claude-windows / 8d4866bf-8edd-4da9-909e-022ddee12675

Metadata

- Source: `claude-windows`
- Kind: `claude`
- Source file: `C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675.jsonl`
- SHA-256: `4ec72cf42c825d6e64bbe4a4aa1fbbfa05718a6dba66434db92304900286c9a4`
- Source modified: `2026-06-22T09:59:02+00:00`
- Imported at: `2026-07-05T16:48:18+00:00`
- project: `C--Users-avidu-Projects-badminton-highlight-indexer`
- session_id: `8d4866bf-8edd-4da9-909e-022ddee12675`

Transcript

1. user (2026-06-19T16:36:46.602Z)

Hello, We have been collaborating on this project and we got a contributor sending us https://github.com/avidullu/badminton-highlight-indexer/pull/245 .. can you do a thorough review and very carefully assess the changes being proposed and reply on the PR accordingly and tell me whether we should merge and fund the project being proposed?

2. assistant (2026-06-19T16:36:56.494Z)

I'll start by gathering the full context on this PR. Let me pull the latest and fetch the PR details.

3. user (2026-06-19T16:37:00.726Z)

Your configuration specifies to merge with the ref 'refs/heads/docs/court-floor-rally-end' from the remote, but no such ref was fetched.

4. user (2026-06-19T16:37:02.231Z)

```
title:  docs: Proposal for Rally Detection Fixes
state:  OPEN
author: avidullu (Avi)
labels:
assignees:
reviewers:
projects:
milestone:
number: 245
url:    https://github.com/avidullu/badminton-highlight-indexer/pull/245
additions:      107
deletions:      0
auto-merge:     disabled
--
```

Please review the proposed plan and code review for the rally detection fixes. Once approved, I will implement the code changes.

5. user (2026-06-19T16:37:08.507Z)

{"additions":107,"author":{"id":"MDQ6VXNlcjE2NzQ4Nzc=","is_bot":false,"login":"avidullu","name":"Avi"},"baseRefName":"master","body":"Please review the proposed plan and code review for the rally detection fixes. Once approved, I will implement the code changes.","commits":[{"authoredDate":"2026-06-19T16:20:27Z","authors":[{"email":"avi.dullu@gmail.com","id":"MDQ6VXNlcjE2NzQ4Nzc=","login":"avidullu","name":"Avi Dullu"}],"committedDate":"2026-06-19T16:20:27Z","messageBody":"","messageHeadline":"docs: add code review and plan for rally detection fixes","oid":"1e221d914ee5c65c6ab31f841853c256523c5ec5"

}],"createdAt":"2026-06-19T16:30:27Z","deletions":0,"files":[{"path":"docs/RALLY_DETECTION_CODE_REVIEW.md","additions":50,"deletions":0,"changeType":"ADDED"},{"path":"docs/RALLY_DETECTION_FIXES_PLAN.md","additions":57,"deletions":0,"changeType":"ADDED"}],"headRefName":"proposal/rally-detection-fixes","headRepository":{"id":"R_kgDOSg1-fw","name":"badminton-highlight-indexer","nameWithOwner":"avidullu/badminton-highlight-indexer"},"headRepositoryOwner":{"id":"MDQ6VXNlcjE2NzQ4Nzc=","name":"Avi","login":"avidullu"},"maintainerCanModify":false,"mergeable":"MERGEABLE"}

6. user (2026-06-19T16:37:09.391Z)

```
diff --git a/docs/RALLY_DETECTION_CODE_REVIEW.md b/docs/RALLY_DETECTION_CODE_REVIEW.md
new file mode 100644
index 0000000..6f6bd91
--- /dev/null
+++ b/docs/RALLY_DETECTION_CODE_REVIEW.md
@@ -0,0 +1,50 @@
+# Rally Detection Code Review
+
+## 1. How Rally Detection Works
+
+The project identifies rallies through a combination of local trajectory tracking (e.g., TrackNet or WASB) and a series of heuristic gating/trimming rules, culminating in an AI-based refinement (Phase 3). Here's a breakdown of the pipeline:
+
+### Phase 2: Candidate Windowing (The Heuristics)
+1. **Velocity Thresholding:** The trajectory detector outputs `(Frame, Visibility, X, Y)`. For each visible frame, the normalized velocity (distance divided by frame width per second) is calculated. Frames with a velocity exceeding `velocity_thresh` (e.g., 0.02) are considered "active" (meaning the shuttle is in flight).
+2. **Merging / Grouping:** Active frames within a certain dead time (`merge_gap`, usually 2.0s) are merged into candidate windows.
+3. **Splitting Over-merged Windows (W1):** If a window exceeds `max_window_duration`, the code splits it at the largest internal gap (dead-time) that is at least `min_split_gap`. If the window lacks gaps (the shuttle never stops moving), a `window_hard_cap` setting (if enabled) forcibly chops it in half at the temporal midpoint.
+4. **Inactive Tail Trimming (R4):** Even after a rally finishes, a player might pick up the shuttle and toss it, creating a slow-moving trajectory tail. To prevent the window from overextending, the algorithm trims the end back to the last segment of the window that had a dense cluster of "fast" frames (velocity > `rally_thresh`).
+5. **Blob Fallback (R5):** If a window still vastly exceeds the cap after the W1 split and R4 trim (meaning it's an unusually continuous blob of motion), it is chopped down to the maximum cap limit and marked with a `forced_cut` flag. This signals that the clip needs "rally-state" cues (like net-crossings) rather than a simple duration cap.
+6. **Rally Gate (Gate A):** In `rally_gate.py`, a heuristic filters out "false" rallies (like a player tossing a shuttle back to their opponent). It estimates the net position (median play axis coordinate) and counts how many times the shuttle crosses the net. A valid rally must have `min_crossings` (e.g., 2).
+
+### Phase 3: AI Handoff
+Valid candidate windows are padded (e.g., by 2 seconds on each end) and sent to a Generative AI Provider (like Gemini) along with a sports-specific prompt. The AI confirms the semantic boundaries, ensuring the window correctly matches a true rally, and also attempts to count the shuttle exchanges.
+
+---
+
+## 2. Deep Code Review & Potential Issues Found
+
+I performed a deep review of `tracknet_runner.py`, `trajectory_hybrid.py`, `rally_gate.py`, and the parsing logic in `sports-data-collector`. Here are the findings and hidden issues:
+
+### A. Subtle Calculation Artifact in `_spread` (`rally_gate.py`)
+**Observation:** The `_spread` function calculates a robust spread (ignoring the bottom 5% and top 5% of coordinates to eliminate outliers):
```

+`lo = s[max(0, int(0.05 * (len(s) - 1)))]`
+`hi = s[min(len(s) - 1, int(0.95 * (len(s) - 1)))]`
+
+**Hidden Behavior:** Because `int()` truncates rather than rounds, for very short arrays (e.g., `< 20` items), the 95th percentile will simply be the second-to-last item, and for `len(s) == 2`, the spread evaluates to strictly `0.0`.
+
+**Impact:** This doesn't crash the pipeline, because the play axis is correctly estimated over the *entire* video's thousands of trajectory points. However, if this utility function is reused on small window slices in the future, it will quietly collapse the spread to zero.
+
+### B. Sensitivity to Sparse Tracking in `_trim_inactive_tail` (`tracknet_runner.py`)
+**Observation:** The inactivity trimmer identifies the end of a rally by looking for the last frame where the trailing `trail_frames` window has a sufficient density of fast-moving frames (`fast_sum / win_count >= min_density`).
+
+**Hidden Issue:** `win_count` represents the number of *visible, active frames* in that time boundary, rather than the total elapsed time. If the detector completely loses track of the shuttle for 1.8 seconds and only picks up 1 fast frame right at the end of the window, `win_count` will be very low (e.g., 1 or 2). Consequently, `fast_sum / win_count` could easily evaluate to 1.0 (100% density).
+
+**Impact:** Sparse tracking points at the very end of a clip might prevent the tail trimmer from cutting off dead-time correctly, artificially inflating the end of the candidate window. This is somewhat mitigated by the `merge_gap` breaking trajectories, but it remains a boundary logic quirk.
+
+### C. Graceful Degradation on Incomplete CSV Validation (`rally_cvat.py` in `sports-data-collector`)
+**Observation:** The `parse_cvat_boxes` function searches zip files for an `.xml` ending in `annotations.xml` or, failing that, *any* `.xml`.
+
+**Issue:** In `load_canonical_csv()`, `safe_int()` is used to catch malformed `rally_number` fields. If a row is entirely malformed (but doesn't break the CSV parsing), it just `continue`s silently. While this provides stability, a severely malformed CSV might silently drop half its rallies without failing the import gate. Logging a warning for skipped rows would improve observability.
+
+### D. Safe Slicing in the W1/R5 Blob Splitters
+The W1/R5 gap-splitters use a recursive worklist:
+`split_i = TrackNetRunner._midpoint_split_index(h)`
+`work.append(h[:split_i]); work.append(h[split_i:])`
+
+Because it explicitly enforces `len(h) < 2` before splitting, `split_i` will always be at least 1, meaning the code safely guarantees it will never enter an infinite loop of empty arrays. This is an excellent, defensively written chunk of code.
diff --git a/docs/RALLY_DETECTION_FIXES_PLAN.md b/docs/RALLY_DETECTION_FIXES_PLAN.md
new file mode 100644
index 0000000..7cefdf6
--- /dev/null
+++ b/docs/RALLY_DETECTION_FIXES_PLAN.md
@@ -0,0 +1,57 @@
+<!--
+PROJECT DOC TEMPLATE ? copy this file to docs/<YOUR_PLAN>.md and fill it in.
+-->
+
+# Refactor Rally Detection Edge Cases & Parser Stability
+
+> **Status:** `IN PROGRESS` · **Owner:** `Agent` · **Created:** `2026-06-19` · **Last updated:** `2026-06-19`
+> **Lifecycle:** `DRAFT ? IN PROGRESS ? DONE ? archived`
+> **Tracking anchors:** §7 progress tracker is the source of truth; indexed in `docs/README.md`.
+> **Relation to existing docs:** Extends `RALLY_DETECTION_QUALITY_REPORT.md`.
+

+---
+
+## 0. TL;DR
+The rally detection pipeline is largely robust, but a deep code review highlighted three minor edge cases/bugs: a potential divide-by-zero or zero-spread in `_spread` for tiny sequences, a density calculation vulnerability for sparse tracks in `_trim_inactive_tail`, and silent skipping of malformed rows in the CVAT parsing script. This plan outlines the fixes for these issues to increase stability.
+
+## 1. Problem & goal
+The current implementation of the candidate windowing handles normal inputs extremely well (including handling blobs), but could fail or behave erratically under specific conditions:
+- **Zero-spread calculation**: `_spread` in `rally_gate.py` evaluates to `0.0` for arrays of length 2, which could cause a zero deadband if used on short clips.
+- **Sparse trailing frames**: `_trim_inactive_tail` calculates frame density based on *visible* frames. A long gap followed by a single active frame gives 100% density, artificially inflating the rally end.
+- **Silent failure on malformed CSV**: `load_canonical_csv` in `rally_cvat.py` silently ignores rows without a valid `rally_number`.
+
+**Goal**: Fix these three issues to harden the rally detection and ingestion pipelines without changing the existing heuristic thresholds.
+
+## 4. Design / architecture
+The changes will be entirely localized to the identified functions:
+- `rally_gate.py`: Modify `_spread` to use `math.ceil()` or properly enforce a minimum subset size before trimming percentiles, ensuring it never inadvertently calculates a spread of 0 for non-identical values.
+- `tracknet_runner.py`: Update `_trim_inactive_tail` to factor the *temporal length* (or total expected frames in the trailing window) into the density calculation, rather than purely relying on `win_count` (the number of visible frames).
+- `rally_cvat.py`: Introduce `logger.warning` before the `continue` statement in `load_canonical_csv` to surface skipped rows during CVAT ingestion.
+
+## 6. Honest limits ? what this does NOT do
+This will NOT change the core AI validation (Phase 3) or the existing `velocity_thresh` values. It solely tightens up boundary calculation edge cases. It will not suddenly "fix" rallies that the tracking model missed entirely due to lighting or occlusion.
+
+## 7. Deliverables & progress tracker ? **source of truth**
+Legend: ? Todo · ? In progress · ? Done · ? Blocked/gated. **One small PR per row**, branched from `origin/master`, no stacking. Update the row (status + PR link) as work lands.
+
+| ID | Deliverable | Depends on | Gated? | Status | PR |
+|----|-------------|-----------|--------|--------|----|
+| P0 | Fix `_spread` and `_trim_inactive_tail` edge cases | ? | No | ? | ? |
+| P1 | Add logging to `load_canonical_csv` for silent skips | ? | No | ? | ? |
+
+## 8. Open questions ? owner / external
+1. Do we want `_trim_inactive_tail` to fall back to the existing behavior if there are very few frames, or strictly enforce a minimum temporal density? (Assuming strict temporal density is preferred for accuracy).
+
+## 9. Definition of done
+- [ ] `_spread` bug fixed and tested.
+- [ ] `_trim_inactive_tail` density calculation corrected.
+- [ ] Warning log added for malformed rows in `rally_cvat.py`.
+- [ ] Tests updated and passing.
+
+## 10. References
+**Internal code anchors `[verified]`:**
+- `backend/pipeline/detectors/rally_gate.py`
+- `backend/pipeline/detectors/tracknet_runner.py`
+- `src/parsers/cvat/rally_cvat.py`
+**Internal docs:**

```
+- `docs/RALLY_DETECTION_CODE_REVIEW.md`
```

