

Pramana

pramāṇa (Sanskrit): a valid means of knowing

An open evidence graph built from research papers, and an app to search it.

How much water does one AI answer use?

Three published answers

519 mL

A bottle, per 100-word email

Ren et al., Making AI Less “Thirsty”

sumption. Additionally, GPT-3 needs to “drink” (i.e., consume) a 500ml bottle of water for roughly 10 – 50 medium-length responses, depending on when and where it is deployed.

Washington Post, Sept 2024, using the method of this UC Riverside paper

0.26 mL

Five drops, per Gemini prompt

Measuring the environmental impact of delivering AI at Google scale

Gemini Apps text prompt over one year. We identify that the median Gemini Apps text prompt use television (0.24 Wh) and consumes the equivalent of five drops of water (0.26 mL). While these in activities, reducing the environmental impact of AI serving continues to warrant important attention

Elsworth et al., Google, arXiv, Aug 2025

0.32 mL

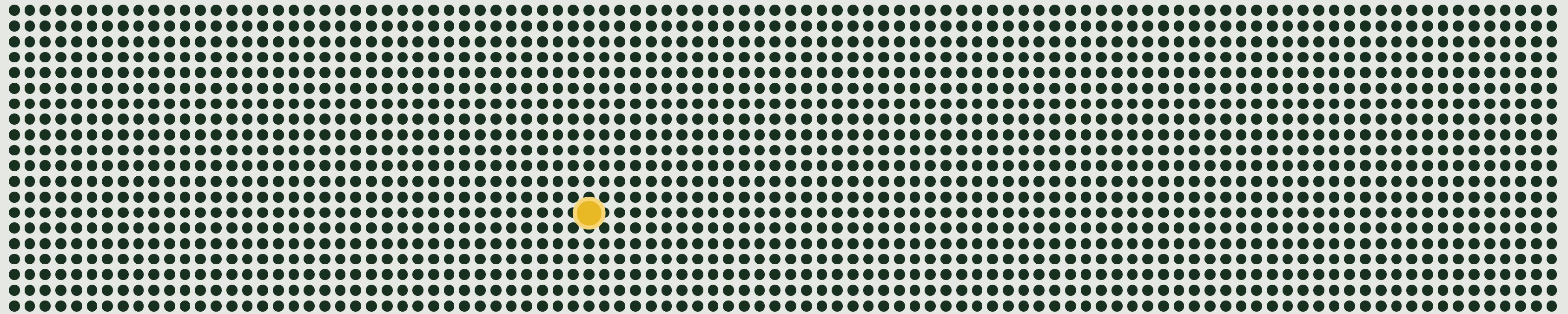
A fifteenth of a teaspoon, per query

The Gentle Singularity

use in a little over one second, or a high-efficiency lightbulb would use in a couple of minutes. It also uses about 0.000085 gallons of water; roughly one fifteenth of a teaspoon.)

Sam Altman, blog, June 2025

The largest answer is 2,000 times the smallest



Each dot is 0.26 mL, Google's figure. All 2,000 together make the Washington Post's 519 mL.

The three figures measure different things

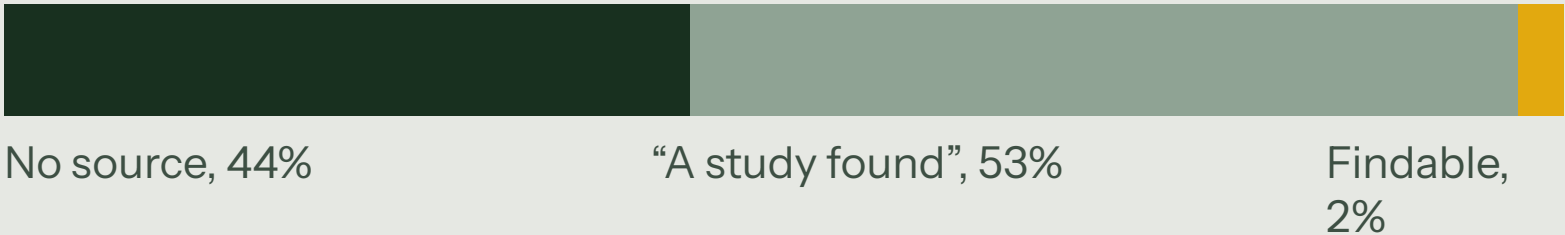
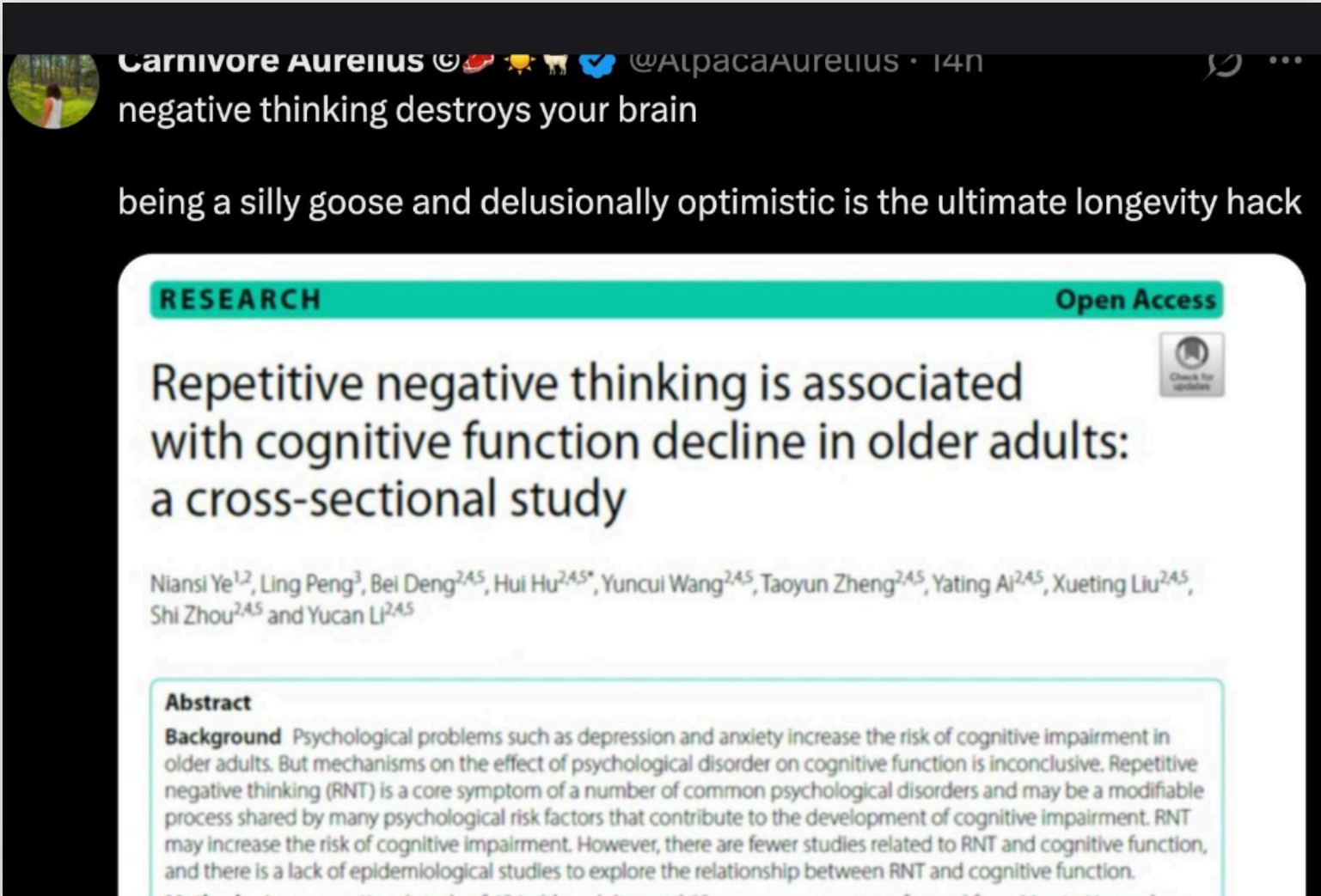
	Washington Post	Google	Sam Altman
Unit	a 100-word email	the median text prompt	the average query
Model, year	GPT-4, 2024	Gemini, 2025	ChatGPT, 2025
Water counted	cooling + power plants	on-site cooling only	not stated
Statistic	one figure	median	average
Method shown	yes, with UC Riverside	yes, arXiv paper	no

