The Compute-Allocation Frontier: A Living, Reproducible Survey of LLM Reasoning and Test-Time Compute

Pengyi Research
P-Research (living corpus pipeline)
github.com/p-research

Draft v0.1 – generated August 19, 2026

Abstract

Large language models are shifting from train-time to test-time scaling: reasoning models now spend deliberate compute at inference through long chains of thought, learned search, and verifier-guided refinement, trained via reinforcement learning with verifiable rewards (RLVR). Yet the field's existing surveys are static snapshots written by individual teams at a single moment, and they treat test-time compute in isolation from the rest of frontier AI. This paper presents a living, reproducible survey of LLM reasoning and test-time compute, generated by a continuously-updating arXiv corpus pipeline that sweeps six pillars weekly into an append-only, evidence-linked paper database. We organize the field under a unifying compute-allocation lens—where compute is spent across training versus inference, in the model versus the environment, and on representation versus reward. Every claim links to a verifiable paper record, making the survey auditable and updatable rather than a frozen artifact.

§1 Introduction

Draft v0.1. Claims cite Evergreen corpus records (`data/papers.jsonl`); only full-text-verified records may survive into the final version.

1 1.1 From train-time to test-time compute

For most of the deep-learning decade, scaling meant training: more data, more parameters, more FLOPs at the optimization step. Since 2023 the field has been re-negotiating that contract. Reasoning models now spend deliberate compute *at inference*: long chains of thought, learned search, and verifier-guided refinement, trained with reinforcement learning on verifiable rewards (RLVR). The frontier has moved from "a bigger model" to "a better allocation of compute" — and the allocation question now spans training vs. inference, model vs. environment, and representation vs. reward. The shift is visible even in abstract-level signals. In our corpus of 574 frontier-AI papers (2023–2026, six pillars), RLVR/GRPO appears in 4 full-text-verified papers from 2024, 18 from 2025, and 6 in the first part of 2026 — while classic chain-of-thought language rises more slowly ($2 \rightarrow 12 \rightarrow 4$). Test-time scaling and verifier/process-reward-model methods follow the same trajectory. (Full corpus: `data/papers.jsonl`; verified subset: 92 records; the first 45 lead-pillar records show a 99% method-tag match rate.)

2 1.2 Why a living survey

Existing surveys on LLM reasoning and test-time compute — e.g. the RLVR survey `arXiv:2501.09686` and the test-time-compute survey `arXiv:2501.02497` — are static snapshots written by a single team at one moment. They are excellent maps, but the territory moves weekly. A static survey cannot show you that multi-agent methods doubled in the verified corpus between 2025 and 2026, or that a method you assumed was reasoning-specific is now converging across pillars.

This survey is generated by a reproducible, continuously-updating corpus pipeline (code on GitHub, append-only database, deterministic structuring, auditable evidence chain from arXiv record to survey claim). It accumulates evidence instead of rotting: every claim links to database records, and the weekly digests archive the evolving frontier.

3 1.3 The compute-allocation lens

We organize the field under one question: **where is compute spent?**

Three axes organize the six pillars of our taxonomy:

1. **train-time vs. test-time** — pre-training and post-training versus inference-time reasoning budgets;
2. **in model vs. in environment** — parameters and activations versus tools, memory, and world interaction;
3. **on representation vs. on reward** — learning what the world is like versus learning what "good" means.

Under this lens, LLM reasoning (pillar 1), agentic deep research (pillar 2), efficient training & inference (pillar 3), RL/alignment/safety (pillar 4), multimodal/world models (pillar 5), and quant×AI (pillar 6) are not separate fields but points on one tradeoff curve. §5 develops this convergence with corpus evidence.

4 1.4 Contributions

1. A **living corpus pipeline** (574 papers, 2023–2026; 92 verified across all pillars papers full-text verified at the time of writing) with deterministic structuring and an auditable evidence chain.
2. A **compute-allocation taxonomy** that places test-time compute in the context of the other five pillars.
3. **Empirical trend sections** derived from the corpus: method migration (RLVR 4→18 verified papers 2024→2025), benchmark saturation (MATH/GSM8K dominance), and cross-pillar convergence (§5–§6).
4. A public, continuously-updated artifact rather than a frozen PDF: digests, database, and survey outline update weekly.

5 1.5 Scope and honesty

- Coverage: arXiv, six pillars, 2023 onward; English-language, abstract

level for unverified records, full text for verified records.