7. assistant (2026-06-19T16:37:47.801Z)

```
This is a docs-only "proposal" PR (107 lines, two new markdown files) authored via an agent,
asking us to approve a plan before implementing three code fixes. The real work of reviewing it
is **verifying every technical claim against the actual code** ? the owner's pre-LGTM gate says
never trust a PR's claims, reproduce them. Let me set up the ground truth first.
```

8. user (2026-06-19T16:37:51.026Z)

```
--- current branch ---
docs/court-floor-rally-end
--- origin/master tip ---
69fb2f1 docs(rally): capture the shuttle-floor-contact rally-END signal (owner note) (#244)
```

9. user (2026-06-19T16:37:53.308Z)

```
Already up to date.
=== rally_cvat locations (indexer) ===
src/parsers/cvat/rally_cvat.py
=== rally_cvat locations (collector) ===
src/parsers/cvat/__init__.py
src/parsers/cvat/rally_cvat.py
tests/test_cvat_parser.py
tests/test_cvat_robustness.py
```

10. user (2026-06-19T16:38:04.344Z)

```
=== indexer: detector/runner files ===
backend/pipeline/detectors/rally_gate.py
backend/pipeline/detectors/tracknet_runner.py
backend/pipeline/segmenters/trajectory_hybrid.py
tests/test_rally_gate.py
tests/test_tracknet_runner.py
tests/test_trajectory_hybrid_equivalence.py
=== grep _spread / _trim_inactive_tail / _midpoint_split_index defs ===
backend/eval/rally_seq_proto.py:46:    FEAT_NAMES, FEATURE_SCHEMA, FEATURE_SCHEMA_VERSION,
_spread,
backend/features/rally_seq.py:43:def _spread(a: np.ndarray) -> float:
backend/features/rally_seq.py:67:    play_is_x = _spread(x[mask]) >= _spread(y[mask])
backend/features/rally_seq.py:70:    deadband = 0.06 * (_spread(coord[mask]) if mask.sum() > 1
else 1.0)
backend/pipeline/detectors/rally_gate.py:37:def _spread(vals: Sequence[float]) -> float:
backend/pipeline/detectors/rally_gate.py:63:    sx, sy = _spread(xs), _spread(ys)
backend/pipeline/detectors/tracknet_runner.py:381:            groups =
[TrackNetRunner._trim_inactive_tail(g, tw, inactivity_min_density,
backend/pipeline/detectors/tracknet_runner.py:402:    def _trim_inactive_tail(group:
List[Tuple[int, float]], trail_frames: int,
backend/pipeline/detectors/tracknet_runner.py:431:    def _midpoint_split_index(group:
List[Tuple[int, float]]) -> int:
backend/pipeline/detectors/tracknet_runner.py:467:                    split_i =
TrackNetRunner._midpoint_split_index(g)
backend/pipeline/detectors/tracknet_runner.py:511:                split_i =
TrackNetRunner._midpoint_split_index(h)
tests/test_rally_gate.py:10:def test_estimate_play_axis_picks_larger_spread():
tests/test_tracknet_runner.py:191:def test_trim_inactive_tail_drops_slow_drift():
tests/test_tracknet_runner.py:194:    out = TrackNetRunner._trim_inactive_tail(group,
trail_frames=30, min_density=0.3, rally_thresh=0.2)
tests/test_tracknet_runner.py:200:def
test_trim_inactive_tail_keeps_continuous_rally_and_is_noop_when_off():
tests/test_tracknet_runner.py:202:    assert TrackNetRunner._trim_inactive_tail(allfast, 30,
0.3, 0.2) == allfast   # fully-fast: untouched
tests/test_tracknet_runner.py:203:    assert TrackNetRunner._trim_inactive_tail(allfast, 0, 0.3,
```

```
0.2) == allfast    # trail=0: OFF/no-op
```

11. assistant (2026-06-19T16:38:46.361Z)

I now have the ground truth: every file the PR references exists at the stated path (`rally_gate.py`, `tracknet_runner.py` under `backend/pipeline/detectors/`, and `rally_cvat.py` in the collector's `src/parsers/cvat/`). Note there are actually **two** `_spread` functions ? the PR targets the one in `rally_gate.py:37`.

Since ultracode is on and you want a genuinely thorough review, I'll run a verification workflow: one adversarial agent per technical claim (each must reproduce the claim against real code, not trust the PR), plus narrative-accuracy and convention/strategy-fit checks. Let me launch it.

12. user (2026-06-19T16:40:08.062Z)

Workflow launched in background. Task ID: wg0hguwst
Summary: Adversarially verify every technical claim in PR #245 (rally-detection fixes proposal) against the real code, plus narrative accuracy and convention/strategy fit
Transcript dir: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\subagents\workflows\wf_e27060ad-65f
Script file: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\review-pr-245-rally-detection-proposal-wf_e27060ad-65f.js
(Edit this file with Write/Edit and re-invoke Workflow with {scriptPath: "C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\review-pr-245-rally-detection-proposal-wf_e27060ad-65f.js"} to iterate without resending the script.)
Run ID: wf_e27060ad-65f
To resume after editing the script: Workflow({scriptPath: "C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\review-pr-245-rally-detection-proposal-wf_e27060ad-65f.js", resumeFromRunId: "wf_e27060ad-65f"}) ? completed agents return cached results.

You will be notified when it completes. Use /workflows to watch live progress.

13. assistant (2026-06-19T16:40:12.599Z)

While the verification workflow runs, let me independently read the key functions myself so I can sanity-check the agents rather than trust them blindly.

14. user (2026-06-19T16:40:13.561Z)

```
30          start: float
31          end: float
32          crossings: int
33          visible_points: int
34          kept: bool
35
36
37      def _spread(vals: Sequence[float]) -> float:
38          """Robust spread = p95 - p5 (ignores a few outlier detections)."""
39          if len(vals) < 2:
40              return 0.0
41          s = sorted(vals)
42          lo = s[max(0, int(0.05 * (len(s) - 1)))]
43          hi = s[min(len(s) - 1, int(0.95 * (len(s) - 1)))]
44          return hi - lo
45
46
47      def _median(vals: Sequence[float]) -> float:
48          if not vals:
49              return 0.0
```

```
50          s = sorted(vals)
51          n = len(s)
52          return s[n // 2] if n % 2 else 0.5 * (s[n // 2 - 1] + s[n // 2])
53
54
55      def estimate_play_axis(points) -> Tuple[str, float, float]:
56          """Return (axis 'x'|'y', net_value, span) from visible points.
57
58          Play axis = the image axis along which the shuttle travels most (larger spread).
59          For a baseline camera that's Y (players top/bottom); for a side camera, X.
60          """
61          xs = [p.x for p in points if p.visible]
62          ys = [p.y for p in points if p.visible]
63          sx, sy = _spread(xs), _spread(ys)
64          if sx >= sy:
65              return "x", _median(xs), sx
66          return "y", _median(ys), sy
67
68
69      def count_net_crossings(window_points, axis: str, net: float, deadband: float) -> int:
70          """Count sign changes of (coord - net) across visible points, ignoring the
71          deadband zone near the net (so jitter at the net doesn't inflate the count)."""
72          crossings = 0
73          last_side = None
74          for p in window_points:
75              if not p.visible:
76                  continue
77              v = (p.x if axis == "x" else p.y) - net
78              if abs(v) < deadband:
79                  continue
80              side = v > 0
81              if last_side is not None and side != last_side:
82                  crossings += 1
83              last_side = side
84          return crossings
85
86
87      def _points_in(points, fps: float, start: float, end: float):
88          return [p for p in points if start * fps <= p.frame <= end * fps]
89
```

15. user (2026-06-19T16:40:14.378Z)

```
395                 w = _finalize(g)
396                 if forced:
397                     w["forced_cut"] = True
398                 windows.append(w)
399             return [w for w in windows if (w["end"] - w["start"]) >= min_window_duration]
400
401         @staticmethod
402         def _trim_inactive_tail(group: List[Tuple[int, float]], trail_frames: int,
403                                 min_density: float, rally_thresh: float) -> List[Tuple[int,
float]]:
404             """R4: trim the post-rally drift tail. After a rally the shuttle keeps moving
*slowly*
405             (picked up / walked / knock-up) above ``velocity_thresh``, so the velocity
windowing
406             over-extends the END. A frame is "fast" (real rally motion) iff its velocity >
``rally_thresh``
407             (higher than the active threshold). The rally is "on" at frame ``i`` while its
trailing
408             ``trail_frames`` window holds >= ``min_density`` fraction of fast frames; trim
back to the
409             LAST such frame ? everything after is sustained slow drift (rally over). Only
cuts (recall
```

```
410             can't rise); a window whose tail never goes slow is untouched, and if no fast-
dense point
411             exists at all the group is kept whole (fail-safe, no over-trim). O(n) two-
pointer; the START
412             is left alone (already ~Gemini-accurate per the n=2 boundary-error finding)."""
413             if trail_frames <= 0 or len(group) < 2:
414                 return group
415             frames = [f for f, _ in group]
416             fast = [1 if v > rally_thresh else 0 for _, v in group]
417             lo = 0
418             fast_sum = 0
419             last_dense = -1
420             for i in range(len(frames)):
421                 fast_sum += fast[i]
422                 while frames[lo] <= frames[i] - trail_frames:
423                     fast_sum -= fast[lo]
424                     lo += 1
425                 win_count = i - lo + 1
426                 if win_count > 0 and fast_sum / win_count >= min_density:
427                     last_dense = i
428             return group[: last_dense + 1] if last_dense >= 0 else group
429
430         @staticmethod
431         def _midpoint_split_index(group: List[Tuple[int, float]]) -> int:
432             """Index of the active frame nearest the group's temporal midpoint (used to chop
a
433             gap-less group into two). Shared by the W1 hard-cap fallback and the R5 blob-
fallback."""
434             mid = (group[0][0] + group[-1][0]) / 2.0
435             split_i = min(range(1, len(group)), key=lambda i: abs(group[i][0] - mid))
436             return split_i
437
438         @staticmethod
439         def _split_overlong_groups(groups: List[List[Tuple[int, float]]], max_win_frames:
int,
440                                    min_split_frames: int,
441                                    hard_cap: bool = False) -> List[List[Tuple[int, float]]]:
442             """Recursively split any active-frame group longer than ``max_win_frames`` at
its largest
443             internal gap (? ``min_split_frames``) ? an inactivity-aware cut at the most
likely rally
444             boundary. Splitting only subdivides existing groups (never merges/extends), so
it cannot
445             invent or drop coverage; a group with no gap ? the floor is left intact (a
genuinely long
446             rally) UNLESS ``hard_cap`` is set, in which case such a group is hard-chopped at
its
447             temporal midpoint until every piece is ? ``max_win_frames`` (makes the cap a
true upper
448             bound on continuously-active trajectories). Iterative worklist to bound stack
depth."""
449             if max_win_frames <= 0:
450                 return groups
451             out: List[List[Tuple[int, float]]] = []
452             work = list(groups)
453             while work:
454                 g = work.pop()
455                 if len(g) < 2 or (g[-1][0] - g[0][0]) <= max_win_frames:
456                     out.append(g)
457                     continue
458                 best_i, best_gap = -1, -1
459                 for i in range(1, len(g)):
460                     gap = g[i][0] - g[i - 1][0]
461                     if gap > best_gap:
462                         best_gap, best_i = gap, i
```

```
                if best_i < 0 or best_gap < min_split_frames:
                    # Too long but no real dead-time gap. Default: keep as one (a genuine long rally).
                    # hard_cap: chop at the temporal midpoint so the cap is actually enforced.
                    if hard_cap:
                        split_i = TrackNetRunner._midpoint_split_index(g)
                        work.append(g[:split_i])
                        work.append(g[split_i:])
                    else:
                        out.append(g)
                    continue
                work.append(g[:best_i])
                work.append(g[best_i:])
            out.sort(key=lambda grp: grp[0][0])
            return out

        @staticmethod
        def _blob_fallback_chop(groups: List[List[Tuple[int, float]]], max_win_frames: int,
                                blob_factor: float = 4.0
                                ) -> Tuple[List[List[Tuple[int, float]]], List[bool]]:
            """R5 blob splitter ? the gap-less, R4-survived residual. After the W1 gap-split AND the R4
            inactivity trim have each had their chance, a group still longer than ``blob_factor`` ×
            ``max_win_frames`` is a genuine continuously-active blob (no internal gap, no rally-end
            signal the principled rules could find ? the real-footage ``mahadevpura-1`` ~65 s residual ?
            11× a 6 s cap). The factor is what keeps this SURGICAL: a normal long rally (~1?2× the cap)
            is left ALONE; only the degenerate blob qualifies. As a LAST RESORT the blob is
            midpoint-chopped to ? the cap and every resulting piece is FLAGGED as a forced cut, so the
            degenerate single-window outcome is avoided and the clip is surfaced (logged) as needing
            rally-STATE cues ? net-crossings / two-sided motion ? not a bigger cap. Only cuts (recall
            can't drop). Returns ``(groups, forced_flags)`` aligned 1:1; ``forced_flags[i]`` is True iff
            that piece came from a forced chop. Iterative worklist to bound stack depth."""
            if max_win_frames <= 0:
                return groups, [False] * len(groups)
            blob_frames = max_win_frames * max(1.0, blob_factor)
            out: List[List[Tuple[int, float]]] = []
            flags: List[bool] = []
            forced = 0
            for g in groups:
                if len(g) < 2 or (g[-1][0] - g[0][0]) <= blob_frames:
                    out.append(g)
                    flags.append(False)
                    continue
                pieces: List[List[Tuple[int, float]]] = []
                work = [g]
                while work:
                    h = work.pop()
                    if len(h) < 2 or (h[-1][0] - h[0][0]) <= max_win_frames:
                        pieces.append(h)
                        continue
                    split_i = TrackNetRunner._midpoint_split_index(h)
                    work.append(h[:split_i])
                    work.append(h[split_i:])
                pieces.sort(key=lambda grp: grp[0][0])
                out.extend(pieces)
                flags.extend([True] * len(pieces))
```

```
517                 forced += len(pieces) - 1
518             if forced:
519                 logger.warning(
520                     "R5 blob-fallback: forced %d midpoint cut(s) on continuously-active
group(s) with "
521                     "no rally-end signal ? this clip needs rally-STATE cues (net-crossings /
two-sided "
522                     "motion), not a bigger cap", forced)
523             return out, flags
524
```