The Post counts water used to generate the electricity; Google counts only cooling water at the data center; Altman does not say. The news coverage did not report which.

Most claims reach the public without their source

0 of 634

claims about AI water, in 77 news stories, link to the study they rest on.

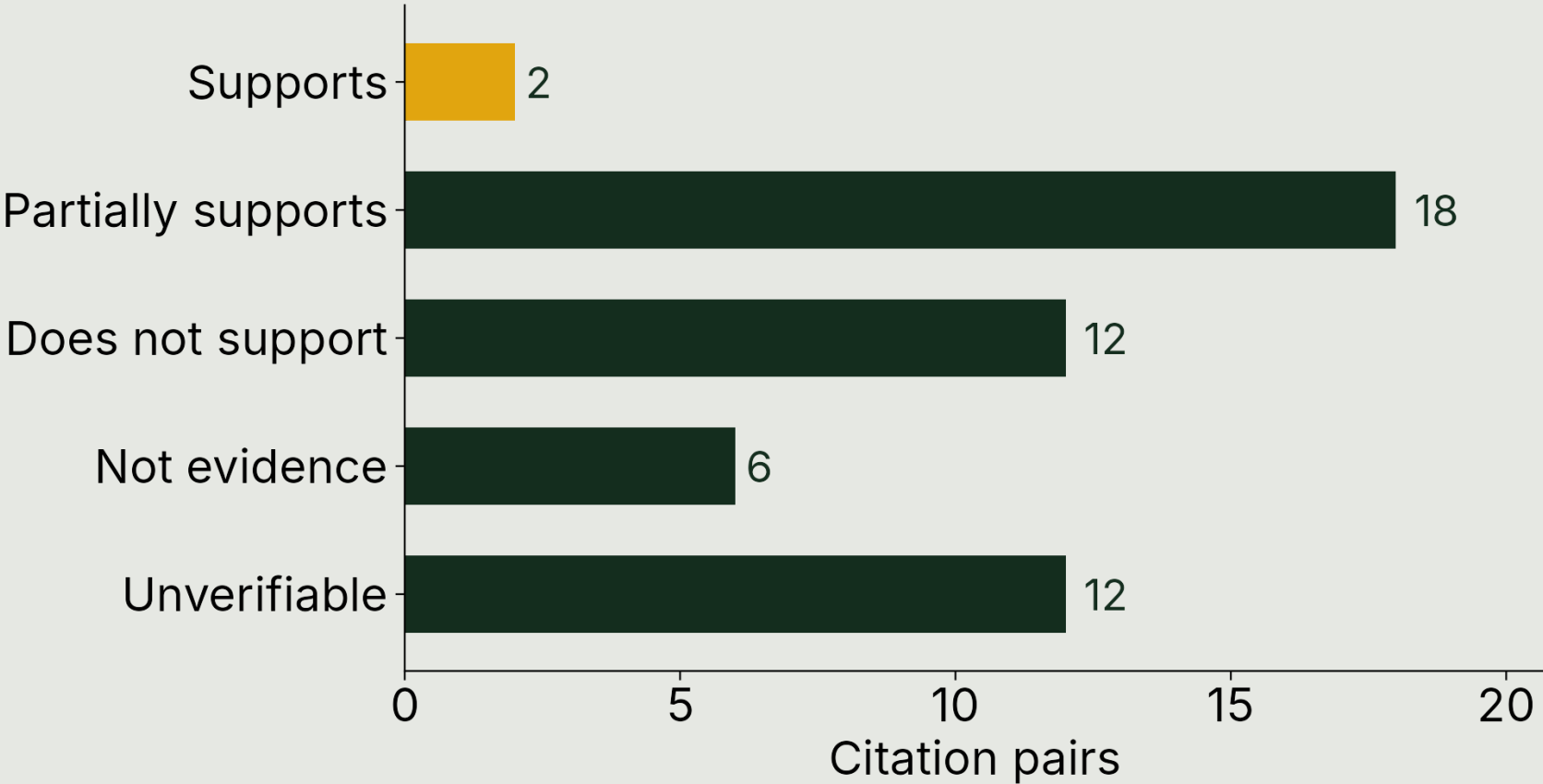


Moses showed me this post on day five. The paper studied 424 adults over 60 in one city, measured once; the most negative thinkers scored 1.6 points lower on a 30-point test.

The drift starts inside the paper

We checked each of the 50 sentences in the Wuhan paper that cite another study against the study it cites.

Two are fully supported. Its line that negative thinking “depletes the brain’s limited resources” cites an eye-tracking experiment on attention.



Pramāṇa, from Nyāya: the means by which knowledge is valid

pratyakṣa

perception

The measurement: 424 people,
one test, one number.

anumāna

inference

What the design permits:
association, not cause.