- Not covered: closed commercial systems without papers; non-ArXiv venues

beyond the verified references we cite.

- Every trend number in this draft is a *corpus* statement, not a claim

about the field's ground truth; where full-text verification is pending,
we say so.

§2 Method: A Reproducible, Continuously-Updating Corpus Pipeline

> Draft v0.1 — every claim must eventually cite `data/papers.jsonl` records. This section documents the pipeline that *generates* this survey.

6 2.1 Design principles

1. **Deterministic and auditable.** Every structuring step is a keyword/regex rule, not an LLM call. A claim about a paper can be re-derived from its abstract and full text by anyone.
2. **Append-only evidence.** The paper database (`data/papers.jsonl`) is append-only; verification metadata is merged via atomic rewrites. The evidence chain runs from arXiv entry → structured record → survey claim.
3. **Living, not frozen.** A weekly cron sweeps the frontier; historical backfills cover 2023 onward. Sections cite the corpus *state*, and digests archive the evolving frontier.
4. **Honest about signal strength.** Abstract-level tags are signals; full-text verification is a separate gate. Only verified records may back survey claims.

7 2.2 Corpus construction

- **Sources**: the arXiv API (export.arxiv.org), six pillar queries over

categories cs.AI, cs.LG, cs.CL, cs.CV, cs.CR, cs.MA, cs.DC, q-fin.

- **Windows**: weekly frontier sweep (last 21 days) + non-overlapping

historical backfill windows (0–360, 360–720, 720–1080, 1080–2160 days).

- **Rate limiting & robustness**: ≥3s between requests, retries with backoff,

time-boxed cache, and salvage of complete `<entry>` blocks from truncated responses.

- **Size**: 574 records (2023–2026) at the time of writing; 92 records

records full-text verified.

8 2.3 Deterministic structuring

Each arXiv entry is mapped to a record with:

- **pillar** — the query of origin (six-pillar taxonomy);
- **methods** — 22 keyword families (RLVR/GRPO, Verifier/PRM, CoT, MCTS,

MoE, distillation, quantization, KV cache, long context, preference optimization, interpretability, safety/jailbreak, multi-agent, tool use, memory/RAG, computer use, deep research, world models, video generation, VLM, quant/trading); short acronyms use word-boundary matching to avoid false positives (e.g. "cot" inside "scotland");

- **benchmarks** — a curated list (AIME, GSM8K, MATH, GPQA, SWE-bench,

MMLU, ARC-AGI, AgentBench, GAIA, ...) with word-boundary rules for short names;

- **models** — regex patterns for common model families;
- **key_results** — sentences carrying result markers (outperform, achieves,

state-of-the-art, ...).

9 2.4 Full-text verification (core-corpus gate)

- **Sources**: ar5iv (LaTeXML HTML) first; arXiv native HTML for papers ar5iv

has not converted. arXiv abstract pages masquerading as fulltext are rejected by `ltx_` marker detection.

- **Procedure**: pull full text → re-run the deterministic taggers on it →

record `verified` plus matched-method overlap vs. the abstract-level tags.

- **Status**: 92 records verified across all six pillars; source split ar5iv/arxiv-html

recorded per record.

10 2.5 Citation tracking (in progress)

Semantic Scholar metadata (citationCount, influentialCitationCount) is fetched per arXiv id and stored in `data/citations.jsonl`; anonymous access degrades gracefully and records are only persisted on success.

11 2.6 Reproducibility artifacts

- `presearch` CLI (stdlib-only Python, ≥3.10): `weekly`, `backfill`,

`verify`, `citations`, `db stats`, `survey`.

- GitHub Actions cron (weekly) commits new research back to the repository.
- MIT code, CC BY 4.0 data.

§3 Foundations: From Chain-of-Thought to RLVR

Draft v0.1. `evg-*` ids cite Evergreen corpus records; only full-text-verified records are cited as evidence.

12 3.1 Chain-of-thought as the seed

Chain-of-thought (CoT) prompting showed that allocating inference tokens to intermediate reasoning improves accuracy without touching weights. In our verified corpus, CoT language appears in 2 verified 2024 papers and 12 verified 2025 papers — but increasingly as a *substrate* of learned reasoning rather than as a prompting trick: reasoning traces became the object of optimization, not just an artifact of prompts.