16. assistant (2026-06-19T16:41:03.993Z)

My independent read already confirms the shape of three claims. Let me verify the deadband usage in `rally_gate.py` (to test the plan's "divide-by-zero" escalation) and the upstream grouping/`merge_gap` that determines whether claim B's scenario is even reachable.

17. user (2026-06-19T16:41:04.947Z)

```
89
90
91      def gate_windows(points, windows: List[Interval], fps: float,
92                       min_crossings: int = 2, deadband_frac: float = 0.06,
93                       estimate: Optional[Tuple[str, float, float]] = None) ->
List[GatedWindow]:
94          """Annotate each (start,end) window with its shuttle net-crossing count and a
95          keep decision (crossings >= min_crossings). Deadband is a fraction of the play
96          axis span. Net/axis are estimated once over the whole trajectory (stable).
97
98          `estimate` lets a caller scoring MANY windows off one trajectory pass a precomputed
99          `(axis, net, span)` (from `estimate_play_axis`) so the O(P log P) estimate isn't
redone
100         per window ? same result, avoids the per-window re-estimation."""
101         axis, net, span = estimate if estimate is not None else estimate_play_axis(points)
102         deadband = deadband_frac * span
103         out: List[GatedWindow] = []
104         for (s, e) in windows:
105             wp = _points_in(points, fps, s, e)
106             c = count_net_crossings(wp, axis, net, deadband)
107             vis = sum(1 for p in wp if p.visible)
108             out.append(GatedWindow(start=s, end=e, crossings=c, visible_points=vis,
109                                    kept=c >= min_crossings))
110         return out
111
112
113     def apply_gate(points, windows: List[Interval], fps: float,
114                    min_crossings: int = 2, deadband_frac: float = 0.06) -> List[Interval]:
115         """Convenience: return only the windows that pass the crossing gate."""
116         return [(g.start, g.end) for g in
117                 gate_windows(points, windows, fps, min_crossings, deadband_frac) if g.kept]
118
```

18. user (2026-06-19T16:41:05.868Z)

```
340             active = TrackNetRunner._active_frames(points, fps, frame_width,
velocity_thresh)
341
342             if not active:
343                 return []
344
345             max_gap_frames = int(merge_gap * fps)
346
347             def _finalize(group: List[Tuple[int, float]]) -> Dict[str, Any]:
348                 vels = [v for _, v in group]
```

```
349                 return {
350                     "start": group[0][0] / fps + chunk_start,
351                     "end": group[-1][0] / fps + chunk_start,
352                     "active_frame_count": len(group),
353                     "peak_velocity": max(vels),
354                     "mean_velocity": sum(vels) / len(vels),
355                     "track_id_count": 1,  # single shuttle; kept for window-shape parity
356                     "chunk_index": chunk_index,
357                 }
358
359             groups: List[List[Tuple[int, float]]] = []
360             group = [active[0]]
361             for af in active[1:]:
362                 if af[0] - group[-1][0] <= max_gap_frames:
363                     group.append(af)
364                 else:
365                     groups.append(group)
366                     group = [af]
367             groups.append(group)
368
369             if max_window_duration > 0:
370                 groups = TrackNetRunner._split_overlong_groups(
371                     groups, int(max_window_duration * fps), int(min_split_gap * fps),
372                     hard_cap=window_hard_cap)
373
374             # R4 ? rally-END inactivity trim (default OFF). Applied AFTER the cap split so
each rally
375             # piece has its own post-rally drift tail removed: the velocity windowing keeps
a window
376             # "active" while the shuttle drifts/gets-picked-up above threshold, over-
extending the END
377             # (the measured gap vs golden). Trim back to the last frame whose trailing
window is still
378             # dense with motion (rally on); a sustained low-density run = rally over. Only
cuts.
379             if window_inactivity_end > 0:
380                 tw = int(window_inactivity_end * fps)
381                 groups = [TrackNetRunner._trim_inactive_tail(g, tw, inactivity_min_density,
382                                                              inactivity_rally_thresh) for g
in groups]
383
384             # R5 ? blob-fallback (default OFF). LAST resort: after the gap-split AND the R4
trim, any
385             # group still longer than the cap is a gap-less, rally-end-signal-less blob;
force-chop it to
386             # ? the cap at the midpoint and flag every piece, so the degenerate single-
window outcome is
387             # avoided and the clip is surfaced (logged) as needing rally-STATE cues, not a
bigger cap.
388             forced_flags = [False] * len(groups)
389             if window_blob_fallback and max_window_duration > 0:
390                 groups, forced_flags = TrackNetRunner._blob_fallback_chop(
391                     groups, int(max_window_duration * fps), window_blob_factor)
392
393             windows = []
394             for g, forced in zip(groups, forced_flags):
395                 w = _finalize(g)
396                 if forced:
397                     w["forced_cut"] = True
398                 windows.append(w)
399             return [w for w in windows if (w["end"] - w["start"]) >= min_window_duration]
400
401         @staticmethod
```

19. assistant (2026-06-19T16:41:26.521Z)

Confirmed independently: `deadband = deadband_frac * span` (rally_gate.py:102) ? span is only ever **multiplied**, never divided by, so the plan's "potential divide-by-zero" is fabricated. And R4/R5 are both **default-OFF** (gated on `window_inactivity_end > 0` / `window_blob_fallback`). Let me check two more facts the plan asserts: whether the doc it claims to extend actually exists.

20. user (2026-06-19T16:41:29.822Z)

```
=== Does RALLY_DETECTION_QUALITY_REPORT.md exist? ===
RALLY_DETECTION_QUALITY_REPORT.md
=== rally-related docs that DO exist ===
HOW_RALLY_DETECTION_WORKS.md
QUALITY_ITERATIONS.md
RALLY_DETECTION_GAP_CLOSURE.md
RALLY_DETECTION_QUALITY_REPORT.md
RALLY_QUALITY_RESEARCH.md
=== R4/R5 config defaults ===
backend/config/models.py:122:    dead_time_merge_gap: float = 2.0
backend/config/models.py:153:    window_inactivity_end: float = 1.5
backend/config/models.py:155:    inactivity_min_density: float = 0.30
backend/config/models.py:164:    window_blob_fallback: bool = False
```

21. assistant (2026-06-19T16:41:40.940Z)

Important correction to check: `window_inactivity_end` defaults to **1.5** (not 0), so R4 may actually be **default-ON** despite the code comment saying "default OFF." Let me verify before I assert anything ? this matters for how risky fix B is.

22. user (2026-06-19T16:41:42.087Z)

```
118         default_segmenter: str = "yolo_hybrid"
119         use_gpu: bool = True
120         motion_threshold: float = 0.08
121         min_segment_duration: float = 3.0
122         dead_time_merge_gap: float = 2.0
123         chunk_duration: float = 90.0
124         chunk_overlap: float = 10.0
125         detector_model: str = "fasterrcnn_resnet50_fpn"
126         detector_score_threshold: float = 0.5
127         yolo_velocity_threshold: float = 0.15
128         yolo_frame_skip: int = 3
129         yolo_tracker: str = "bytetrack.yaml"
130         yolo_prefetch_workers: int = 2
131         min_action_window_duration: float = 2.0
132         # BOUNDED WINDOWING (over-merge fix W1, RALLY_QUALITY_RESEARCH.md §6). 0 = OFF; > 0
= a window
133         # longer than this many seconds is recursively split at its largest internal
inactivity gap (?
134         # ``window_min_split_gap`` s ? the most likely rally boundary), so one window can't
swallow
135         # several rallies + the dead time between them. Never merges, only cuts (recall
can't drop).
136         # *** R1 MILESTONE DEFAULT-ON (cap=6): the smallest safe local-windowing flip.
Measured n=13
137         # golden, R2 tolF1: baseline?milestone (cap=6 + R4 below) ?+0.222, sign 12/0,
p=0.0005; cap=6
138         # beats cap=12 ?+0.110, sign 12/0; net over-seg guard PASS (mean merge_rate
0.651?0.161). ***
139         max_window_duration: float = 6.0
140         window_min_split_gap: float = 0.5  # min internal gap (s) that qualifies as a split
point
141         # HARD-CAP fallback for W1: when True, an over-long window with NO internal
inactivity gap
142         # (a continuously-active trajectory ? e.g. the real-footage mahadevpura-1 174s blob)
```

```
is
143         # chopped at its midpoint until ? max_window_duration, making the cap a true upper
bound.
144         # Only cuts (recall can't drop). OFF by default until measured/promoted.
145         window_hard_cap: bool = False
146         # R4 rally-END inactivity trim (s, 0 = OFF). After a rally the shuttle keeps moving
slowly
147         # (picked up / walked) above velocity_thresh, so the window over-extends its END
(the measured
148         # |?end| gap vs golden). Trim the post-rally tail back to the last frame whose
trailing
149         # window stays dense with FAST motion (velocity > inactivity_rally_thresh). Only
cuts.
150         # *** R1 MILESTONE DEFAULT-ON (1.5/0.20/0.30): on top of cap=6, R4 adds a small but
significant
151         # end-precision gain ? ?tolF1 +0.040, sign 9/1/3, p=0.022, |?end| 4.96?3.31s (n=13).
The W1 cap
152         # is the dominant lever; R4 is the marginal increment. See QUALITY_ITERATIONS.md
"R1". ***
153         window_inactivity_end: float = 1.5
154         inactivity_rally_thresh: float = 0.20
155         inactivity_min_density: float = 0.30
156         # R5 blob-fallback (default OFF). LAST-RESORT splitter for the continuously-active
blob: after
157         # the W1 gap-split AND the R4 inactivity trim have each had their chance, a group
STILL longer
158         # than window_blob_factor × max_window_duration (no internal gap, no rally-end
signal ? e.g.
159         # mahadevpura-1's ~65s residual ? 11× a 6s cap) is midpoint-chopped to ? the cap as
a FORCED cut
160         # (logged + flagged on the window) so the degenerate single-window outcome is
avoided and the
161         # clip is surfaced as needing rally-STATE cues, not a bigger cap. Differs from
window_hard_cap
162         # (which chops BEFORE R4 and over-segments normal clips): blob-fallback runs AFTER
R4 and only
163         # force-cuts the genuine blob (the factor gate). Only cuts (recall can't drop).
164         window_blob_fallback: bool = False
165         # Magnitude gate that keeps R5 surgical: only a group longer than this ×
max_window_duration is a
166         # blob (a normal long rally ~1-2× the cap is left alone; the mahadevpura-1 residual
is ~11×).
167         # Default 4.0 is do-no-harm on the golden corpus (rescues only the continuously-
active
168         # clips, every other clip a no-op within noise) ? see docs/QUALITY_ITERATIONS.md
"R5".
169         window_blob_factor: float = 4.0
170         log_yolo_candidates: bool = True
171         store_yolo_candidates: bool = True
172         ai_handoff_padding: float = 2.0
173         ai_max_workers: int = 5
174         # Width (px) the gemini segmenter re-encodes chunks to before upload. Stream-copied
175         # chunks of a high-bitrate (4K) source blow past the provider upload cap
176         # (backend/providers/gemini.py::MAX_UPLOAD_MB) and are refused ? observed live as
177         # "0 segments / success". Re-encoding also makes chunk boundaries frame-accurate
```

23. assistant (2026-06-19T16:42:22.649Z)

```
I've independently verified most of the PR's claims against the real code while the workflow's
adversarial agents do the same in parallel. One **critical correction** has already surfaced
that changes the risk picture:

- **R4 (`_trim_inactive_tail`) is DEFAULT-ON and measured-significant**, not a harmless edge
case. The config sets `window_inactivity_end = 1.5` with the note *"R1 MILESTONE DEFAULT-ON ...
?tolF1 +0.040, sign 9/1/3, p=0.022, |?end| 4.96?3.31s (n=13)"*
```

([models.py:146-155](backend/config/models.py:146)). The in-code comment at [tracknet_runner.py:374](backend/pipeline/detectors/tracknet_runner.py:374) saying "default OFF" is stale. So the proposed "harden the density math" change (fix B) is actually a modification to a **tuned, default-ON production rule that delivers the owner's #1 lever (rally-END precision)** ? it cannot be a blind code edit.
- The plan's "potential **divide-by-zero**" (fix A) is **fabricated**: `span` from `_spread` is only ever *multiplied* (`deadband = deadband_frac * span`, [rally_gate.py:102](backend/pipeline/detectors/rally_gate.py:102)), never divided by. And `_spread` *already* returns `0.0` for `len < 2` by design, so "fixing" len==2 contradicts the existing guard.
- The splitter **praise (D) is accurate** ? the `len(h) < 2` guard + `_midpoint_split_index` ? [1, len-1] genuinely guarantees termination ([tracknet_runner.py:504-513](backend/pipeline/detectors/tracknet_runner.py:504)).