upamāna

comparison

Every other study on the same
question.

śabda

testimony

The headline, the post, the press
release.

Almost everything we know about science arrives as śabda. Pramana walks each piece of testimony back to the perception and inference underneath it, and sets it beside the comparisons.

A claim is a set of fields, compared one by one

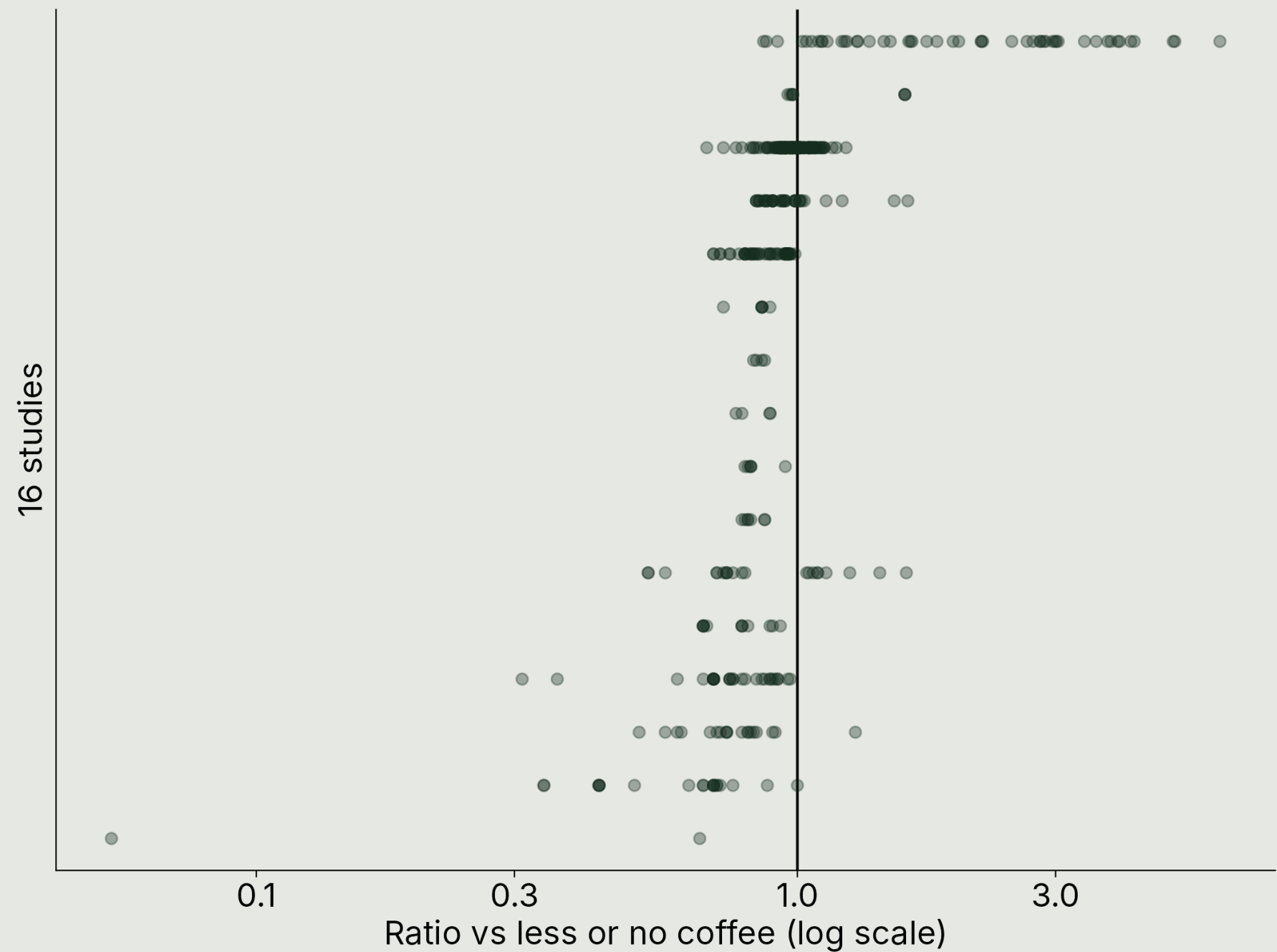
negative thinking destroys your brain

Field	Paper	Post	Difference, fidelity lens v1
Subject	repetitive negative thinking	negative thinking	none
Outcome	MoCA score, a 30-point test	your brain	measure read as organ
Relation	associated with	destroys	association read as cause
Size	1.6 points lower	destroys	inflated
Population	424 adults over 60, Wuhan	you	overgeneralized
Certainty	“further studies needed”	stated as fact	hedge dropped

Coffee and mortality: 16 studies, 437 estimates

Each row is a study; each dot is one estimate it reports, for one dose, subgroup or model. Most sit just below 1.