13 3.2 Verifiable rewards

The RLVR recipe — reinforcement learning where the reward is a mechanically checkable function (a final answer, a unit test, a compiler verdict) — removes the need for a learned reward model on many tasks. Verifier/process-reward-model methods grow in the verified corpus from 1 (2024) to 9 (2025) to 5 (partial 2026). Representative verified records:

- `evg-e04110c6e72bd0da` (arXiv:2408.15565, 2024): self-improving code-assisted

mathematical reasoning — an early bridge between code execution and verifiable rewards.

- `evg-a1327781cf590ac8` (arXiv:2508.16921, 2025): reward/verifier machinery

applied beyond math, to affective-alignment diagnosis — a sign of the stack spreading across pillars.

14 3.3 GRPO, PPO, and the policy-optimization zoo

Within RLVR, group-relative policy optimization (GRPO) became the workhorse because it drops the critic model, cutting memory and compute — an *efficiency* choice that doubles as an *allocation* choice (reward vs. representation). Our corpus shows preference-optimization language in 3 verified 2024 papers, 20 verified 2025 papers, and 7 in 2026 — the fastest growth of any method family in the verified subset. Policy optimization now appears far outside its origin: in embodied robot manipulation (`evg-4e48d42ef4f7ad7e`, BATON), in inverse RL for control (`evg-3f12f30b5bc710da`), and in dialogue model-based RL (`evg-964669cd2625570d`).

15 3.4 Search and self-correction

Test-time search (MCTS-style tree search, beam variants, best-of-N) turns inference into an optimization loop over candidate traces. Verified examples: `evg-4e48d42ef4f7ad7e` (subtask-level exploration composed into long-horizon plans), `evg-37a3465e623bd6ff` (AutoSR: symbolic regression by searching "research states"), `evg-6742b060cb568d54` (Le Critique:

privileged value functions guiding LLM reinforcement learning), and `evg-f945ded0c59d6837` (PuzzleJAX, a benchmark designed for reasoning-and-learning under search). Search methods in the verified corpus grow 1 (2024) → 6 (2025) → 3 (partial 2026).

16 3.5 Reward hacking and the verifier-quality wall

As verifiable rewards scale, so does gaming them. The safety pillar already feeds back into reasoning: verified record `evg-204795ac0a186819` (What Do Compliance Detectors Read?) audits activation probes and guard models — detector *representations* are now part of the reasoning stack's failure modes. We flag an open problem here: **verifier quality is the new bottleneck** — a claim we develop with corpus evidence in §7.

17 3.6 What the corpus shows

Method family (verified full-text)	2024	2025	2026 (partial)
RLVR / GRPO	4	18	6
Preference optimization	3	20	7
Chain-of-thought	2	12	4
Verifier / PRM	1	9	5
Search / MCTS	1	6	3
Test-time scaling	0	7	5

Corpus caveat: weekly sweeps weight recent years; 2026 is a partial year. These are corpus frequencies, not field-wide prevalence.

18 3.7 Open threads for §4

- How do verifiers and search interact with *inference-time scaling laws*?
- Where does the budget go — trace length, search width, or verifier depth?
- What transfers from the reasoning pillar to agents, efficiency, and quant×AI (§5)?

§4 Test-Time Compute

Draft v0.1. `evg-*` ids cite Evergreen corpus records (full-text verified unless noted). Formal scaling-law results are summarized from the survey literature; corpus evidence covers the *system designs*.

19 4.1 The allocation question

Test-time compute is the budget a model spends at inference — on tokens, candidate traces, verifier calls, or tool interactions — to raise answer quality without touching weights. The design space decomposes into three questions we use to organize this section:

1. **Where does search happen?** — tree search, beam variants, best-of-N,

or learned routing over candidate traces.
2. **What judges a trace?** — outcome verifiers, process reward models, learned value functions, or external tools.
3. **How does the budget scale?** — inference-time scaling laws and the regime where more compute stops helping.

20 4.2 Search: from fixed traces to composed plans

Verified corpus examples show search moving from "sample many answers" to "search over structured states":

- `evg-4e48d42ef4f7ad7e` (BATON) — subtask-level exploration: each subtask is explored in a cheap short-horizon regime, and long-horizon trajectories are *composed* from stored solutions, turning multiplicative exploration cost (T to the power K) into additive (T times K).
- `evg-37a3465e623bd6ff` (AutoSR) — symbolic regression framed as searching "research states", with a verifier-governed transition model.
- `evg-6742b060cb568d54` (Le Critique) — privileged value functions steer LLM policy search, reintroducing critic-shaped guidance into the RLVR loop.
- `evg-f469e87d0194dc45` (Learning from Diverse Reasoning Paths) — routing and collaboration across diverse reasoning paths rather than single-trace decoding.
- `evg-f945ded0c59d6837` (PuzzleJAX) — a benchmark built for reasoning *and learning* under search, i.e. search as a first-class evaluation axis.