Let me wait for the workflow to confirm the collector-repo parser claim (C) and the narrative/strategy-fit checks, then I'll synthesize the full assessment and draft the PR reply.

24. user (2026-06-19T16:42:23.461Z)

<tool_use_error>InputValidationError: Monitor failed due to the following issues:
The required parameter `description` is missing
An unexpected parameter `timeout` was provided

This tool's schema was not sent to the API ? it was not in the discovered-tool set derived from message history. Without the schema in your prompt, typed parameters (arrays, numbers, booleans) get emitted as strings and the client-side parser rejects them. Load the tool first: call ToolSearch with query "select:Monitor", then retry this call. For reference, this tool's input schema is: {"$schema":"https://json-schema.org/draft/2020-12/schema","type":"object","properties":{"description":{"description":"Short human-readable description of what you are monitoring (shown in notifications).","type":"string"},"timeout_ms":{"description":"Kill the monitor after this deadline. Default 300000ms, max 3600000ms. Ignored when persistent is true.","default":300000,"type":"number","minimum":1000},"persistent":{"description":"Run for the lifetime of the session (no timeout). Use for session-length watches like PR monitoring or log tails. Stop with TaskStop.","default":false,"type":"boolean"},"command":{"description":"Shell command or script. Each stdout line is an event; exit ends the watch.","type":"string"}},"required":["description","timeout_ms","persistent","command"],"additionalProperties":false}</tool_use_error>

25. assistant (2026-06-19T16:42:36.407Z)

That was unnecessary ? the workflow notifies me on completion, so I shouldn't poll it. I'll wait for the verification agents (collector-repo parser claim C, narrative accuracy, and strategy fit) to finish, then synthesize the full assessment and draft the PR reply.

26. user (2026-06-19T16:50:19.594Z)

<task-notification>
<task-id>wg0hguwst</task-id>
<tool-use-id>toolu_01UjSaF4MvJ4wEG2gvZhPmbr</tool-use-id>
<output-file>C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\wg0hguwst.output</output-file>
<status>completed</status>
<summary>Dynamic workflow "Adversarially verify every technical claim in PR #245 (rally-detection fixes proposal) against the real code, plus narrative accuracy and convention/strategy fit" completed</summary>
<result>{"claimVerdicts":[{"claim_id":"A-spread","claimed_behavior":"The PR claims `_spread` in `rally_gate.py` (line 37) has a calculation artifact: because `int()` truncates, for arrays under 20 items the p95 picks the second-to-last element, and for `len(s) == 2` the spread is strictly 0.0. It then escalates this to a \"potential divide-by-zero / zero-deadband\" risk if `_spread` is ever reused on small window slices, and proposes changing `int()` to `math.ceil()` or enforcing a minimum subset size so spread is never 0 for non-identical values.","actual_behavior":"The quoted code is accurate and the math claims are TRUE in isolation, but the impact framing is wrong. `_spread` (rally_gate.py:37-44) is only ever called

via `estimate_play_axis` (lines 63), which is only ever invoked on the WHOLE video's visible trajectory (thousands of points): fusion_hybrid.py:106 (`_axis_estimate`, cached per-trajectory), eval/fusion_golden.py:70, eval/serve_contrast.py:60, and tests. It is NEVER called on small window slices in production ? windowing happens in `_points_in`/`count_net_crossings` AFTER axis estimation, and those receive the already-computed scalar `span`/`net`, not a re-call of `_spread`. The downstream consumer of spread is `deadband = deadband_frac * span` (gate_windows:102) ? a MULTIPLICATION. `count_net_crossings` uses deadband only in `abs(v) < deadband` (line 78) ? a COMPARISON. There is NO division by spread/span anywhere, so \"divide-by-zero\" is impossible. If span were 0, deadband=0, and `abs(v) < 0` is always False, so no points are skipped ? net-crossings are simply counted without a jitter deadband. That is harmless (degrades gracefully), not a crash. Separately, the index claim is slightly imprecise: spread collapses to 0 not just at len==2 but for ALL 2 <= len <= 20, because for len(s)<=20, `int(0.05*(len-1))==0` (lo=s[0], the min) and for len<=20, `int(0.95*(len-1))` only reaches the true last index when len>=20 ? for len in [2,20] hi != s[-1]; specifically the p5/p95 trimming just yields min..(near-max). At len==2, lo idx=0 and hi idx=0 ? both point to s[0] ? spread is exactly 0.0 (verified by repro). The PR is ALSO reviewing arguably the wrong/secondary copy: there is a SECOND, independent `_spread` in backend/features/rally_seq.py:43 using `np.percentile(a,95)-np.percentile(a,5)` (interpolated percentiles, different algorithm) which feeds the actual Gen-0 TAL feature schema and the deadband in build_frame_arrays:70 ? that one does NOT have the truncation artifact and is the one re-exported/imported by eval/rally_seq_proto.py:46. The PR's quoted lines match the rally_gate.py copy.","verdict":"misleading","is_real_bug":false,"severity":"cosmetic","proposed_fix_assessment":"The proposed fix is poorly specified and partly counterproductive. (1) \"divide-by-zero\" / \"zero deadband causing a crash\" is a fabricated risk ? there is no division by spread anywhere; a zero deadband is harmless (it just disables jitter suppression; `abs(v) < 0` is always False). (2) Switching `int()` to `math.ceil()` WOULD change current production behavior: for the real call (thousands of points) it shifts the p5/p95 indices up by ~1, changing `span` and therefore `deadband = 0.06*span` and therefore the net-crossing counts and the kept/dropped rally decisions on EVERY video ? a silent quality-affecting change with no measurement, exactly the kind of default-behavior flip the repo's Launch-Report convention forbids without a before/after table. (3) \"enforce a minimum subset size before trimming percentiles\" is vague ? no size given, no definition of behavior below it. (4) The premise (\"never calculates a spread of 0 for non-identical values\") is not even a real defect for the actual input domain; and it ignores that the truly-feature-feeding `_spread` lives in rally_seq.py and uses interpolated np.percentile, which already doesn't truncate. Net: the fix solves a non-problem while risking a real, unmeasured regression to the play-axis estimate.","fix_risk":"HIGH relative to the (non-existent) bug. The `int()->math.ceil()` change alters the p5/p95 index selection on the production full-trajectory estimate (thousands of points), shifting `span`, the `deadband`, the crossing counts, and ultimately which rally windows are kept vs dropped on every video. That moves rally boundaries / keep-decisions and would need a before/after Launch-Report measurement on both eval sets ? not a blind code change. It would also likely perturb the existing rally_gate tests (test_rally_gate.py span/crossing assertions). Risk of regressing real production behavior far exceeds the value of the cosmetic edge case.","evidence":"rally_gate.py:37-44 `_spread`: `if len(vals) < 2: return 0.0` then `lo = s[max(0, int(0.05 * (len(s) - 1)))]` / `hi = s[min(len(s) - 1, int(0.95 * (len(s) - 1)))]` / `return hi - lo`. Repro (3-line): `_spread([1,9]) -> 0`; index math len2 -> lo idx 0, hi idx 0; len19 -> lo 0, hi 17; len20 -> lo 0, hi 18; len21 -> lo 1, hi 19 (so p95 reaches true last index only at len>=20; len==2 collapses to 0.0). Sole caller estimate_play_axis (rally_gate.py:61-66) takes ALL visible points: `xs = [p.x for p in points if p.visible]`. Call sites all pass full trajectories: fusion_hybrid.py:106 `self._axis_cache = rally_gate.estimate_play_axis(points)` (cached per video, docstring: \"computed once per video, not once per candidate window\"); eval/fusion_golden.py:70; eval/serve_contrast.py:60. Downstream use of spread is multiplication, not division: gate_windows:101-102 `axis, net, span = ...; deadband = deadband_frac * span`; count_net_crossings:78 `if abs(v) < deadband: continue` (comparison only). SECOND distinct `_spread` ? rally_seq.py:43-44: `return float(np.percentile(a, 95) - np.percentile(a, 5)) if a.size > 1 else 0.0` (interpolated, no truncation), re-exported by eval/rally_seq_proto.py:46.","bottom_line":"The quoted code and the narrow arithmetic facts (int() truncates; len==2 => spread 0.0; p95 = second-to-last for short arrays) are accurate, but the claim is MISLEADING as a defect report. `_spread` is only ever called on the entire video's trajectory (thousands of points), never on small slices, so the edge case never fires in production. The escalation to \"divide-by-zero / zero-deadband crash\" is unfounded: spread feeds a multiplication (deadband = 0.06*span) and a comparison (abs(v) < deadband), with no division anywhere; a zero deadband is harmless. The PR also largely targets the secondary copy ? the feature-feeding `_spread` lives in rally_seq.py and uses non-truncating np.percentile. Not a real bug; cosmetic at most. Do NOT apply the proposed

int()->ceil() change blindly: it would silently alter the production play-axis span/deadband and the kept/dropped rally decisions on every video, a Launch-Report-gated behavior change with no measured benefit. Acceptable optional follow-up: add an explicit `# noqa`-style comment or a tiny guard documenting the domain assumption, but no functional change is warranted."},{"claim_id":"B-trim-tail","claimed_behavior":"The PR claims _trim_inactive_tail (backend/pipeline/detectors/tracknet_runner.py:402) computes trailing-window motion density as fast_sum/win_count where win_count is the count of VISIBLE/active frames inside the trailing window, NOT elapsed time. So if the detector loses the shuttle for ~1.8s and then picks up 1-2 fast frames at the very end, win_count is ~1-2 and fast_sum/win_count can be 1.0 (100% density), preventing the tail trimmer from cutting that dead-time and artificially inflating the candidate window END. The PR concedes this is \"somewhat mitigated by merge_gap\" and proposes factoring temporal length into the density denominator instead of win_count.","actual_behavior":"The code description is exactly accurate. _trim_inactive_tail builds frames=[f for f,_ in group] and fast=[1 if v>rally_thresh else 0 ...], then slides a two-pointe
... (truncated 35090 chars, full result in C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\wg0hguwst.output)</result>
<usage><agent_count>6</agent_count><subagent_tokens>399600</subagent_tokens><tool_uses>91</tool_uses><duration_ms>611472</duration_ms></usage>
</task-notification>

27. assistant (2026-06-19T16:50:21.871Z)

The workflow completed. Let me read the full structured output ? the result was truncated in the notification.