A headline picks one dot. Pramana shows the whole field, with each dot linked to its sentence.



Pramana app: search the evidence behind a claim

Paste a headline, a post or a sentence.
The app breaks it into subject,
outcome, population, effect size and
certainty, and compares each with
the papers that measured it.

It does not output true or false. It
shows the numbers and where they
came from.

Pramana

Home

Receipts5

Evidence map1,190

Paper network

About

Q How much water does one AI answer use?Ctrl K

Fidelity lens v1Raw data

How much water does one AI answer use?

How much water a single AI response consumes, and what each estimate counts.

Mapped by Claude:electricity zoneDrinking WaterWater SupplyWaterwater footprintdata centerlarge language modelinferencecooling

Subject	Outcome	Relation	Direction	Population	Certainty	
AI model inference	water consumption per response	quantity	not stated	not stated	question	

Summary from the evidence · Claude reading of the records shown · lens: none

The records on AI water use come mostly from modelling studies that combine published energy figures with water-per-energy factors PMC11140606. Estimates per unit of computing span wide ranges because studies differ in whether they count on-site cooling water only or also the water used to generate the electricity W7134815866. Several papers model regional effects, for example on public water supply near data centers W4362707075. The records hold no direct measurement of water per individual chatbot response.

Evidence network1 pair · click a paper

● lower● interval includes no effect● higher● no interval reported

+

-

↗

electricity zone → Water Supply

How much water does one AI answer use?

Drag to move, scroll to zoom

Papers40 matches

Environmental footprints of the data center service sector in Sweden

Cross-sectional environmental footprint estimation (modelled from engineering and grid-mix · PMC11140606

Temperature-Aware Scheduling of LLM Inference in Large-Scale Geo-Distributed Edge Data Centers with Distributed Optimization

not_reported · W7134815866

Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models

not_applicable · W4362707075

How Hungry is AI? Benchmarking Energy, Water, and Carbon Footprint of LLM Inference

Modelling and benchmarking study. Per-query energy, water and carbon are estimated from AP · W4415137654

Navigating the Environmental Paradox of AI: A Decision Framework for Clean Technology Practitioners

Systematic literature review with narrative synthesis; framework development (Environmenta · W7150920407

MITIGATING THE AI WATER CRISIS THROUGH ECO-ADAPTIVE SPATIO-TEMPORAL WORKLOAD SCHEDULING

not_reported · W7158681742

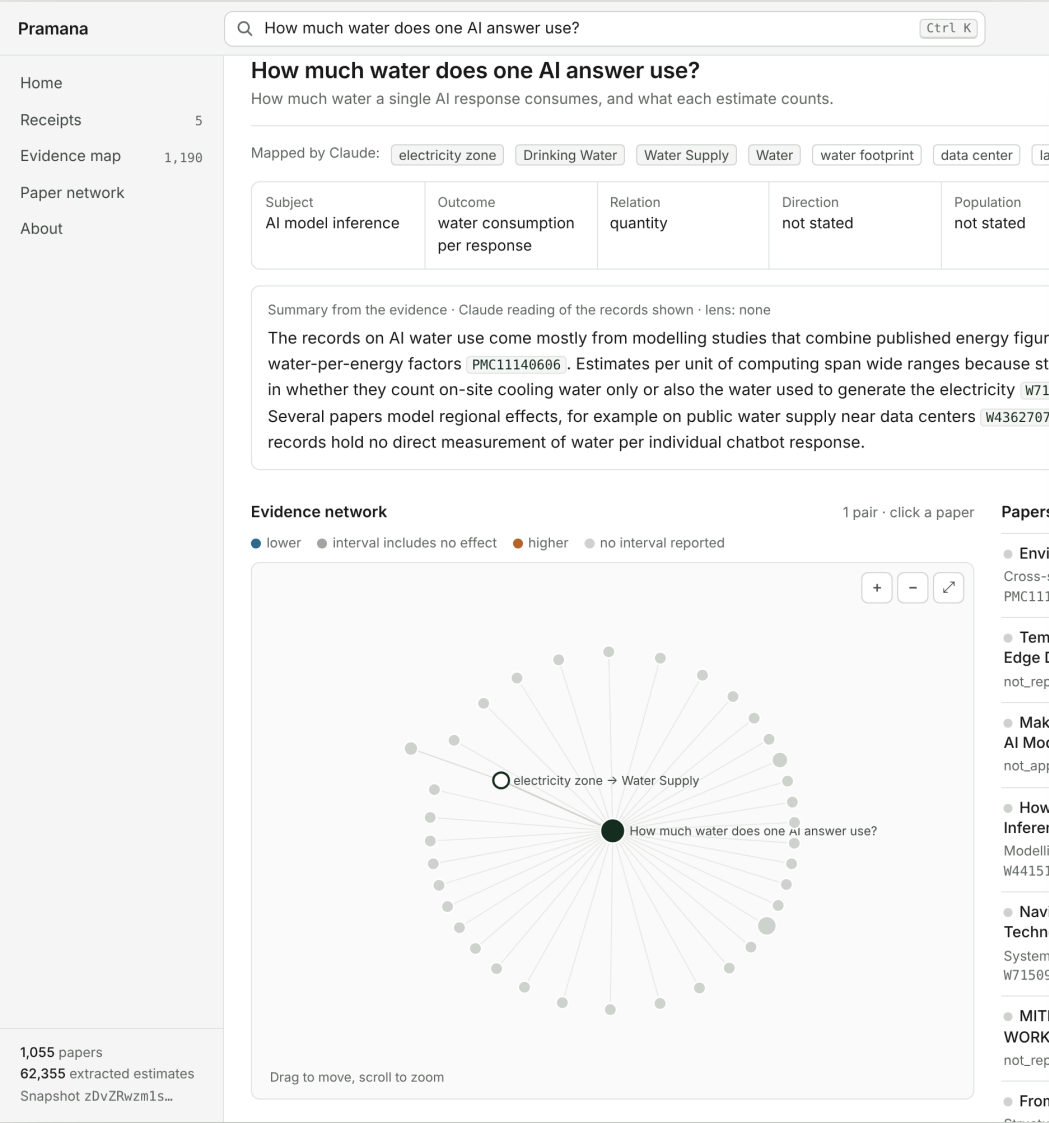
From Cradle to Cloud: A Life Cycle Review of AI's Environmental Footprint

