — NEXT

Find work linked to a project's methods

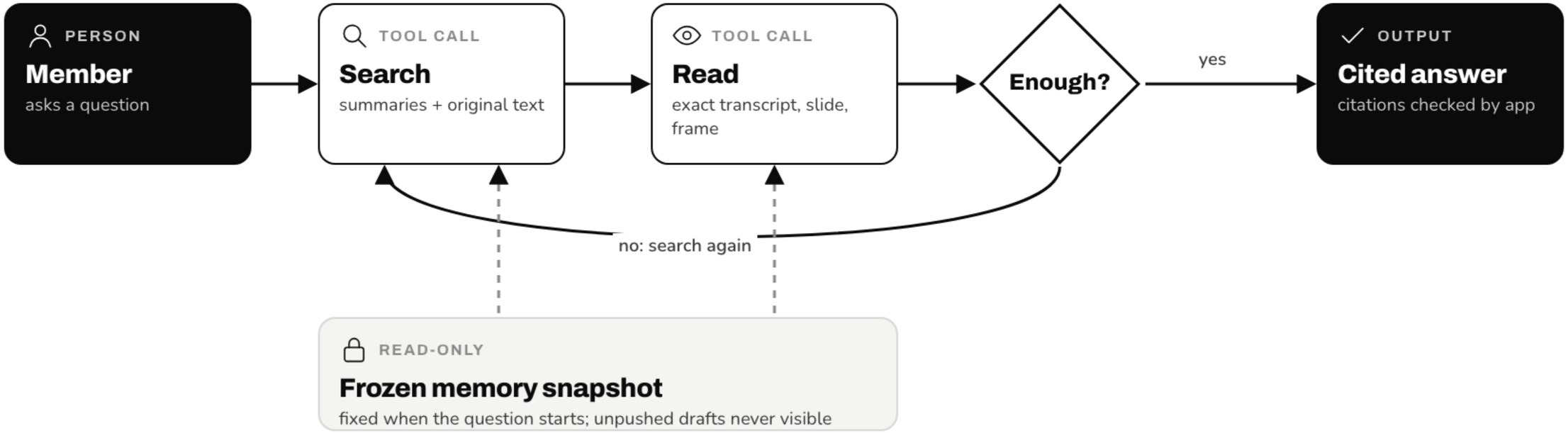
Lab chatbot — NEXT

Help new members catch up on a project

DOWNSTREAM AGENT · RESEARCHER

It decides what to read next.