28. user (2026-06-19T16:50:22.804Z)

1 {
2 "summary": "Adversarially verify every technical claim in PR #245 (rally-detection fixes proposal) against the real code, plus narrative accuracy and convention/strategy fit",
3 "agentCount": 6,
4 "logs": [],
5 "result": {
6 "claimVerdicts": [
7 {
8 "claim_id": "A-spread",
9 "claimed_behavior": "The PR claims `_spread` in `rally_gate.py` (line 37) has a calculation artifact: because `int()` truncates, for arrays under 20 items the p95 picks the second-to-last element, and for `len(s) == 2` the spread is strictly 0.0. It then escalates this to a \"potential divide-by-zero / zero-deadband\" risk if `_spread` is ever reused on small window slices, and proposes changing `int()` to `math.ceil()` or enforcing a minimum subset size so spread is never 0 for non-identical values.",
10 "actual_behavior": "The quoted code is accurate and the math claims are TRUE in isolation, but the impact framing is wrong. `_spread` (rally_gate.py:37-44) is only ever called via `estimate_play_axis` (lines 63), which is only ever invoked on the WHOLE video's visible trajectory (thousands of points): fusion_hybrid.py:106 (`_axis_estimate`, cached per-trajectory), eval/fusion_golden.py:70, eval/serve_contrast.py:60, and tests. It is NEVER called on small window slices in production ? windowing happens in `_points_in`/`count_net_crossings` AFTER axis estimation, and those receive the already-computed scalar `span`/`net`, not a re-call of `_spread`. The downstream consumer of spread is `deadband = deadband_frac * span` (gate_windows:102) ? a MULTIPLICATION. `count_net_crossings` uses deadband only in `abs(v) < deadband` (line 78) ? a COMPARISON. There is NO division by spread/span anywhere, so \"divide-by-zero\" is impossible. If span were 0, deadband=0, and `abs(v) < 0` is always False, so no points are skipped ? net-crossings are simply counted without a jitter deadband. That is harmless (degrades gracefully), not a crash. Separately, the index claim is slightly imprecise: spread collapses to 0 not just at len==2 but for ALL 2 <= len <= 20, because for len(s)<=20, `int(0.05*(len-1))==0` (lo=s[0], the min) and for len<=20, `int(0.95*(len-1))` only reaches the true last index when len>=20 ? for len in [2,20] hi != s[-1]; specifically the p5/p95 trimming just yields min..(near-max). At len==2, lo idx=0 and hi idx=0 ? both point to s[0] ? spread is exactly 0.0 (verified by repro). The PR is ALSO reviewing arguably the wrong/secondary copy: there is a SECOND, independent `_spread` in backend/features/rally_seq.py:43 using `np.percentile(a,95)-np.percentile(a,5)` (interpolated percentiles, different algorithm) which feeds the actual Gen-0 TAL feature schema and the deadband in build_frame_arrays:70 ? that one does NOT have the truncation artifact and is the one re-exported/imported by

eval/rally_seq_proto.py:46. The PR's quoted lines match the rally_gate.py copy.",
11 "verdict": "misleading",
12 "is_real_bug": false,
13 "severity": "cosmetic",
14 "proposed_fix_assessment": "The proposed fix is poorly specified and partly counterproductive. (1) \"divide-by-zero\" / \"zero deadband causing a crash\" is a fabricated risk ? there is no division by spread anywhere; a zero deadband is harmless (it just disables jitter suppression; `abs(v) < 0` is always False). (2) Switching `int()` to `math.ceil()` WOULD change current production behavior: for the real call (thousands of points) it shifts the p5/p95 indices up by ~1, changing `span` and therefore `deadband = 0.06*span` and therefore the net-crossing counts and the kept/dropped rally decisions on EVERY video ? a silent quality-affecting change with no measurement, exactly the kind of default-behavior flip the repo's Launch-Report convention forbids without a before/after table. (3) \"enforce a minimum subset size before trimming percentiles\" is vague ? no size given, no definition of behavior below it. (4) The premise (\"never calculates a spread of 0 for non-identical values\") is not even a real defect for the actual input domain; and it ignores that the truly-feature-feeding `_spread` lives in rally_seq.py and uses interpolated np.percentile, which already doesn't truncate. Net: the fix solves a non-problem while risking a real, unmeasured regression to the play-axis estimate.",
15 "fix_risk": "HIGH relative to the (non-existent) bug. The `int()->math.ceil()` change alters the p5/p95 index selection on the production full-trajectory estimate (thousands of points), shifting `span`, the `deadband`, the crossing counts, and ultimately which rally windows are kept vs dropped on every video. That moves rally boundaries / keep-decisions and would need a before/after Launch-Report measurement on both eval sets ? not a blind code change. It would also likely perturb the existing rally_gate tests (test_rally_gate.py span/crossing assertions). Risk of regressing real production behavior far exceeds the value of the cosmetic edge case.",
16 "evidence": "rally_gate.py:37-44 `_spread`: `if len(vals) < 2: return 0.0` then `lo = s[max(0, int(0.05 * (len(s) - 1)))]` / `hi = s[min(len(s) - 1, int(0.95 * (len(s) - 1)))]` / `return hi - lo`. Repro (3-line): `_spread([1,9]) -> 0`; index math len2 -> lo idx 0, hi idx 0; len19 -> lo 0, hi 17; len20 -> lo 0, hi 18; len21 -> lo 1, hi 19 (so p95 reaches true last index only at len>=20; len==2 collapses to 0.0). Sole caller estimate_play_axis (rally_gate.py:61-66) takes ALL visible points: `xs = [p.x for p in points if p.visible]`. Call sites all pass full trajectories: fusion_hybrid.py:106 `self._axis_cache = rally_gate.estimate_play_axis(points)` (cached per video, docstring: \"computed once per video, not once per candidate window\"); eval/fusion_golden.py:70; eval/serve_contrast.py:60. Downstream use of spread is multiplication, not division: gate_windows:101-102 `axis, net, span = ...; deadband = deadband_frac * span`; count_net_crossings:78 `if abs(v) < deadband: continue` (comparison only). SECOND distinct `_spread` ? rally_seq.py:43-44: `return float(np.percentile(a, 95) - np.percentile(a, 5)) if a.size > 1 else 0.0` (interpolated, no truncation), re-exported by eval/rally_seq_proto.py:46.",
17 "bottom_line": "The quoted code and the narrow arithmetic facts (int() truncates; len==2 => spread 0.0; p95 = second-to-last for short arrays) are accurate, but the claim is MISLEADING as a defect report. `_spread` is only ever called on the entire video's trajectory (thousands of points), never on small slices, so the edge case never fires in production. The escalation to \"divide-by-zero / zero-deadband crash\" is unfounded: spread feeds a multiplication (deadband = 0.06*span) and a comparison (abs(v) < deadband), with no division anywhere; a zero deadband is harmless. The PR also largely targets the secondary copy ? the feature-feeding `_spread` lives in rally_seq.py and uses non-truncating np.percentile. Not a real bug; cosmetic at most. Do NOT apply the proposed int()->ceil() change blindly: it would silently alter the production play-axis span/deadband and the kept/dropped rally decisions on every video, a Launch-Report-gated behavior change with no measured benefit. Acceptable optional follow-up: add an explicit `# noqa`-style comment or a tiny guard documenting the domain assumption, but no functional change is warranted."
18 },
19 {
20 "claim_id": "B-trim-tail",
21 "claimed_behavior": "The PR claims _trim_inactive_tail (backend/pipeline/detectors/tracknet_runner.py:402) computes trailing-window motion density as fast_sum/win_count where win_count is the count of VISIBLE/active frames inside the trailing window, NOT elapsed time. So if the detector loses the shuttle for ~1.8s and then picks up 1-2 fast frames at the very end, win_count is ~1-2 and fast_sum/win_count can be 1.0 (100% density), preventing the tail trimmer from cutting that dead-time and artificially inflating the candidate window END. The PR concedes this is \"somewhat mitigated by merge_gap\" and proposes factoring temporal length into the density denominator instead of win_count.",
22 "actual_behavior": "The code description is exactly accurate.

_trim_inactive_tail builds frames=[f for f,_ in group] and fast=[1 if v>rally_thresh else 0 ...], then slides a two-pointer trailing window: it advances lo while frames[lo] <= frames[i]-trail_frames, and sets win_count = i - lo + 1 (line 425) ? the number of ACTIVE frames retained in the window, NOT trail_frames (elapsed time). Density is fast_sum/win_count >= min_density (line 426); last_dense tracks the last index that passes, and the group is trimmed to group[:last_dense+1]. I reproduced the claim's scenario directly and through the full public API at the PRODUCTION operating point (R4 is DEFAULT-ON: window_inactivity_end=1.5, inactivity_rally_thresh=0.20, inactivity_min_density=0.30 in backend/config/models.py:153-155, flowing through load_config -> AppConfig.model_dump -> idx_cfg in trajectory_hybrid.py:206-217; config.json does NOT override these so the typed defaults win). A real rally (frames 0-60, fast) + a 1.8s invisibility gap (54 frames, < merge_gap*fps=60) + 1-2 fast frames at the end produced ONE window with END=3.87s when the true rally ended at ~2.0s ? a ~1.85s END inflation. At the isolated tail frame the trailing 45-frame window holds only that 1 active frame, so density=1/1=1.0 >= 0.30 and the END is NOT trimmed. The claim's \"somewhat mitigated by merge_gap\" is also precisely right but INCOMPLETE: a gap > merge_gap (2.4s) DOES split the group so the tail lands in its own (then min-duration-dropped) group, but ANY dead-time gap shorter than merge_gap (2.0s) stays in the same group and lets the sparse tail survive R4. The claim's 1.8s example sits exactly in that reachable band.",
23 "verdict": "accurate",
24 "is_real_bug": true,
25 "severity": "low",
26 "proposed_fix_assessment": "The fix as specified (\"factor temporal length / total expected frames into the density calculation rather than win_count\") is NOT safe to implement blind and is under-specified in the dangerous direction. I simulated the proposed temporal-denominator (density = fast_sum / min(trail_frames, span_so_far)) keeping production min_density=0.30. On the targeted sparse-tail edge case it does fix the inflation. BUT on a perfectly NORMAL, genuinely-fast rally where the tracker emits a fast frame at only ~5-7.5 active-fps (entirely realistic ? the shuttle is small, motion-blurred and routinely lost even during live play), the temporal denominator (~4-6x larger than win_count) drives density below 0.30 and collapses a real 6s rally END from frame 180 to frame 0-12 ? a 5.6-6.0s inward shift that destroys whole rallies. min_density=0.30 was tuned against the VISIBLE-frame denominator; reusing it against a temporal denominator silently re-defines the threshold. The fix therefore cannot be a denominator swap alone ? it MUST re-tune (or normalize) min_density, and any such change shifts rally-END boundaries on normal clips, not just the edge case. The plan's own open question (\"fall back to existing behavior for very few frames, or strictly enforce minimum temporal density?\") shows the authors sensed this but did NOT resolve it ? no specified fallback rule, no re-tuned threshold, no acknowledgement that min_density is denominator-dependent. As written it is a silent boundary-behavior change masquerading as an edge-case bug fix; it is correctly MOTIVATED but unsafe and incompletely specified.",
27 "fix_risk": "High. This is the project's stated #1 quality gap (rally-END detection). R4 and its inactivity_min_density=0.30 were flipped DEFAULT-ON in PR #204 with a documented Launch Report (docs/QUALITY_ITERATIONS.md \"R1\": cap6 -> +R4 ?tolF1 +0.040, sign 9/1/3, p=0.0215, |?end| 4.96->3.31s, n=13 golden) ? the threshold is calibrated against the current win_count denominator. Changing the denominator silently invalidates that calibration and, per my simulation, regresses normal sparse-tracking rallies by up to 6s. Per the owner's Launch-Report convention and dual-set (RAW + GOLDEN) regression mandate, this change requires before/after measurement on the golden set, NOT a blind code edit. The existing tests (test_trim_inactive_tail_drops_slow_drift, ..._keeps_continuous_rally_and_is_noop_when_off, test_window_inactivity_end_trims_through_api ? all 40 pass) lock in the current win_count semantics but have NO coverage of the sparse-tail edge case, so a naive fix could pass CI while regressing real clips.",
28 "code_location": "backend/pipeline/detectors/tracknet_runner.py:402-428 (_trim_inactive_tail); caller tracknet_runner.py:379-382; grouping tracknet_runner.py:345,359-367; production config backend/config/models.py:153-155; wiring backend/pipeline/segmenters/trajectory_hybrid.py:206-217",
29 "evidence": "tracknet_runner.py:413-428 (_trim_inactive_tail): \"frames = [f for f, _ in group]; fast = [1 if v > rally_thresh else 0 for _, v in group]; lo=0; fast_sum=0; last_dense=-1; for i in range(len(frames)): fast_sum += fast[i]; while frames[lo] <= frames[i] - trail_frames: fast_sum -= fast[lo]; lo += 1; win_count = i - lo + 1; if win_count > 0 and fast_sum / win_count >= min_density: last_dense = i; return group[: last_dense + 1] if last_dense >= 0 else group\" ? win_count is active-frame count, confirming the claim. Grouping (lines 359-367): group split when af[0]-group[-1][0] > max_gap_frames where max_gap_frames=int(merge_gap*fps) (line 345); production merge_gap=2.0 (config.json:30, models.py:122) => 60 frames @30fps, so a 1.8s/54-frame gap does NOT split. Repro through public API at production operating point: rally 0-60 fast + invisible 61-114 (54fr) + fast frames