Structural literature review (descriptive mapping of research practices and methods across

Pramana

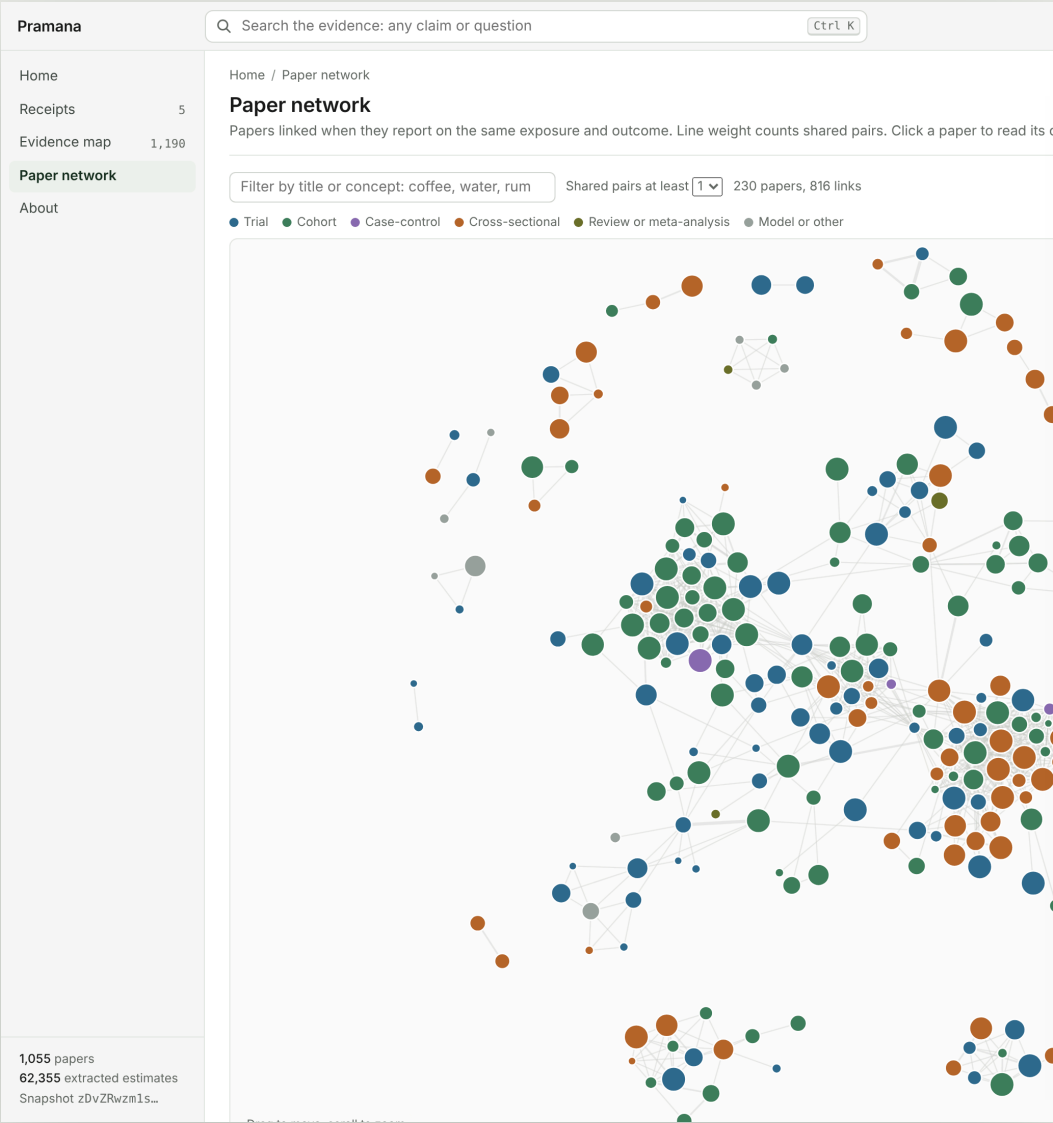
11

Search, networks and papers



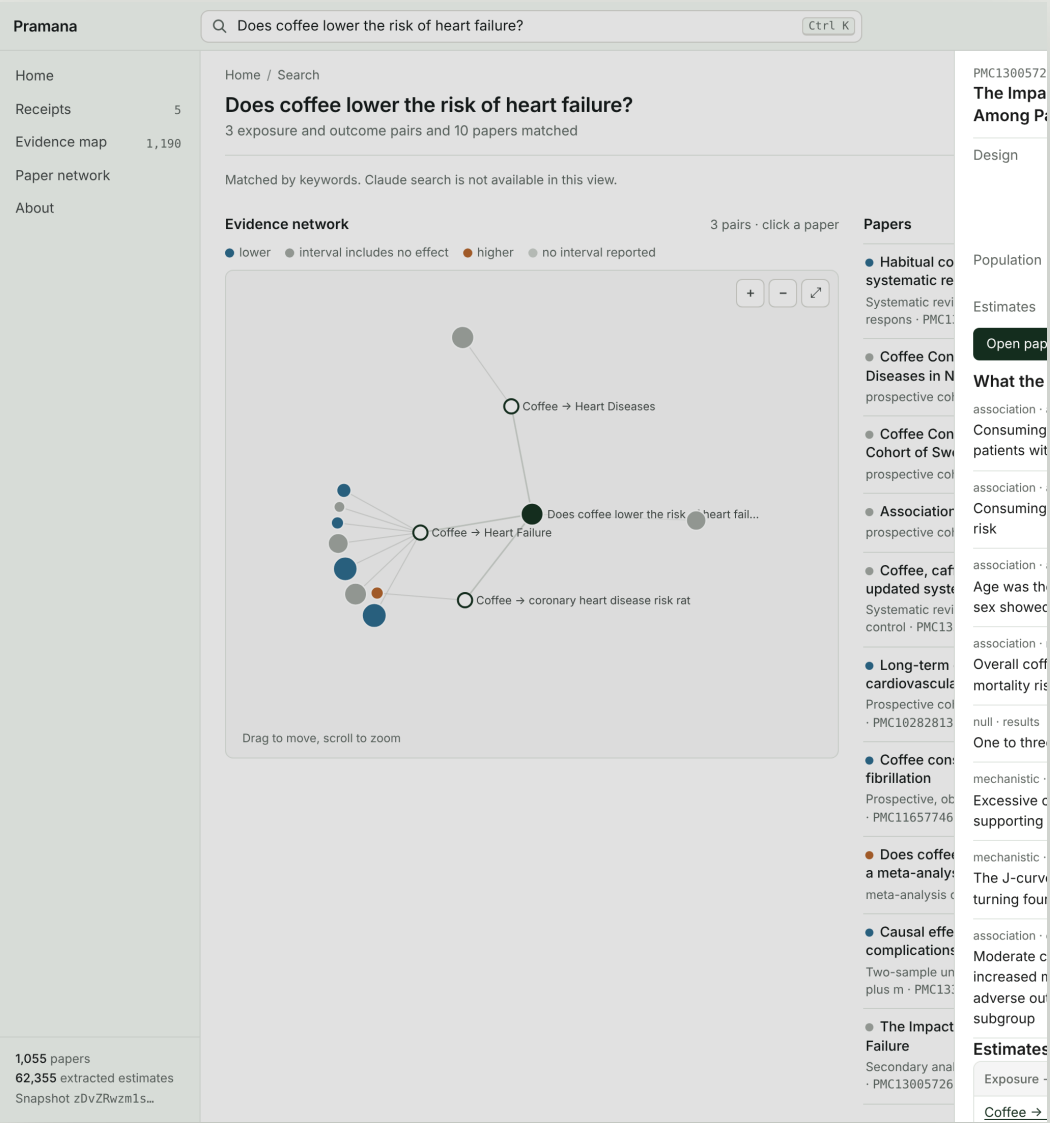
Search

Any question. Claude maps it to the knowledge