Skills Audit Documentation

Progressive-disclosure design pattern for ALAI skill system.

Pillar #4 — Skills Audit Overview

Pillar #4 — Skills Audit Overview

Date: 2026-05-04
MC: #99131
Phase: DESIGN + PoC (Phase 2)
Spec: `agentic-os-pillar4-skills-audit-2026-05-04.md`

Executive Summary

This audit covers the ALAI skill system progressive-disclosure refactor: 79 skills inventoried, top-20 refactor priorities identified, L0-L3 rubric established, and PoC analysis completed for `task-postflight` skill.

Key Findings

- 79 active skill directories on disk; 94 rows in skill-registry.db (32 phantoms, 17 unregistered)
- Only 15 skills have any log invocations in the 19-day measurement window
- `mehanik` (186 hits) and `update-config` (1 hit) appear in logs but have no disk directory — ghost invocations
- 9 skills with references/ dir; 70 are monolithic (L0/L1)
- 12 TOB skills have nested structure — invisible to Claude Code flat-discovery loader
- Highest-priority refactor target: `task-postflight` (5,367 tokens × 21 measured invocations = priority_score 82.05)
- **Reality anchor:** At current ALAI scale (Claude Max flat-rate subscription), context-bloat incremental cost is approximately $0-2/month. The value of this audit is **context window capacity management**, not dollar cost reduction.

Inventory Summary

Metric	Value	Source
Active skill dirs on disk	79	`ls ~/.claude/skills/ | grep -v _archived | wc -l`

Metric	Value	Source
Archived skills	32	`ls ~/.claude/skills/_archived/ | wc -l`
skill-registry.db rows	94	`sqlite3 skill-registry.db 'SELECT COUNT(*) FROM skills;'`
DB-only phantoms	32	comm comparison
Disk-only unregistered	17	comm comparison
Skills with references/ dir	9	find query
Skills with invocations in window	15	log grep
Measurement window	19 days	2026-04-16 to 2026-05-05
Total invocations in window	267	awk filter
Ghost invocations (mehanik)	186	log grep — mehanik not on disk

---

Aggregate Savings Projection

Skills loaded per turn	Tokens saved vs. baseline	% context window recovered (128K window)
Only task-postflight (PASS path)	3,500 tokens	2.7%
task-postflight + prompt-forge	4,700 tokens	3.7%
Top-5 hot-path skills (ranks 1-5)	7,300 tokens	5.7%
All top-20 (max benefit, full session)	19,500 tokens	15.2%
All 79 skills at L3 (theoretical max)	~35,000 tokens	27.3%

Assumes 40-50% body-token reduction per skill post-refactor. These are per-turn estimates derived from body-token reduction; monthly projections without measured session counts would be phantom claims.

Related Documentation

- **Skills Inventory:** Top 20 Priority table with per-skill triage
- **L0-L3 Design Pattern:** Progressive-disclosure rubric and anti-pattern catalog
- **PoC Analysis:** task-postflight refactor demonstration (541 → 194 LOC, 64.7% reduction)

Source spec: `~/system/specs/agentic-os-pillar4-skills-audit-2026-05-04.md` (479 lines)
HiveMind record: TBD
MC: #99131 (ready_for_review)

Skills Inventory — Top 20 Priority

Skills Inventory — Top 20 Priority

Measurement window: 19 days (2026-04-16 to 2026-05-05)
Priority formula: `log10(skill_md_tokens_est) * (1 + invocations_30d)`

Top-20 Refactor Priority Table

Rank	Skill Name	LOC	Tokens	Inv 30d	Est $/mo (current)	Est $/mo (post-L3)	Savings $/mo	Priority Score	Owner
1	task-postflight	541	5,367	21	$0.547	$0.078	$0.469	82.054	john
2	prompt-forge	224	2,372	20	$0.350	$0.070	$0.280	70.877	john
3	plan-with-team	140	1,177	13	$0.105	$0.042	$0.063	42.991	john
4	build-plan	90	923	7	$0.126	$0.063	$0.063	23.722	john
5	ask-board	307	2,623	3	$0.125	$0.038	$0.087	13.675	john
6	build	79	838	3	$0.113	$0.057	$0.056	11.693	john
7	sentinel	105	990	2	$0.116	$0.058	$0.058	8.987	john
8	sync	46	346	2	$0.087	$0.087	$0.000	7.617	john
9	learning-opportunity	165	1,433	1	$0.067	$0.034	$0.033	6.313	john