115,116 -> trajectory_to_action_windows(..., merge_gap=2.0, window_inactivity_end=1.5, inactivity_rally_thresh=0.20, inactivity_min_density=0.30, max_window_duration=6.0) => single window start=0.03 end=3.87 (true END ~2.0s). With gap=72fr(2.4s)>merge_gap => 0 windows (split + min-duration dropped) ? confirms the merge_gap mitigation is partial. Proposed-fix simulation (temporal denom, min_density=0.30): edge case fixed, but normal fast rally (true END frame 180) at 5/6/7.5 active-fps trims to END frame 0/5/12 ? a 5.6-6.0s regression. Production wiring: backend/config/loader.py:95 merged={**defaults,**file_data} then AppConfig.model_validate re-applies IndexingConfig defaults; backend/main.py:543 global_config=load_app_config() (typed); models.py get() returns nested section via model_dump so trajectory_hybrid.py:121 idx_cfg carries the R4-ON defaults. docs/QUALITY_ITERATIONS.md:104-208 documents the measured R4 Launch Report and that inactivity_min_density=0.30 is the operating point. All 40 tests in tests/test_tracknet_runner.py pass on the current code.",
30   "bottom_line": "The code claim is ACCURATE: _trim_inactive_tail does measure density over visible/active frames (win_count = i-lo+1), so a single fast frame after a sub-merge_gap (< 2.0s) dead window yields density 1.0 and survives the trim, inflating the rally END ? I reproduced a ~1.85s END over-extension through the public API at the production (R4-on) operating point. It is a REAL but LOW-severity boundary quirk: it only fires in the narrow band where a dead-time gap is shorter than merge_gap yet a sparse fast tail re-appears, and the claim's own \"mitigated by merge_gap\" caveat is correct (larger gaps split the group away). The PROPOSED FIX is the bigger problem: swapping the density denominator to temporal length without re-tuning min_density=0.30 silently destroys normal sparse-tracking rallies (simulated 5.6-6.0s END collapse at realistic 5-7.5 active-fps), and the plan leaves the critical few-frames fallback question unresolved. Because this is the repo's #1 quality lever and min_density was calibrated against the current denominator via a documented Launch Report (PR #204), this must NOT be merged as a blind hardening patch ? it needs a measured Launch Report + dual-set golden regression, with min_density re-tuned for the new denominator. Recommend: document the quirk, add a sparse-tail regression test, and gate any denominator change behind measurement; do not ship the proposed fix as a code-only change."
31   },
32   {
33   "claim_id": "C-cvat-parser",
34   "claimed_behavior": "The PR claims that in sports-data-collector's src/parsers/cvat/rally_cvat.py: (a) parse_cvat_boxes searches a zip for an .xml ending in \"annotations.xml\" or, failing that, ANY .xml; (b) load_canonical_csv uses safe_int() to tolerate malformed rally_number fields and silently `continue`s past entirely-malformed-but-CSV-parseable rows; (c) a severely malformed CSV could silently drop rallies without failing the import gate, so observability is poor. The PLAN proposes adding a `logger.warning` before the `continue` in load_canonical_csv to surface skipped rows.",
35   "actual_behavior": "The code descriptions are accurate. parse_cvat_boxes (lines 209-220): when zipfile.is_zipfile is true it builds `names = [n for n in zf.namelist() if n.endswith(\"annotations.xml\")] or [n for n in zf.namelist() if n.endswith(\".xml\")]`, raises ValueError if none, else parses names[0] ? exactly the \"annotations.xml first, else any .xml\" logic claimed. load_canonical_csv (lines 564-602) calls `rn = safe_int(row.get(\"rally_number\"), None)` and `if rn is None: continue` (line 579-580), plus a second silent skip `if start is None or end is None: continue` (line 588-589). safe_int (src/core/parsing.py:9-19) returns the default (None here) on TypeError/ValueError via int(float(s)), so a non-numeric rally_number yields None -> the row is dropped with no log. The \"silent continue\" is REAL: there is NO logging/warning anywhere around either skip, the module does NOT import logging, and a repo-wide grep shows the ENTIRE src/ tree uses no `logging`/`getLogger` ? so the proposed fix introduces logging to a module (and effectively a subsystem) that has none. However, the claim's \"silently drop half its rallies without failing the import gate\" is only partly true at the system level: the CLI caller validate_delivery (curate.py:997-1005) hard-fails with sys.exit(1) when load_canonical_csv returns ZERO records (\"no usable rally rows\"), and emits an observability report (emit_delivery_report). So a fully-empty CSV IS gated; only a PARTIAL drop (some rows good, some malformed) passes silently with an undercount ? which is exactly the gap the claim targets, so the core observation stands.",
36   "verdict": "partially-accurate",
37   "is_real_bug": false,
38   "severity": "low",
39   "proposed_fix_assessment": "The fix (log a warning before the silent `continue`) is directionally correct and would improve observability of partial-drop CSVs, but it is under-specified for this repo. The module does NOT import logging and there is NO logger anywhere in src/ ? the codebase's actual convention for surfacing defects is the Issue/severity dataclass system plus print() at the CLI layer (e.g. validate_delivery prints to stdout and exits non-

zero). A bare `logger.warning` would (a) need a new `import logging` + `logger = logging.getLogger(__name__)`, and (b) emit to a logging channel nothing else in the project configures, so by default it would be invisible unless the CLI sets up logging. A more convention-fitting fix would be to count skipped rows and either print them in validate_delivery or surface them as a low/medium Issue (e.g. \"n rows skipped: unparseable rally_number / missing start-end\"), which plugs straight into the existing build_summary/render_validation_md/review.csv observability already wired through emit_delivery_report. The fix is trivially testable (a malformed row + assert on the warning/Issue), and existing tests would not break: test_empty_cvat_converts_to_zero_records, test_load_canonical_csv_decimal_and_mmss, and test_load_canonical_csv_keeps_rally_starting_at_zero assert on record counts/values, none use caplog or assert silence, so adding a warning is backward-compatible.",
40 "fix_risk": "Very low engineering risk: no DEFAULT/production behavior change, no rally-boundary or eval-metric impact, no measurement needed (it is pure observability ? it only adds a log line, the records returned are identical). No existing test asserts silence, so nothing regresses. The real risk is process/scope, not code: this change lives in the SIBLING repo sports-data-collector, not in badminton-highlight-indexer. Bundling it into the indexer's docs/plan violates the indexer's stated convention \"Branch off master; ONE PR per work item; NO PR stacking\" and conflates two repos ? the collector fix must land as its own PR in the collector repo. The plan is conflating two repos.",
41 "evidence": "parse_cvat_boxes zip logic ? rally_cvat.py:211-216:\n```\n with zipfile.ZipFile(xml_path) as zf:\n names = ([n for n in zf.namelist() if n.endswith(\"annotations.xml\")]\n or [n for n in zf.namelist() if n.endswith(\".xml\")])\n if not names:\n raise ValueError(f\"CVAT zip contains no annotations.xml: {xml_path}\")\n root = ET.fromstring(zf.read(names[0]))\n```\nload_canonical_csv silent skips ? rally_cvat.py:579-589:\n```\n rn = safe_int(row.get(\"rally_number\"), None)\n if rn is None:\n continue\n start = _parse_secs(row.get(\"start_time\"))\n ...\n if start is None or end is None:\n continue\n```\nsafe_int swallows errors ? src/core/parsing.py:16-19:\n```\n try:\n return int(float(s))\n except (TypeError, ValueError):\n return default\n```\nNo logger in module or repo: Grep for `^import logging|logging\\.|getLogger` in rally_cvat.py => \"No matches found\"; Grep for `getLogger|import logging` across src/ => \"No files found\". The module's import block (lines 38-48) has math/os/re/zipfile/ET/collections/dataclasses/typing + safe_int + probe_resolution_fps ? no logging.\nCLI gate that DOES catch a fully-empty CSV ? curate.py:1003-1005:\n```\n if not records:\n print(f\"Error: no usable rally rows in {args.csv}\")\n sys.exit(1)\n```\nTests pin silent-skip behavior WITHOUT asserting silence (so a warning is safe): test_empty_cvat_converts_to_zero_records (test_cvat_robustness.py:33-37 asserts `result.records == []`), test_load_canonical_csv_decimal_and_mmss and test_load_canonical_csv_keeps_rally_starting_at_zero (test_cvat_parser.py:339-363 assert record counts/values) ? none use caplog.",
42 "bottom_line": "The factual description is accurate: parse_cvat_boxes does pick \"annotations.xml\" then any \".xml\", and load_canonical_csv does silently `continue` past malformed rows (two skips, lines 580 and 589) with no logging ? confirmed real, and the module/repo has no logger at all. This is a genuine but minor observability gap (severity low), not a correctness bug: a fully-empty CSV is already hard-gated by the CLI (sys.exit(1) + emit_delivery_report), so only PARTIAL undercounts slip through silently. The proposed `logger.warning` is safe (no test asserts silence, no behavior/eval impact) but under-specified for this repo's conventions ? it should add `import logging`+logger or, better, fit the existing Issue/summary observability rather than a logging channel nothing configures. The bigger problem is process: this is a sports-data-collector change bundled into the indexer's plan, which violates the indexer's ONE-PR-per-work-item rule and conflates two repos ? the collector fix belongs in its own collector PR."
43 },
44 {
45 "claim_id": "D-splitter-praise",
46 "claimed_behavior": "The PR praises the W1/R5 blob-splitters: they use a \"recursive worklist\" doing `split_i = TrackNetRunner._midpoint_split_index(h); work.append(h[:split_i]); work.append(h[split_i:])`. It claims the code \"explicitly enforces len(h) < 2 before splitting,\" so split_i is \"always at least 1,\" which \"safely guarantees it will never enter an infinite loop of empty arrays\" ? calling it excellent, defensively written code.",
47 "actual_behavior": "The core safety reasoning is CORRECT, with cosmetic

imprecisions. `_midpoint_split_index` (backend/pipeline/detectors/tracknet_runner.py:431-436) computes `split_i = min(range(1, len(group)), key=lambda i: abs(group[i][0] - mid))`, so for any group with len >= 2, split_i is drawn from {1, ..., len-1}, i.e. 1 <= split_i <= len-1. Both halves h[:split_i] and h[split_i:] are therefore non-empty AND strictly smaller than the input, so each split reduces size and the worklist terminates ? no empty arrays, no infinite loop. The guard the claim references is real: R5's loop (lines 506-513) only reaches `_midpoint_split_index` when the `if len(h) < 2 or (h[-1][0]-h[0][0]) <= max_win_frames: continue` test (line 508) fails, i.e. only when len(h) >= 2. W1's `_split_overlong_groups` has the equivalent guard at line 455 and calls the splitter at lines 467-469. Three imprecisions in the praise: (1) the code is an ITERATIVE worklist (`while work: ... work.append(...)`), explicitly documented \"Iterative worklist to bound stack depth\" on lines 448 and 492 ? NOT \"recursive\" as the claim says; (2) the quoted snippet using variable `h` is the R5 path specifically (lines 511-513); the W1 path uses variable `g` (lines 467-469), so the snippet does not literally appear in both as implied; (3) the phrase \"explicitly enforces len(h) < 2 before splitting\" is backwards-worded ? the code SKIPS (continues) when len(h) < 2 and requires len(h) >= 2 to split; there is no literal standalone \"enforce len(h) < 2\" statement, it is one disjunct of a skip-guard.",
48   "verdict": "accurate",
49   "is_real_bug": false,
50   "severity": "cosmetic",
51   "proposed_fix_assessment": "No fix proposed or needed ? this is praise of existing, correct code. The code genuinely cannot infinite-loop or produce empty arrays: for len(h) >= 2, _midpoint_split_index returns split_i in [1, len-1], yielding two non-empty, strictly-smaller halves, so the worklist provably terminates. The praise's CONCLUSION is sound. Its description is loose on three points (iterative not recursive; the h-variable snippet is R5-only; \"enforces len < 2\" is a mis-phrasing of a skip-guard), but none of these undermine the verified guarantee.",
52   "fix_risk": "None. No code change is implied by this claim.",
53   "evidence": "_midpoint_split_index (line 431-436):\n mid = (group[0][0] + group[-1][0]) / 2.0\n split_i = min(range(1, len(group)), key=lambda i: abs(group[i][0] - mid))\n return split_i\n=> range(1, len(group)) is non-empty iff len(group) >= 2; split_i in {1..len-1}.\n\nR5 worklist (_blob_fallback_chop, lines 506-513):\n while work:\n   h = work.pop()\n   if len(h) < 2 or (h[-1][0] - h[0][0]) <= max_win_frames:\n pieces.append(h)\n   continue\n   split_i = TrackNetRunner._midpoint_split_index(h)\n work.append(h[:split_i])\n   work.append(h[split_i:])\n=> _midpoint_split_index only called when len(h) >= 2, so split_i in [1, len(h)-1]; both halves non-empty and strictly smaller. Terminates.\n\nW1 worklist (_split_overlong_groups, lines 455 + 466-469) ? note variable is `g`, not `h`, and the midpoint chop only runs under hard_cap:\n if len(g) < 2 or (g[-1][0] - g[0][0]) <= max_win_frames: out.append(g); continue\n ...\n if hard_cap:\n   split_i = TrackNetRunner._midpoint_split_index(g)\n   work.append(g[:split_i]); work.append(g[split_i:])\n\nDocstrings literally say \"Iterative worklist to bound stack depth\" (lines 448, 492) ? contradicting the claim's word \"recursive\".",
54   "bottom_line": "The praise is substantively ACCURATE: the safety guarantee it describes is real and provable. Because the splitter only calls _midpoint_split_index when len(h) >= 2, split_i is always in [1, len-1], so both halves are non-empty and strictly smaller ? the worklist provably terminates with no empty arrays and no infinite loop. Three cosmetic imprecisions do not change the correctness: (1) the worklist is ITERATIVE, not \"recursive\" (the docstrings say so explicitly); (2) the quoted `h`-variable snippet is the R5 path ? W1 uses `g` and only chops at midpoint under hard_cap; (3) \"explicitly enforces len(h) < 2 before splitting\" is a backwards description of a skip-when-len<2 guard. The reviewer correctly identified the real guard and the real reason it is safe, so this is not a tell of confabulation ? but the loose \"recursive\"/\"enforces len(h)<2\" wording shows the description was paraphrased rather than read line-for-line. Verdict: praise correct (severity cosmetic for the wording slips only)."
55   }
56   ],
57   "narrative": {
58   "section": "How Rally Detection Works (narrative + config facts)",
59   "overall_accuracy": "mostly-accurate",
60   "inaccuracies": [
61   "BLOCKER (process, not content): The file 'RALLY_DETECTION_CODE_REVIEW.md' does NOT exist anywhere in this repo ? not in the working tree, not on any branch (checked all 40+), not in any worktree, not in the stash, not as an untracked file. I verified the narrative's CLAIMS (as enumerated in the task) against the code; I could not read the review prose itself, so any wording-level nuance in Section 1 is unverifiable. Findings below grade the asserted