base, then summarizes only the records it found.



Paper network

Papers linked when they report on the same exposure and outcome.



Paper details

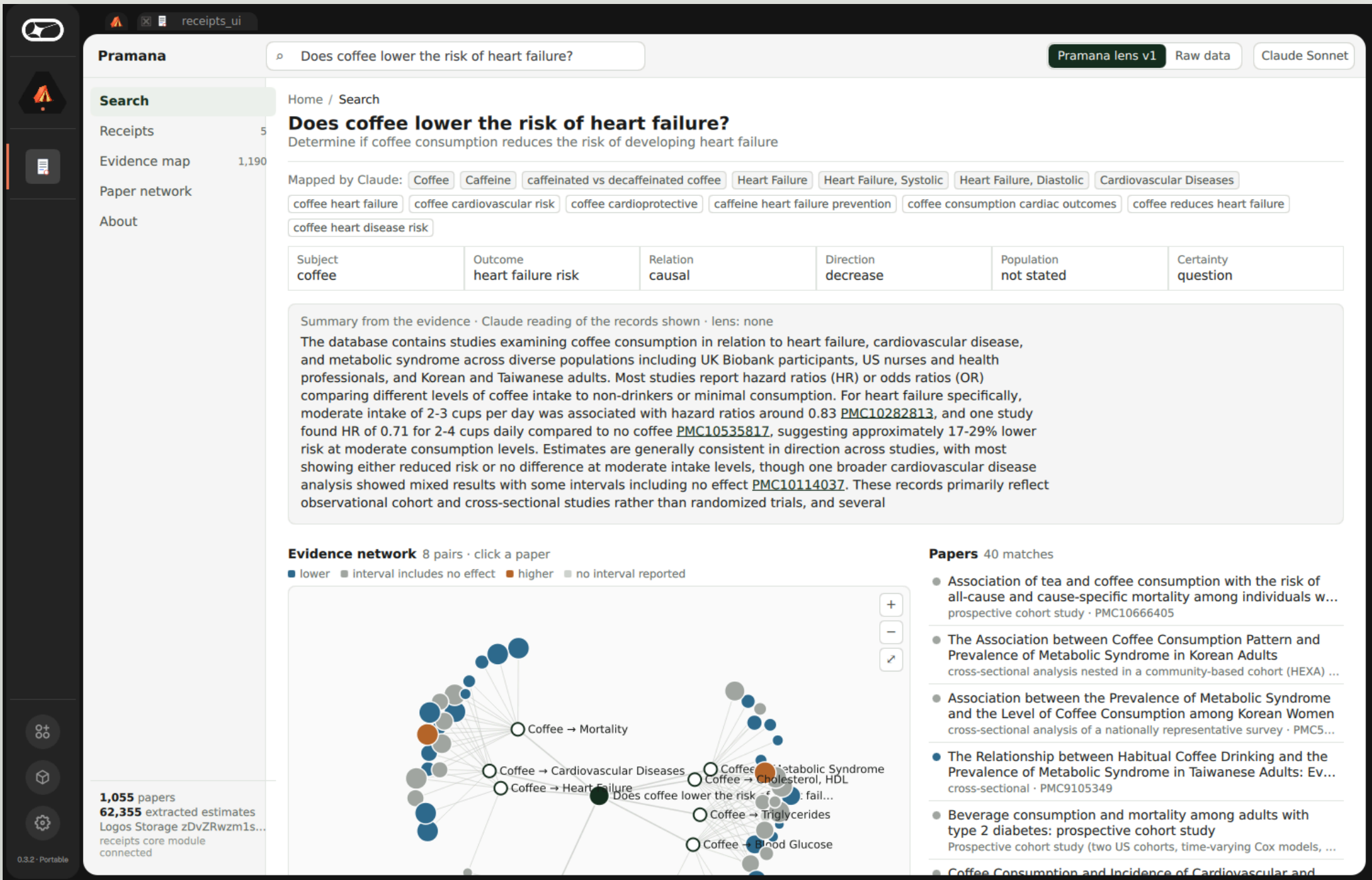
Click any paper for its conclusions, the section each came from, and every estimate.