Rank	Skill Name	LOC	Tokens	Inv 30d	Est $/mo (current)	Est $/mo (post-L3)	Savings $/mo	Priority Score	Owner
10	vault-unlock	117	1,312	1	$0.142	$0.071	$0.071	6.236	john
11	incident-response	122	1,051	1	$0.067	$0.034	$0.033	6.043	john
12	youtube-learning	93	877	1	$0.136	$0.068	$0.068	5.886	john
13	code-review	87	674	1	$0.002	$0.001	$0.001	5.657	john
14	lightrag-upload	87	659	1	$0.117	$0.059	$0.058	5.638	john
15	lightrag-status	101	625	1	$0.121	$0.061	$0.060	5.592	john
16	product-lifecycle	491	5,103	0	$0.081	$0.041	$0.040	3.708	john
17	skill-creator	362	4,911	0	$0.088	$0.044	$0.044	3.691	john
18	doc-coauthoring	375	4,274	0	$0.208	$0.104	$0.104	3.631	john
19	mcp-builder	236	2,457	0	$0.135	$0.068	$0.067	3.390	john
20	plan-build-test	293	2,437	0	$0.099	$0.050	$0.049	3.387	john

est_$/mo (post-L3) = estimate assuming 50% body-token reduction via progressive disclosure

---

Per-Skill Triage — Top 5

#1 task-postflight

- **Current footprint:** 541 LOC / 5,367 tokens
- **Why bloated:** BLOAT_LOC_GT_300 — Contains anomaly decision tree (Section 3), learning-opportunity dispatch template (Section 4), memory writer procedure (Section 5), and failure mode reference table (Section 8) all inline in one file. Most of this content is only needed after an anomaly is detected.

- **Recommended action:** Split — progressive-disclose. Trigger skeleton ≤200 LOC stays in SKILL.md; Sections 3-5+8 move to references/.
- **Predicted savings:** ~3,500 tokens/session on typical PASS flows (63% context reduction); full 5,367 tokens only loaded on ANOMALY path.

#2 prompt-forge

- **Current footprint:** 224 LOC / 2,372 tokens
- **Why bloated:** Single references/agent-briefs.md exists but body still contains full 5-panelist dispatch protocol, model tier assignments, and synthesis rules inline.
- **Recommended action:** Split — move per-panelist briefs and synthesis rules to references/; keep trigger condition and dispatch skeleton in core.
- **Predicted savings:** ~1,200 tokens/session (50% reduction).

#3 plan-with-team

- **Current footprint:** 140 LOC / 1,177 tokens
- **Why bloated:** No references/ dir. Builder/validator role descriptions, round-robin protocol, and output templates are all inline. Frequently invoked (13x in window).
- **Recommended action:** Progressive-disclose — move builder brief and validator brief to references/. Keep selection logic in SKILL.md.
- **Predicted savings:** ~700 tokens/session (59% reduction) across 13 monthly invocations.

#4 build-plan

- **Current footprint:** 90 LOC / 923 tokens
- **Why bloated:** No references/ dir. Moderate size but high invocation frequency (7x). Output templates and TaskList format examples inline.
- **Recommended action:** Progressive-disclose — move TaskList format examples and edge-case handling to references/quick-ref.md.
- **Predicted savings:** ~400 tokens/session (43% reduction).

#5 ask-board

- **Current footprint:** 307 LOC / 2,623 tokens
- **Why bloated:** BLOAT_LOC_GT_300 — 5-agent dispatch briefs are fully inline. Each panelist persona description (50-80 lines each) loads for every board invocation.
- **Recommended action:** Split — move per-panelist briefs to references/panelist-<name>.md. Keep dispatch skeleton (trigger, model tier, synthesis format) in SKILL.md.
- **Predicted savings:** ~1,800 tokens/session (69% reduction).

Full per-skill triage for all 20: See main audit spec `~/system/specs/agentic-os-pillar4-skills-audit-2026-05-04.md` §4.

CSV inventory: `~/system/specs/agentic-os-pillar4-skills-inventory.csv` (79 skills, 20 columns)