tool calls · 24 max

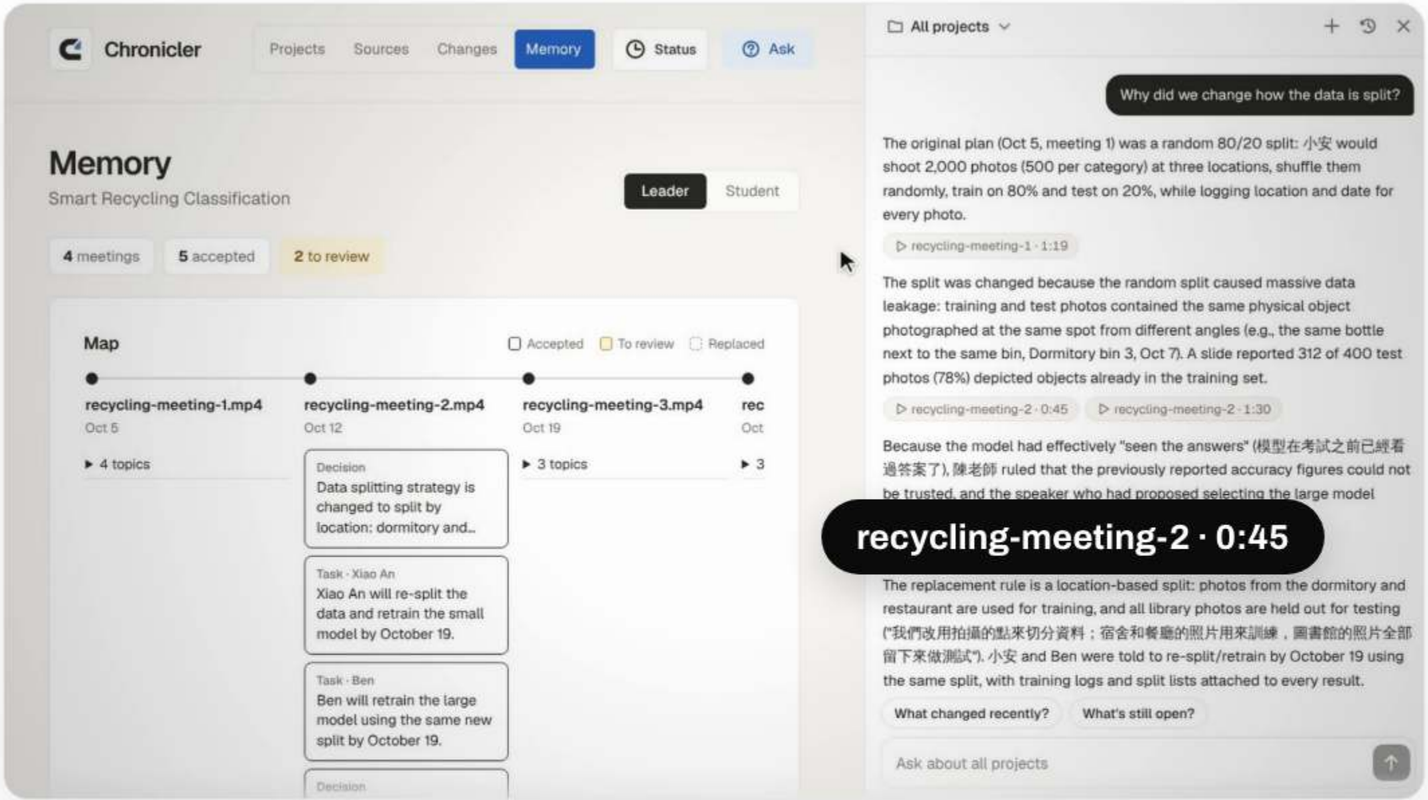


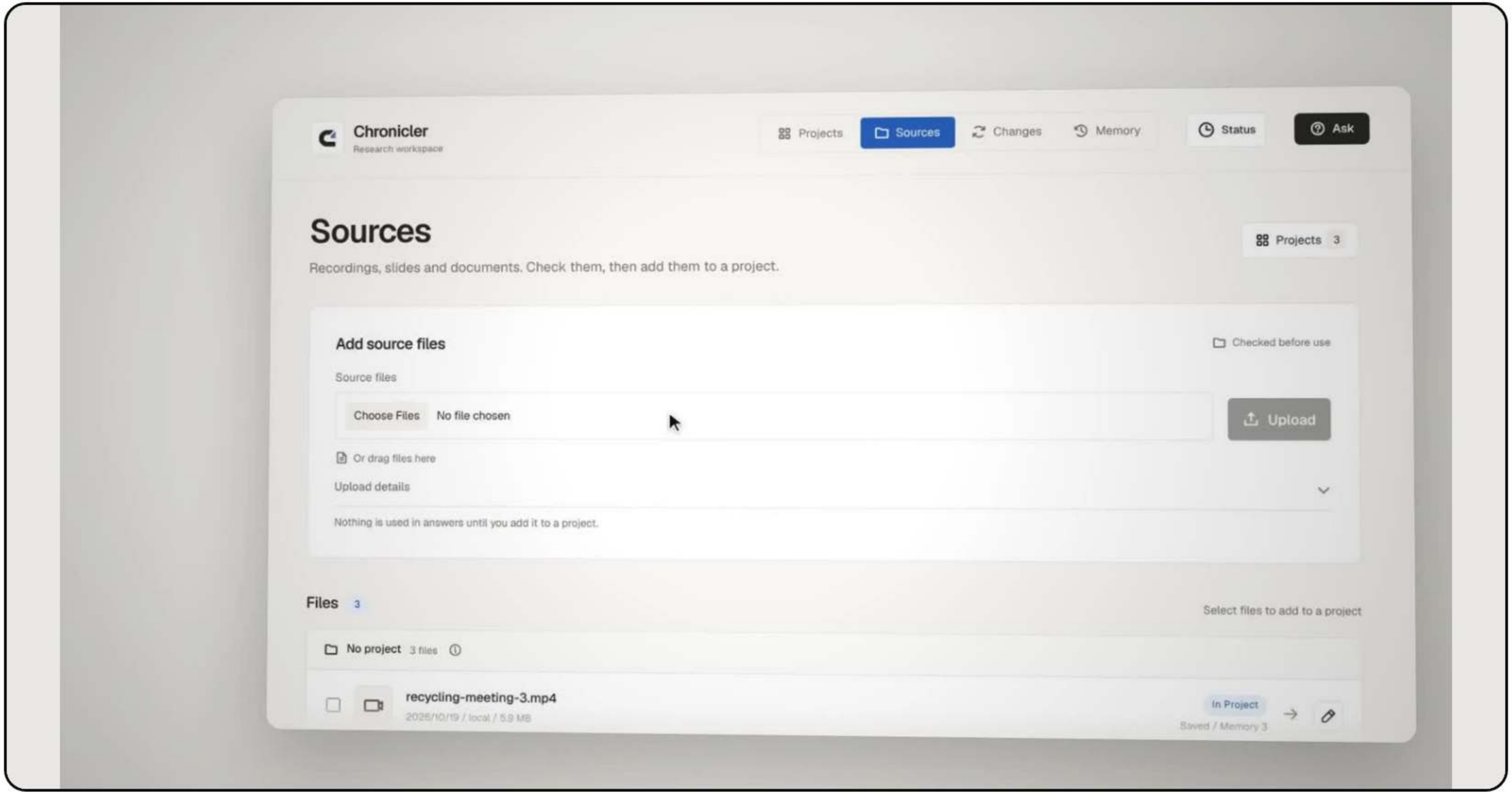
Per question: at most 24 tool calls, 10 rounds and 12 minutes. Text inside sources is treated as evidence, never as instructions.

04

It answers with the source.

"Why did we change how the data is split?" Each claim links to the exact moment in the recording.




Chronicler
Research workspace
Projects
Sources
Changes
Memory
Status
Ask
Sources
Recordings, slides and documents. Check them, then add them to a project.
Projects 3
Add source files
Checked before use
Source files
Choose Files
No file chosen
Upload
Or drag files here
Upload details
Nothing is used in answers until you add it to a project.
Files 3
Select files to add to a project
No project 3 files
recycling-meeting-3.mp4
2026/10/19 / local / 5.9 MB
In Project
Saved / Memory 3

L3–4

Does it work?

10 students tried it

3-week lab pilot

35 benchmark cases

Ten students tried it.

MEAN RATING, 1–5, N = 10



Tasks: find a past decision · check a citation · correct a draft · push to memory

U03 · new master's positive

I didn't have to watch weeks of meeting videos to find the past method and its source.

U09 · senior graduate positive

Keeping old revisions, with a manual push into memory, lowers the risk of citing AI errors.

U02 · project PM problem

At first I couldn't tell "correction saved" from "pushed to memory".

U04 · PhD problem

I still had to work out which decision was the latest one.

Feedback translated and shortened from Mandarin.

FIELD PILOT · TIME TO ANSWER

58.5% faster.

Median query time, 12 paired queries: by hand vs with Chronicler



One research group’s weekly meetings, Oct 19 – Nov 9 · 6 members (4 graduate students, 1 PM, 1 professor) · 3 meetings.

FIELD PILOT · QUALITY

More found. More verified.