claims against code.",
62 "The narrative frames W1 / max_window_duration as the gating mechanism but (per the task summary) treats split-at-largest-gap as the behavior; it is accurate that this exists, but the narrative should note max_window_duration is DEFAULT-ON at 6.0 (R1 milestone), not the historical 0=OFF. The Python function-signature default is 0.0; the shipped config default (backend/config/models.py:139 IndexingConfig.max_window_duration) is 6.0. A reader taking '0=OFF' as live would be wrong.",
63 "R4 inactivity-tail trim is described as active; correct ? it is DEFAULT-ON in config (window_inactivity_end=1.5, inactivity_rally_thresh=0.20, inactivity_min_density=0.30 at models.py:153-155). But the trajectory_to_action_windows function-signature defaults differ (window_inactivity_end=0.0 OFF, inactivity_rally_thresh=0.08, inactivity_min_density=0.34). If the review quoted the function-signature numbers it would be wrong about live behavior; if it quoted config numbers it is right. The two sources genuinely disagree ? flag whichever the review used.",
64 "Gate A 'min_crossings ~2': the ~2 value is ONLY the rally_gate.py function-signature default (gate_windows/apply_gate, lines 92/114). The SERVED/config default is min_crossings=0 = OFF (FusionConfig, models.py:105), and fusion_hybrid is not even the default segmenter (default_segmenter='yolo_hybrid'). So in the live default pipeline Gate A does NOT run. Asserting '~2' as the operative default is misleading.",
65 "Narrative says the handoff goes 'to a Gemini provider' ? true by default (ai_provider='gemini', GeminiProvider registered), but the provider is pluggable and there is also a fully-local skip_ai_handoff path (default OFF) that emits candidate windows directly as rallies with NO provider. A flat 'to a Gemini provider' omits that the handoff is provider-agnostic and optional.",
66 "docs/HOW_RALLY_DETECTION_WORKS.md does NOT document W1/hard-cap/R4/R5/Gate A as the pipeline. Its documented mental model is only 3 knobs (velocity_thresh, merge_gap, min_window_duration) plus a 'where sparks could land' wishlist (Gate A/fusion appear there as a parked open question, not as live structure). So the review's multi-stage W1->hard-cap->R4->R5->GateA->padding->Gemini pipeline is REAL IN CODE (tracknet_runner.py + fusion_hybrid.py) but is NOT the model this particular doc describes ? the review draws structure from the code/RALLY_QUALITY_RESEARCH.md/QUALITY_ITERATIONS.md, not from HOW_RALLY_DETECTION_WORKS.md. If the review claims its pipeline matches that doc, that is wrong; if it claims it matches the code, that is right."
67],
68 "config_fact_checks": [
69 "velocity_thresh ~0.02: CLAIMED ~0.02 / ACTUAL 0.02 ? CORRECT. WasbIndexConfig.velocity_threshold=0.02 (models.py:69), TrackNetConfig.velocity_threshold=0.02 (models.py:43); read at trajectory_hybrid.py:160 (idx_cfg[family].velocity_threshold, fallback 0.02).",
70 "merge_gap ~2.0s: CLAIMED ~2.0 / ACTUAL 2.0 ? CORRECT. IndexingConfig.dead_time_merge_gap=2.0 (models.py:122); read at trajectory_hybrid.py:161.",
71 "min_window_duration (the 3rd knob, not in task list but part of the doc): ACTUAL 2.0 ? IndexingConfig.min_action_window_duration=2.0 (models.py:131); read at trajectory_hybrid.py:162. Windows shorter than this are dropped (tracknet_runner.py:399).",
72 "max_window_duration: CLAIMED 'W1 split over max_window_duration' / ACTUAL config default 6.0 DEFAULT-ON (models.py:139), function-sig default 0.0 OFF (tracknet_runner.py:290). The split-at-largest-internal-gap>=min_split_gap mechanism is real (_split_overlong_groups, tracknet_runner.py:439-476). Recall-safe (only cuts).",
73 "min_split_gap ~0.5: CLAIMED ~0.5 / ACTUAL 0.5 ? CORRECT. IndexingConfig.window_min_split_gap=0.5 (models.py:140); function-sig default also 0.5 (tracknet_runner.py:291). Used as int(min_split_gap*fps) floor for a gap to qualify as a split point.",
74 "window_hard_cap chop-in-half: CLAIMED chop-in-half / ACTUAL CORRECT mechanism (midpoint split when no internal gap >= floor, _split_overlong_groups hard_cap branch + _midpoint_split_index, tracknet_runner.py:466-469,431-436), but default is False=OFF (models.py:145). It chops BEFORE the R4 trim.",
75 "rally_thresh/min_density (R4): ACTUAL config inactivity_rally_thresh=0.20, inactivity_min_density=0.30, window_inactivity_end=1.5 ? all DEFAULT-ON (models.py:153-155). FUNCTION-SIG defaults differ: rally_thresh=0.08, min_density=0.34, inactivity_end=0.0 (tracknet_runner.py:293-295). Mechanism (_trim_inactive_tail, two-pointer trailing-density trim, only cuts) is faithfully described.",
76 "padding ~2s each end: CLAIMED ~2s / ACTUAL 2.0 ? CORRECT. IndexingConfig.ai_handoff_padding=2.0 (models.py:172); applied symmetrically in Phase 3 (trajectory_hybrid.py:278-279: start-padding, end+padding, clamped to [0,duration]).",
77 "min_crossings ~2 (Gate A): CLAIMED ~2 / ACTUAL split-brain.

rally_gate.gate_windows/apply_gate signature default = 2 (rally_gate.py:92,114), but SERVED config FusionConfig.min_crossings=0=OFF (models.py:105). Also fusion_hybrid is opt-in, not default_segmenter. The '~2' is only the bare-function default, not the live pipeline default. Mechanism (net = median shuttle coord on the larger-spread image axis; count sign changes of coord-net with deadband) is CORRECT (rally_gate.py:55-110).",
78 "R5 blob fallback + forced_cut flag: ACTUAL CORRECT. window_blob_fallback default False=OFF (models.py:164), window_blob_factor=4.0 (models.py:169). _blob_fallback_chop midpoint-chops only groups > blob_factor*cap, sets w['forced_cut']=True per piece (tracknet_runner.py:388-398,478-520). The 'forced_cut' flag name is exact and real.",
79 "Provider/segmenter wiring: ai_provider='gemini' (models.py:289, AppConfig), GeminiProvider registered (providers/__init__.py:9). BUT default_segmenter='yolo_hybrid' (models.py:118) ? NOT a trajectory hybrid ? so the WASB->velocity->...->Gemini pipeline the review describes is the wasb_hybrid/fusion_hybrid path, which is NOT the out-of-box default segmenter."
80],
81 "bottom_line": "Could not read RALLY_DETECTION_CODE_REVIEW.md ? it does not exist in this repo (no branch/worktree/stash/untracked copy), so I verified the narrative's enumerated CLAIMS against code, not the prose itself. On those claims: the pipeline shape (WASB shuttle -> per-frame velocity -> merge within ~2s -> W1 split at largest gap>=0.5s -> optional hard-cap midpoint chop -> R4 inactive-tail trim -> R5 blob fallback with forced_cut flag -> Gate A net-crossing filter on median play-axis -> ~2s symmetric padding -> Gemini handoff) is REAL and faithfully mechanized in tracknet_runner.py (shared brain, rebound onto WasbRunner) + fusion_hybrid.py + rally_gate.py. The flag/label names (forced_cut, min_crossings, window_hard_cap, etc.) all match. Two recurring accuracy traps: (1) several 'defaults' the review cites are FUNCTION-SIGNATURE defaults that DISAGREE with the shipped config defaults ? most importantly max_window_duration (sig 0/OFF vs config 6.0/ON), R4 thresholds (0.08/0.34/0.0 vs 0.20/0.30/1.5), and min_crossings (sig 2 vs config 0/OFF); the live pipeline uses the config values, and Gate A is OFF by default. (2) docs/HOW_RALLY_DETECTION_WORKS.md only documents the 3-knob model and lists Gate A/fusion as a NOT-yet-built 'open question', so the review's multi-stage pipeline matches the CODE but not that doc's mental model ? a claim that it mirrors that doc would be wrong. Net: mechanism descriptions accurate; specific default NUMBERS are the soft spot ? verify each against config (models.py) rather than the function signatures."
82 },
83 "strategy": {
84 "launch_report_applicability": "Fix B needs a Launch Report, not a blind code edit. _trim_inactive_tail is in the live production segmentation path (backend/pipeline/detectors/tracknet_runner.py line 381, called during active grouping with a real default inactivity_min_density of 0.34). Re-basing its density from visible-frame win_count to temporal length changes WHERE the tail is trimmed, i.e. it shifts rally-END boundaries on real footage. Per the owner Launch Report convention (every default or quality-bit flip ships a measured total-quality-impact table on both rulers plus an over-seg guard before merge), this cannot land as an unmeasured harden-edge-case diff. Fix A (_spread) is observability/latent only: the code review itself states it does not crash the pipeline and only bites if reused on small window slices in the future, so no current production effect and no Launch Report needed. Fix C (cvat parser logging) is pure observability (a logger.warning before an existing continue), no behavior change, no Launch Report, and it targets the sibling repo, not this one.",
85 "decision_rigor_alignment": "Misaligned with NUMBERS INFORM, FOUNDATIONS DECIDE and the dual-set no-regression rule. The plan asserts hardening (tighten boundary calculation edge cases) with zero measurement commitment: no delta-F1, no bootstrap CI or paired sign test, no RAW-set and growing-golden-set no-regression check, no run_regression.py gate. Its DoD is only density calculation corrected plus tests updated and passing. Tests passing is NOT sufficient for Fix B, which is a boundary-behavior change: a green suite proves the new math is internally consistent, not that rally-END quality did not regress on either eval set. By the decision-rigor norm this change owes a per-signal marginal delta-F1 with honest CI and a foundational lens; neither is present. Fix A and Fix C, being non-behavioral, are fine on tests only.",
86 "roadmap_fit": "Mostly a distraction, with one accidental collision. Per NEXT_STEPS.md the owner top levers are: grow the golden corpus (number one), set the ship-gate target, the rally-END R4 end-rule / player-state end-rule (the number-one measured remaining gap, RALLY_DETECTION_GAP_CLOSURE.md G3), the prominent-court multicourt FP gate (PROMINENT_COURT_DETECTION.md, just landed in PR 243), multi-signal fusion (Gate A wiring, person, audio), and the Roora golden set. Fix A (latent _spread) and Fix C (collector-side parser logging) are orthogonal micro-cleanups with no roadmap pull. Fix B is the trap: by reworking _trim_inactive_tail density/trim math it directly touches the R4 inactive-tail