Reading: search is absorbing the *planning* role that classical AI assigned to planners — but now the search space is language traces and the value function is a learned or verifiable judge (§4.3).

21 4.3 Verifiers and process reward models

The judge is the new bottleneck (§3.5). Verified examples:

- `evg-58ed686e9dc31903` (GRIP) — grounded reasoning via information-restricted premises: restricting what a trace may assume is itself a verifier mechanism.
- `evg-e804bfb18640fc6e` (ClawGym II) — black-box RL on an agent harness: the environment as the ultimate verifier for agentic policies.
- `evg-204795ac0a186819` (Compliance Detectors) — audits of activation probes and guard models; detector *representation* quality becomes part of the reasoning stack's failure modes.

For the formal taxonomy of outcome vs. process reward models and their failure modes (reward hacking, credit assignment over long traces), the reader should consult `arXiv:2501.09686` and `arXiv:2501.02497`; our corpus contributes the empirical observation that verifier-style machinery now appears in 15 verified lead-pillar records — and in guard-model audits far outside math reasoning.

22 4.4 Inference-time scaling laws

The quantitative literature (compute-optimal inference, budget allocation, self-refinement limits) is summarized by the test-time-compute survey `arXiv:2501.02497`; several more recent scaling-law surveys exist, and we cite only ids verified against the live arXiv API.
Our corpus's contribution is *systemic* rather than formal: verified records show test-time-scaling language in 7 papers from 2025 and 5 from the partial year 2026 — and, importantly, it has begun to span pillars (lead: 3, efficiency: 1, quant: 1 at abstract level). The scaling-law question we can now ask the corpus: **does the frontier treat test-time compute as a lever (spend more when it helps) or a budget (allocate a fixed envelope)?** — a question flagged for §7.

23 4.5 Summary

Mechanism	Verified examples	What it allocates
Structured search	BATON, AutoSR, Le Critique	trace space × time
Verifier / PRM	GRIP, ClawGym II, Compliance Detectors	judge quality
Routing/collaboration	Diverse Reasoning Paths	width vs. depth
Benchmark pressure	PuzzleJAX	evaluation budget

All records full-text verified; see §2.4 for the verification procedure.

§5 Cross-Pillar Convergence under a Compute-Allocation Lens

Draft v0.1. Counts are abstract-level corpus frequencies (574 records); pillar assignment = query of origin. Verified subsets are noted where used.

24 5.1 Method families that refuse to stay in one pillar

A method family that appears in several pillars at once is either a fad, a foundation, or a bridge. The corpus matrix (papers per pillar) for families spanning ≥ 5 pillars:

Family	Reasoning	Agents	Efficiency	RL/Align	Multimodal	Quant
RLVR / GRPO	29	18	5	9	2	7
Preference optimization	14	1	8	24	17	4
VLM	7	3	8	4	57	2
Interpretability	8	4	5	31	2	7
Memory / RAG	9	8	16	2	5	9
Deep Research	4	8	1	1	3	7
Safety / Jailbreak	11	6	1	12	5	4
Distillation	3	1	37	1	5	3
Video generation	6	2	5	1	19	5

Reading guide: rows are allocation strategies; columns are where compute is spent. The table is generated from `data/papers.jsonl` and regenerated on every sweep.

25 5.2 RLVR is the lingua franca

RLVR/GRPO is the strongest convergence signal in the corpus: present in all six pillars, 79 records at the abstract level. The recipe — optimize a policy against a verifiable reward — no longer belongs to math reasoning. Embodied agents (BATON), black-box agent harnesses (ClawGym II), and on-chain/finance pipelines all adopt it. Under the compute-allocation lens, RLVR is a *reward-side* allocation: you spend compute deciding what "good" means, and the policy follows.

26 5.3 The efficiency pillar is the mirror image

Where reasoning spends compute at inference, efficiency spends it at design time — distillation (37 records), quantization (27), MoE (12), KV-cache work (7) — to *buy back* the inference budget. The two pillars are not opposed; they are the same curve. Distillation from a reasoning teacher to a cheap student is literally test-time-compute crystallized into weights. This is the sharpest single observation of the compute-allocation lens, and §6 quantifies it with trend data.