Running inside Logos Basecamp

The same app as the web, native in Basecamp: Claude search, evidence network, paper detail, receipts and the evidence map, all in QML.

A C++ core module holds the graph in SQLite and carries Claude requests out of the sandbox. Browsing stays on the machine.

The same snapshot is published on Logos Storage under its content ID.



Pramana protocol: shared data, separate lenses

Contributors

Run the pinned pipeline

Same prompts, model versions and schema, each pinned by hash. Contributors bring their own API keys.

Data layer, on Logos Storage

What each paper reported

Estimates, intervals, populations and designs, each with a pointer to its sentence. Two independent extractions must agree. No verdicts.

Lenses

Signed rules that interpret the data

A lens decides what counts as faithful, inflated or causal. Pramana's fidelity lens v1 is one. Anyone can publish another.

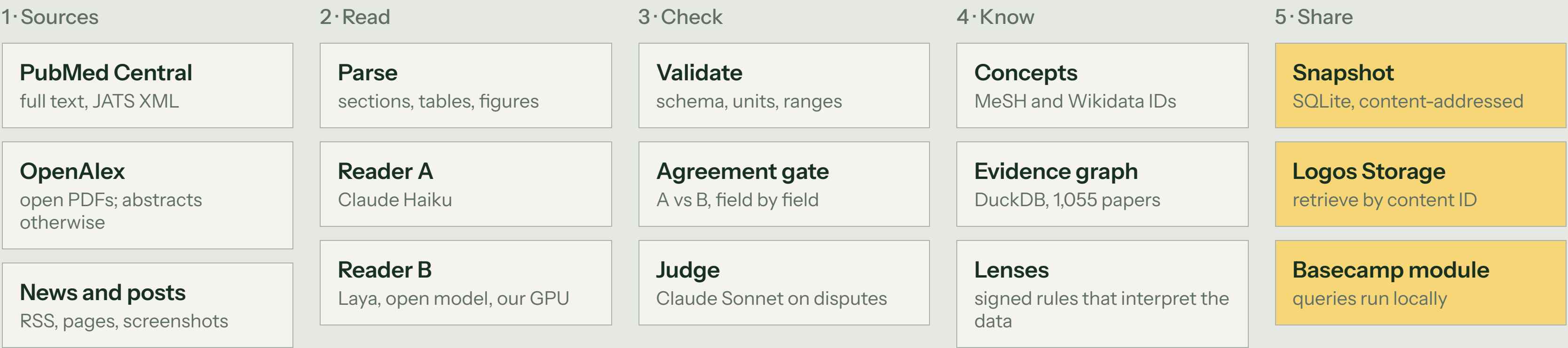
Readers

Any client

Query a local copy through any lens, or none. Where two lenses disagree, the reader sees both.

Judgments never enter the data layer. People can argue about lenses because the facts underneath are shared.