FIELD PILOT · EFFORT

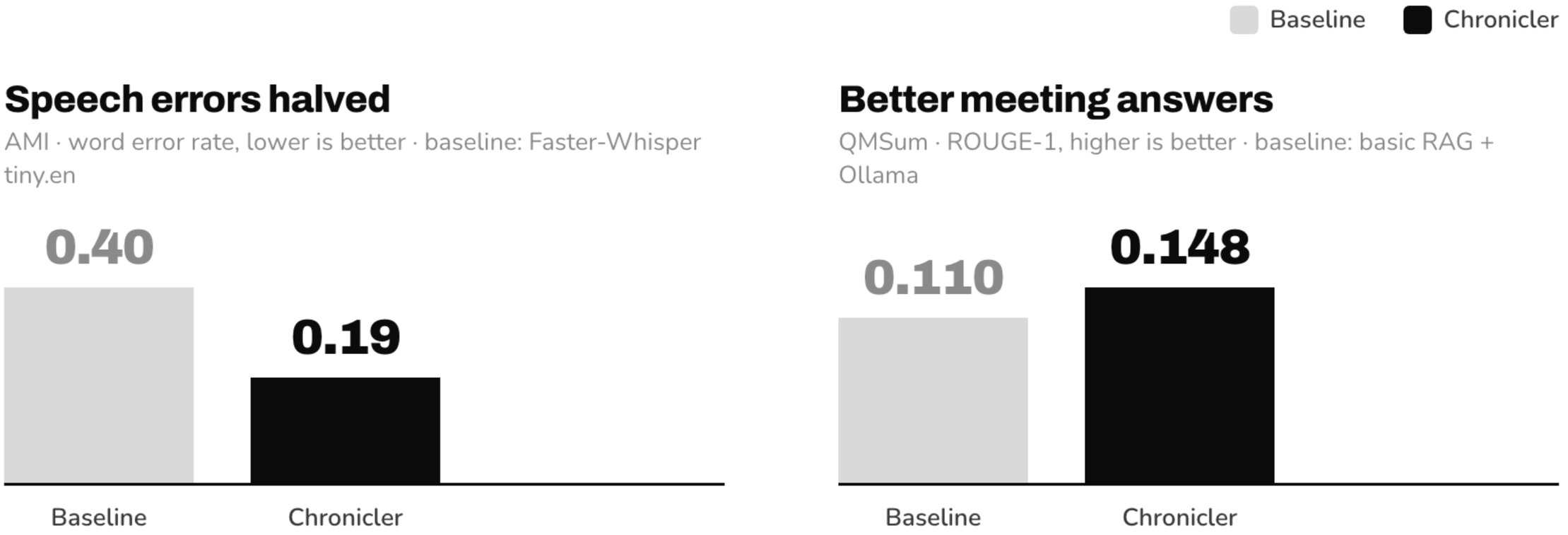
8 minutes back, every meeting.

Median over 3 weekly meetings



PUBLIC BENCHMARKS

Tested beyond our own lab.



35/35 cases ran. Baselines are simpler team-built pipelines with different models, so this compares whole systems, not the agent alone. LongMemEval-S: 0.101 → 0.103.

AGENT TRACE · BENCHMARK RUN

Every step is traced.

240 tool calls across 32 benchmark questions, all saved for audit



WHO CAN CHANGE MEMORY

Only a person approves.



Interpreter

Writes drafts only



Reviewer

Approves and pushes



Chronicler

Commits after approval



Researcher

Read-only

PRIVACY

Nothing leaves campus.

One campus machine

Recordings
and slides

Speech model
local, Mandarin

Qwen
self-hosted GPU

Memory
with full history

Reviewers
lab members

Consent
for every recording



No cloud calls

FUTURE WORK

From one lab to the campus.

NOW

File uploads

Recordings and decks uploaded and reviewed by hand, at one field site.

FIRST, BEFORE SCALING

Live agents in meetings

Notes and answers while the meeting runs, not after.

Your own agents, connected

Personal Claude Code or Codex agents read lab memory through MCP and skills.

MOST IMPORTANT

Data ownership

Every source has an owner, every request a requester. Access is granted, logged and revocable.

NEXT

One lab, many projects

Every project in a lab, with scoped access.

LATER

Library service

Any research group on campus can run its own memory.

NAPAI AGENTIC AI · FINALS
2026.11.21

Chronicler

Thank you. Questions welcome.

TEAM	TRACK	THEME
07 · 牛排	Campus self-proposed	Library, research & knowledge services