27 5.4 Memory and verifiers meet in the middle

Memory/RAG spans all six pillars (9/8/16/2/5/9) and verifier/PRM language leaks from reasoning into efficiency and alignment. The mechanism is shared: both are ways to *move compute out of the model* — into stored context (memory) or into an external judge (verifier). Agents, long-context models, and retrieval-augmented pipelines are the same allocation move at different timescales.

28 5.5 Quant×AI as a convergence sink

Quant×AI hosts 56 quant/trading records but also imports the frontier stack: RLVR (7), deep research (7), interpretability (7), memory (9), video generation (5), preference optimization (4). Finance is where the reasoning pillar's *verifiable reward* (P&L, risk limits, backtests) is native — which makes it a natural stress test for §7's open problems.

29 5.6 Caveats

- Abstract-level tags overstate co-occurrence: a paper mentioning

"safety" is not a safety paper. Full-text verification (§2.4) is the gate that turns these frequencies into citable claims; verified-subset versions of this table ship with the final draft.

- Pillar assignment follows the query of origin; cross-pillar numbers

therefore measure *query overlap* plus true method migration. We report both effects rather than hiding either.

30 5.7 What §5 claims, once verified

1. RLVR/GRPO is the most widely shared method family across pillars.
2. Efficiency methods are the counter-allocation to test-time compute.
3. Memory and verifiers are the two dominant "out-of-model" compute moves.
4. Quant×AI is the frontier stack's most demanding verifiable-reward testbed.

31 5.8 Verified-subset table (n = 92, updated 2026-08-19)

Full-text verification now spans all six pillars (45 reasoning, 10 RL/alignment, 10 quant, 9 agents, 9 efficiency, 9 multimodal). The verified-subset matrix — computed only from full-text re-tagged records — confirms the abstract-level convergence claims:

Family	Reasoning	Agents	Efficiency	RL/Align	Multimodal	Quant
Preference optimization	30	2	6	8	7	5
RLVR / GRPO	28	3	6	4	3	3
Memory / RAG	25	7	5	3	3	4
Interpretability	18	3	3	8	7	5
Quant / Trading	19	2	3	6	3	8
VLM	14	4	4	1	8	2
Multi-Agent	11	7	4	3	5	2
Deep Research	12	4	2	6	3	3
Verifier / PRM	15	2	4	1	3	0
Test-time Scaling	12	3	2	2	4	2

Sampling bias disclosure: 45 of 92 verified records come from the lead pillar, so lead-pillar counts dominate. The table proves the span of each family across pillars — not relative pillar intensity. The 99% abstract-to-fulltext tag-match rate (§2.4) is the bridge that lets the abstract level stand in where verification has not yet reached.

§6 Empirical Trends from the Living Corpus

Draft v0.1. All numbers are corpus statements, regenerated from `data/papers.jsonl` on every sweep. Verified subsets are preferred where available; abstract-level counts are labeled as such.

32 6.1 Method migration: the RLVR wave (verified)

Full-text-verified records show the sharpest single trend in the corpus: RLVR/GRPO quadrupled from 4 (2024) to 18 (2025) verified papers, with 6 already in the partial year 2026; preference-optimization language grew $3 \to 20 \to 7$. Chain-of-thought ($2 \to 12 \to 4$) grew more slowly, consistent with the §3 reading that CoT is being absorbed as a *substrate* of learned reasoning rather than remaining a prompting technique.
Abstract-level counts across all six pillars tell the same story with more noise (RLVR: 17/18/21/14 over 2023-2026) — the 2023 elevation is a backfill-window artifact (§6.4).

33 6.2 Benchmark concentration — and where the hard ones hide

In verified lead-pillar full texts, three benchmarks dominate: MATH (13), DROP (12), GSM8K (8). High-difficulty suites (AIME: 2, MMLU: 2) appear far less often — but note the *location* asymmetry: AIME/GPQA almost never appear in abstracts, while MATH/GSM8K do. Abstract-level benchmark mining therefore systematically under-counts hard benchmarks; full-text verification corrects it. For survey purposes this matters: a reader judging "reasoning progress" from abstracts alone will see an easier field than the papers actually evaluate on.