Architecture



Pramana protocol

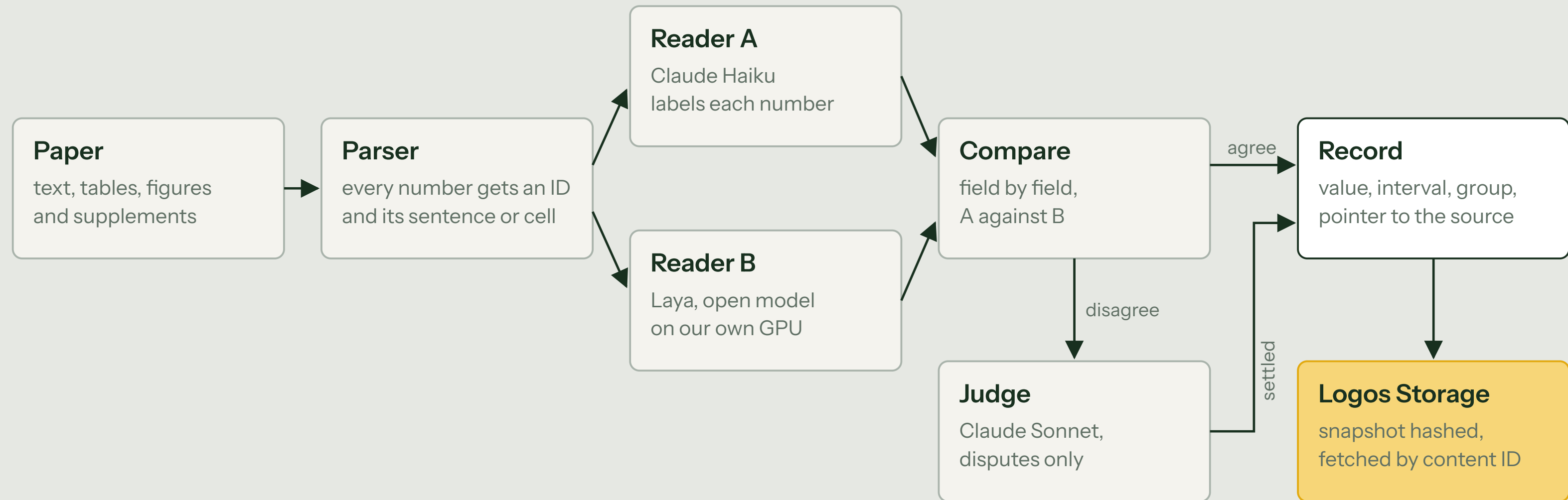
Stages 1 to 4, pinned by hash. Anyone can run them with their own API keys. Lenses sit on top.

Pramana app

One reader of the shared knowledge base.

The pipeline records input and output hashes in an append-only ledger, and every extracted number keeps a pointer to the sentence, table or figure it came from.

How a paper is read



Readers never type a number

They cite the parser's ID, so a value cannot be invented.

Two blind reads

Different models, no shared context.

Judged, never averaged

Each dispute gets a decision and a reason.

Arithmetic as a check

An estimate outside its own interval is flagged.

How we chose the models, and what a paper costs

Model	Tested as	Agreement with blind consensus	Cost
Claude Sonnet + Opus	two readers, 44 papers	0.990 median	too costly to scale
Qwen3 30B, local	reader	κ 0.07, dropped	free
Jev	judge	0.90 exact, 0.99 family	2¢ a paper
Laya, out of the box	judge	0.09	free
Laya, one night of tuning	judge	0.83 to 0.89	free, our GPU
Haiku + Laya, Sonnet judge	production	1,055 papers	13¢ a paper

\$13

for 100 papers

1,000 papers\$134

10,000 papers\$1,340

Measured: \$141.27 of Anthropic API spend for the 1,055-paper build. The open model adds no per-paper fee.

λ Logos

What Logos provides

Without it, every user would have to trust my server, my copy of the data and my logs.

What a receipt needs	Logos does the work	Status
Proof it wasn't edited after printing	Storage: every evidence snapshot has a content ID	live
Looking up a claim reveals nothing about you	Basecamp: the graph runs on your machine	live
Readers add the source a post left out	Messaging: notes travel peer to peer	next
Notes agreed across sides outrank one side	Blockchain: stake on bridging ratings	next

Demo

1

Paste the water question
into Pramana in Basecamp

2

Read the receipt, field by
field

3

Open the sentence each
number came from

4

Match the snapshot's
content ID on Logos
Storage

Then yours: bring any claim about coffee, aging, diabetes or AI water.

Threat model

Threat	Defense
A tampered snapshot	Content ID mismatch; the module refuses it
A model misreads a paper	Two readers, a judge, and every number points to its sentence
Someone logs what you check	Queries never leave your device
A biased verdict	Verdicts live in signed, swappable lenses, never in the data
Brigaded notes	Bridging ratings with stake (next)

Known limits

Open-access papers read in full; paywalled ones from the abstract only, and labeled so.

1,055 papers is a sample, not a library.

Extraction accuracy against human annotators: not yet measured.

English text only.

Who can see what

Who	What they can see
You	Your claim, your receipt, the full evidence graph on your own disk
Me	Nothing. No account, no server, no analytics
Logos Storage nodes	The public snapshot, never the questions asked of it
A note you choose to post	Signed with your key

Lookups run on the user’s machine, so checking a health or political claim leaves no record anywhere. Two of the seven topics came from residents’ requests this week.

Where a receipt changes a decision

Newsrooms

A receipt beside every science claim, checked before it ships.

Community notes

Notes that carry the numbers, not just a link.

AI assistants

Answers that cite receipts instead of memory, through an MCP server.

Regulators

Water permits for data centers in Rajasthan decided on counted water, not headlines.

Courts

Every number in an expert report traced to its source.

Journals and reviewers

A citation audit of each manuscript before it is accepted.

Funders

A map of where evidence is thin, to decide which replications to pay for.

Prediction markets

Questions that resolve on the evidence record.

Users after the residency

Science desks and fact-checkers

Paste a line before it ships. The receipt goes in the footnote.

Researchers writing reviews

The graph is the first pass of a systematic review, at 14¢ a paper instead of a month of reading.

Contributors

Labs and journal clubs add the papers in their field, with their own API keys.

Feedback: disputes filed by readers become labeled examples for retraining the open extraction model. An MCP server lets AI agents query the same knowledge base.

Next three milestones

1 · Release

Protocol spec and contributor tool published. Basecamp app on testnet. Snapshots on Logos Storage. 10,000 papers.

2 · Pilot

Two desks use it for a month. Accuracy measured against human annotators. A public report of what broke.

3 · Adoption

Agreement checking across independent contributors. Notes over Messaging. MCP server. Self-audit. A named operator running a node.

Goal

A public record of what the evidence says,
grown by anyone and owned by no one.

Next time someone tells you how much water your question used, ask which water they counted.

Amruth Deepak Bhat

adb3502.github.io/pramana-protocol · github.com/ad3502/pramana

λ Logos GRAM^Λ