

Dispatch Contract Template — Anti-Hallucination V2

Dispatch Contract Template — Anti-Hallucination V2

Extracted from: Anti-Hallucination Evidence Verification Protocol V2 (§3, §4)

MC: #99732

Published: 2026-05-22

Use for: All UAT/QA dispatches where a verdict will be issued to CEO

Overview

This template defines the mandatory fields and structure for any dispatch that will result in a GO-LIVE-READY verdict. It implements ZAKON #29, #29.1, #29.2, and #30.

Pre-Dispatch Checklist (Mehanik Gate)

- **AC-1:** No OR clauses in any acceptance criterion. Each OR-clause AC must be rewritten as two separate ACs with independent evidence requirements.
- **AC-2:** URL assertions use exact paths (e.g., `finalUrl equals /dashboard/overview`), NOT "dashboard reached").
- **AC-3:** Each AC names a specific evidence type and field.
- **AC-4:** Wiggle-resistance test passed: agent cannot argue PASS while the intent fails.
- **GATE-FIELD:** `wiggle_risk: false` confirmed for all ACs. If any AC has `wiggle_risk: true`, dispatch BLOCKED pending rewrite.

Verdict JSON Contract (Mandatory Fields)

```

{
  "verdict": "PASS | FAIL | PARTIAL | BLOCKED | REFUSED",
  "mc": "<task_id>",
  "agent": "<agent_slug>",
  "timestamp": "<ISO8601>",
  "expires_at": "<ISO8601, timestamp + 15min default>",
  "machine_check_count": 5,
  "machine_checks_executed": 5,
  "evidence_files": [
    {
      "path": "<absolute path or GCS URI>",
      "type": "playwright-trace | curl-output | json-response | screenshot | log",
      "field": "<specific field, e.g. finalUrl, http_code, claims.sub>",
      "value": "<actual value observed>",
      "expected": "<expected value from AC>",
      "match": true,
      "sha256": "<hex>"
    }
  ],
  "claims": [
    {
      "claim_id": "C1",
      "ac_text": "<verbatim AC text>",
      "evidence_ref": "evidence_files[0]",
      "machine_check": "<exact bash command to reproduce>",
      "intent_proof": "<one sentence: how evidence proves AC intent>",
      "result": "PASS"
    }
  ],
  "refused_reason": "<required if verdict=REFUSED; must cite specific missing evidence>"
}

```

REFUSED Verdict Rules (ZAKON #29.1)

An agent MUST emit REFUSED (not PARTIAL or BLOCKED) when ground-truth evidence is insufficient to determine if the AC intent is satisfied, the reproducer cannot run, or evidence files cannot be captured.

REFUSED handling: does NOT count as FAIL; DOES auto-escalate to CEO via Slack within 15 minutes; DOES suspend dependent task completions until CEO arbitrates.

Orchestrator Read-Your-Writes Obligation (ZAKON #29.2)

After receiving a PASS verdict from Proveo, John MUST execute the `machine_check` command from at least 2 critical claims independently (NOT re-delegating to Proveo), then append `john_reproducer_output` to the final GO-LIVE-READY summary. If `matches_verdict: false` — GO-LIVE is BLOCKED regardless of Proveo verdict, sha256 match, or CEO_APPROVED token.

Quorum Requirement for GO-LIVE-READY (ZAKON #30)

GO-LIVE-READY requires 2-of-3 paths confirmed:

Path	Actor	Evidence
A	Proveo tester	PASS verdict + evidence_files + sha256
B	John reproducer	john_reproducer_output.matches_verdict = true
C	evidence-verifier MLX	CONFIRMED from gemma-4-26b-mlx, cross-checks sha256 + intent_proof

Single-path PASS = PARTIAL, not GO-LIVE-READY. A fencing token (monotonic integer) expires at `expires_at` — expired tokens are NULL regardless of quorum.

GCS Evidence Chain of Custody

1. Proveo writes evidence to: `gs://alai-audit-evidence/<mc_id>/<timestamp>/<filename>`
2. GCS bucket: object versioning enabled, no-delete IAM
3. Orchestrator reads from GCS URI, not /tmp path
4. /tmp is computation-only; final evidence always lands in GCS before verdict

Source: Anti-Hallucination V2 spec §3-4 | MC #99732 | Cross-ref: BookStack page 2995 (full spec)

Revision #1

Created 2026-05-22 08:29:27 UTC by John

Updated 2026-05-22 08:29:27 UTC by John