34 6.3 The frontier moves weekly

The weekly digests (`data/weekly/`) archive the moving frontier. The 2026-W34 sweep (the first) ingested 69 papers across the six pillars with zero degraded queries; convergence signals that week included VLM spanning 4 pillars (12 papers) and preference optimization spanning 4 (10 papers). Each digest carries the sweep's convergence signals and degraded-pillar log, so trend claims in this survey can be audited week by week.

35 6.4 Corpus-construction biases (reported, not hidden)

1. **Window weighting.** Weekly sweeps cover the last 21 days; backfills cover fixed windows (0-360, 360-720, 720-1080, 1080-2160 days). Recent years are over-represented; 2026 is partial.
2. **Query-of-origin pillar assignment.** A paper's pillar follows the query that found it, not a content classifier. Cross-pillar counts measure query overlap plus true migration (§5.6).
3. **Abstract-level noise.** Keyword tags overstate co-occurrence; only full-text-verified records (45, method match rate 99%) back hard claims.
4. **arXiv-only.** Closed-model reports, non-English venues, and non-arXiv publications are out of scope (§1.5).

36 6.5 What §6 contributes

- A *method-migration* estimate with verified backing (RLVR 4→18).
- A *benchmark-visibility asymmetry* (hard benchmarks hide in full text).
- An *auditable cadence* — every trend links to weekly digests and DB

records, so the survey's numbers can be re-derived, not just re-cited.

§7 Open Problems, Risks, and Outlook

Draft v0.1. Open problems are ordered by how directly the corpus can test them.

37 7.1 Verifier quality is the new bottleneck

RLVR scales only as far as the reward signal is trustworthy. The corpus already shows the failure mode migrating: verifier/guard-model audits (`evg-204795ac0a186819`), affective-hallucination diagnosis (`evg-a1327781cf590ac8`), and black-box agent-harness RL (`evg-e804bfb18640fc6e`) all circle the same question — **how do you verify the verifier?** Candidate research directions: adversarial verifier evaluation, reward calibration under distribution shift, and process-reward attribution over long traces.

38 7.2 Test-time compute: lever or budget?

§4.4 asked whether the frontier treats test-time compute as a lever (spend more when it helps) or a budget (fixed envelope). The corpus cannot answer it yet — scaling-law papers are under-represented in our queries, and the formal literature lives in the surveys we cite (`arXiv:2501.02497`). A corpus upgrade (dedicated scaling-law query family) is planned.

39 7.3 The closed frontier is invisible to us

Closed commercial systems (o-series, Claude, Gemini reasoning modes) publish no full text; our corpus sees them only through third-party evaluations. Any survey claim about "the frontier" is therefore conditional on the open literature. Mitigation: position claims as "open-literature frontier" and track closed-model evaluations as a separate, clearly-labeled source class.

40 7.4 Citation lag and novelty scoring

2026 papers have immature citation graphs; Semantic Scholar coverage lags submission by weeks. Novelty scoring from citations (§2.5) must therefore blend citation counts with *method-overlap novelty* (how unusual a paper's method vector is within its pillar cohort) — a direction we are implementing.

41 7.5 Reproducibility of a living artifact

A "living survey" is only as living as its hosting. Risks: repo abandonment, API changes, DB corruption. Mitigations in progress: full DB checksums in digests, deterministic regeneration from raw arXiv responses (cached), and periodic Zenodo snapshots of `data/`.

42 7.6 Publication risk (arXiv endorsement)

First-time submissions to cs.AI/cs.LG/cs.CL require endorsement, and arXiv tightened its policy in January 2026. Status and fallback venues are tracked in `positioning.md`; the survey's substance is designed to remain valuable regardless of venue (open data + code + digests).

43 7.7 Outlook

Three bets close the survey:

1. **Reward-side compute** (RLVR + verifiers) keeps diffusing into every pillar — Quant×AI is its most demanding testbed because P&L is the harshest verifier.
2. **Efficiency methods become reasoning methods' mirror**: distillation crystallizes test-time compute into weights; the two pillars converge on one allocation curve.
3. **Memory and verifiers merge**: the next reasoning stack will treat stored context and external judges as one out-of-model compute substrate.

The corpus will re-test these bets every Monday.

Acknowledgments

This survey is generated by an open-source, continuously-updating arXiv corpus pipeline. The full database, weekly digests, and verification metadata are released under CC BY 4.0. All claims cite corpus records; only full-text-verified records are treated as evidence.