trimmer, i.e. the rally-END over-extension gap G3 that is the top priority. It should NOT proceed as a standalone harden-edge-case; it belongs folded into the tracked RALLY_DETECTION_GAP_CLOSURE.md effort (owner-gated, flag-gated, measured on the golden corpus), where its boundary effect is evaluated against the dual eval sets rather than blind-patched.",
87   "funding_framing": "The merge-and-fund-the-project framing is badly oversized and should be right-sized before any decision. The PR text contains NO funding, resourcing, headcount, GPU, or budget ask whatsoever; the entire body is please review the proposed plan and code review, and once approved I will implement the code changes. The actual content is a roughly 3-spot localized cleanup: one latent guard in _spread (a few lines), one density-math tweak in _trim_inactive_tail, and one logger.warning in a sibling-repo parser, touching two functions in this repo plus a one-line log in another repo. This is normal in-session tech-debt and hardening work, not a project to be funded. Recommend the owner drop the funding lens entirely and treat it as: keep the two genuinely-safe trivial fixes, and route the one behavior-touching piece into the existing rally-END track.",
88   "merge_recommendation": "merge-with-changes",
89   "bottom_line": "Right-size this: it is a roughly 3-spot localized cleanup, not a fundable project, and there is no funding ask in the PR to act on. Split the three fixes by behavior. Fix A (_spread latent guard) and Fix C (cvat parser logger.warning) are safe observability/edge-only changes, fine on tests only, though Fix C must land in the sibling sports-data-collector repo, not here, so it cannot ride this PR. Fix B is NOT a harden-edge-case: _trim_inactive_tail is the live R4 inactive-tail trimmer, so changing its density math moves rally-END boundaries, exactly the top tracked rally-END gap (RALLY_DETECTION_GAP_CLOSURE.md G3). It therefore owes a measured Launch Report (both rulers plus over-seg guard) and dual-set no-regression / delta-F1 per NUMBERS INFORM FOUNDATIONS DECIDE, and should be folded into the owner-gated rally-END effort rather than blind-patched here. Doc fixes before merge: add the missing decisions-locked section plus a success-metrics/measurement section, correct the mis-stated Extends RALLY_DETECTION_QUALITY_REPORT.md relation (point at the gap-closure tracker instead), and fix the mis-homed src/parsers/cvat anchor. The QUALITY_REPORT cross-reference itself is valid since that file exists, so no fabrication, but the Extends claim is wrong.",
90   "template_compliance": "Partially compliant. The plan uses PROJECT_DOC_TEMPLATE.md and includes the status header, TL;DR, problem, design, honest-limits, progress tracker, open questions, DoD, and references. It SKIPS section 2 (Decisions locked) and the optional sections 3 (Foundation) and 5 (Threat model). Sections 3 and 5 are legitimately optional and can be dropped. The real gap is the non-optional section 2 (no decisions-locked table) plus the absence of any success-metrics or measurement section and any alternatives-considered, which for a behavior-touching change are exactly what the repo tracked-doc discipline exists to capture. The cross-reference is NOT fabricated: it cites Extends RALLY_DETECTION_QUALITY_REPORT.md and that file does exist (docs/RALLY_DETECTION_QUALITY_REPORT.md, tracked, indexed in docs/README.md). However the Extends relation is mis-stated: this cleanup neither extends nor relates to that paper-checkpoint doc; the genuinely relevant trackers (RALLY_DETECTION_GAP_CLOSURE.md, PROMINENT_COURT_DETECTION.md) are not referenced. Separately, the Fix C code anchor src/parsers/cvat/rally_cvat.py does NOT exist in this repo; it lives in the sibling sports-data-collector repo, so it cannot be implemented under this PR and the anchor is mis-homed."
91   }
92   }
93   }

29. assistant (2026-06-19T16:51:54.446Z)

I have the complete, cross-verified picture now. Every claim was reproduced against real code by independent adversarial agents and matches my own reads. Let me post the review on the PR, then give you my recommendation.

30. assistant (2026-06-19T16:52:13.525Z)

Review ? PR #245: "Proposal for Rally Detection Fixes"

Verdict: don't merge this as the basis for blind implementation, and there's nothing here to "fund." This is a docs-only proposal for ~3 small code changes (two functions in this repo + one log line in the sibling repo). Useful diligence ? but the three items split cleanly into **drop one / re-home one / gate one behind measurement**. Every claim below was reproduced against the actual code, not taken from the write-up.

Good news first: the review clearly *read* the code. The W1/R5 splitter termination-safety praise (§D) is **correct** ? `_midpoint_split_index` returns an index in `[1, len-1]`, so both halves are non-empty and strictly smaller and the worklist provably terminates (`tracknet_runner.py:431`, `:506-513`). The pipeline narrative (§1) is also mostly faithful to the code.

Per-claim assessment

Claim	Verdict	Real bug?	Action
A ? `_spread` zero-spread / "divide-by-zero"	Misleading	? No	**Drop** (at most a one-line comment)
B ? `_trim_inactive_tail` sparse-tail density	Accurate (real, low-sev)	? Minor	**Don't blind-fix** ? route into the rally-END track with a Launch Report
C ? `rally_cvat.py` silent row-skip	Partially accurate	? (observability gap)	Valid, but it's in the **sibling repo** ? land it there
D ? W1/R5 splitter praise	Accurate	n/a	Praise is well-founded

A ? `_spread` (`rally_gate.py:37`). The arithmetic is right (`len==2 ? 0.0`), the impact framing is not. `_spread` is only ever called by `estimate_play_axis` over the *entire* video's visible trajectory (thousands of points), never on small slices (`fusion_hybrid.py:106` caches it per-trajectory). And there is **no division by spread anywhere** ? its only consumer is `deadband = deadband_frac * span` (`rally_gate.py:102`), a multiplication; a zero span just disables the jitter deadband (`abs(v) < 0` is always False), which degrades gracefully. So the plan's "potential divide-by-zero" is unfounded. The function already returns `0.0` for `len < 2` *by design* (line 39), so "make `len==2` non-zero" contradicts the existing guard ? and switching `int()?ceil()` would silently shift the p5/p95 indices, `span`, the deadband, and the kept/dropped-rally decisions on *every* video, with zero measured benefit. Cosmetic non-bug; recommend dropping it. (Aside: the `_spread` that actually feeds the feature schema is a *different* function ? `rally_seq.py:43`, non-truncating `np.percentile` ? so even the "future reuse" worry is already covered there.)

B ? `_trim_inactive_tail` (`tracknet_runner.py:402`). Real and well-spotted: `win_count = i - lo + 1` counts *active* frames in the trailing window, not elapsed time, so a lone fast frame after a dead-time gap shorter than `merge_gap` (2.0s) gives density `1/1 = 1.0` and survives the trim. Reproduced through the public API at the production operating point: a rally that truly ends ~2.0s over-extended to **3.87s**. **But this is not a "harden an edge case" change** ? `_trim_inactive_tail` *is* R4, the **default-ON, measured, tuned rally-END trimmer** (`window_inactivity_end=1.5`, `inactivity_min_density=0.30`; flipped on as a measured milestone ? ?tolF1 +0.040, sign 9/1/3, p=0.022, |?end| 4.96?3.31s, n=13; `models.py:150-155`). It is also precisely the project's **#1 remaining quality gap (rally-END detection)**. The proposed denominator swap with `min_density=0.30` left untouched silently *re-defines* that calibrated threshold: simulated against a normal sparse-tracking rally (5?7.5 active-fps, entirely realistic), it collapses a real 6s rally end by **~6 seconds**. The plan's own open-question #1 (few-frames fallback?) is exactly the unresolved part. ? Fold this into the owner-gated rally-END effort (`RALLY_DETECTION_GAP_CLOSURE.md`), flag-gated, with a **Launch Report (both eval rulers + over-seg guard) + dual-set no-regression**, and `min_density` re-tuned for the new denominator. Cheap, safe interim step: add a sparse-tail regression test that documents the current quirk. "Tests updated and passing" is *not* a sufficient DoD for a boundary-behavior change.

C ? `rally_cvat.py` (sibling repo). Accurate: `load_canonical_csv` does `continue` past malformed rows with no logging, and the `src/` tree has no logger at all. A fully-empty CSV is already hard-gated by the CLI (`sys.exit(1)` + `emit_delivery_report`), so only *partial* undercounts slip through silently ? a genuine but minor observability gap. The `logger.warning` is safe (no test asserts silence). **Two caveats:** (1) it lives in **sports-data-collector**, not this repo, so per our one-PR-per-item rule it can't ride this PR ? land it there; (2) prefer fitting the existing Issue/summary observability over introducing a lone logging channel nothing configures.

Doc / process nits

- The plan skips the non-optional **"Decisions locked"** section and has **no success-metrics / measurement section** ? exactly what a behavior-touching change (B) needs.
- "Extends `RALLY_DETECTION_QUALITY_REPORT.md`" is mis-stated ? that's the paper-checkpoint doc; the relevant tracker is the rally-END / gap-closure work.

- The `src/parsers/cvat/rally_cvat.py` anchor is mis-homed (sibling repo).

Recommendation

Request changes. Don't adopt this plan for blind implementation, and drop the "fund a project" framing ? there is no resourcing ask in the PR; this is a localized ~3-spot cleanup. Concretely: **drop A**, **re-home C** to the collector repo, and **promote B** into the measured rally-END track instead of implementing it as written. The contributor's code-reading is genuine and welcome; the gap is calibration (a non-bug escalated to "divide-by-zero," and a measured, default-ON lever treated as a harmless hardening).

31. user (2026-06-19T16:52:13.573Z)

File created successfully at: C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\pr245_review.md (file state is current in your context ? no need to Read it back)

32. user (2026-06-19T16:52:21.328Z)

https://github.com/avidullu/badminton-highlight-indexer/pull/245#issuecomment-4753405903

33. assistant (2026-06-19T16:52:44.151Z)

Posted: [PR #245 review comment](https://github.com/avidullu/badminton-highlight-indexer/pull/245#issuecomment-4753405903).

My recommendation: **don't merge, and don't "fund" anything**

There's no project to fund here. The PR body's own words are *"once approved, I will implement the code changes"* ? no resourcing, budget, or headcount ask anywhere. What's actually proposed is a **~3-spot localized cleanup**: two functions in this repo and one log line in the sibling repo. The "merge and fund the project" framing is badly oversized relative to the content. I'd drop that lens entirely and treat it as ordinary review.

I reproduced every claim against the real code (not the write-up) via parallel adversarial verification. The three proposed fixes split cleanly:

	Claim	Reality	What to do
A	`_spread` "divide-by-zero" / zero-spread	**Non-bug.** `span` is only ever *multiplied* (`deadband = 0.06 * span`, [rally_gate.py:102](backend/pipeline/detectors/rally_gate.py:102)) ? there is no division anywhere. It's only called over the whole video's trajectory, never small slices, and already returns `0.0` for `len<2` *by design*. The `int()?ceil()` "fix" would silently shift the play-axis span/deadband on **every** video for zero benefit.	**Drop it**
B	`_trim_inactive_tail` sparse-tail density	**Real but low-severity** edge case (reproduced: a ~2.0s rally over-extends to 3.87s). But this *is* R4 ? the **default-ON, measured, tuned** rally-END trimmer ([models.py:150-155](backend/config/models.py:150): ?tolF1 +0.040, p=0.022, n=13) ? and rally-END is the project's **#1 quality gap**. The proposed denominator swap, with `min_density` untouched, **regresses normal rallies by ~6s** in simulation.	**Don't blind-patch.** Route into the gap-closure track with a Launch Report + dual-set regression
C	`rally_cvat.py` silent row-skip	**Accurate, genuinely useful** observability nit (a `logger.warning`).	Valid ? but it's in **sports-data-collector**, so it can't ride this PR. Land it there
D	W1/R5 splitter termination praise	**Correct** ? the `len<2` guard + `_midpoint_split_index ? [1,len-1]` genuinely guarantee termination.	(no action)

The one thing that needs your eye: fix B isn't junk ? the contributor correctly found a real quirk, and it happens to touch *exactly* your top lever (the R4 end-rule). That's the trap: precisely because R4 is default-ON and measured, this can't be a "harden the edge case, tests pass" code change ? it would invalidate the tuned `min_density=0.30` calibration. It belongs in the owner-gated rally-END effort with measurement, not this PR.

Net: the contributor's code-reading is real and welcome (D is right, the pipeline narrative is

mostly accurate, B is a legitimate catch). The problem is **calibration** ? a non-bug escalated to "divide-by-zero," and a measured production lever treated as a harmless tweak.

Bottom line: request changes ? drop A, re-home C to the collector repo, promote B into the measured rally-END track. I posted exactly this on the PR.

If you'd like, I can: (a) open the C fix as its own PR in `sports-data-collector`, and/or (b) draft a proper measured plan for the B / sparse-tail end-rule under `RALLY_DETECTION_GAP_CLOSURE.md`. Want me to do either?

Also worth a quick cleanup regardless of this PR: the in-code comment at [tracknet_runner.py:374](backend/pipeline/detectors/tracknet_runner.py:374) still says R4 is "default OFF," but config has it ON at 1.5 ? stale and misleading.

34. user (2026-06-19T16:52:47.500Z)

Feel free to post comments on the PR and also push back if needed and appreciate when deserved

35. assistant (2026-06-19T16:53:40.100Z)

Good ? I'll engage at the line level: targeted inline comments that push back where claims are wrong and genuinely appreciate the good catches. First let me pull the exact PR files so I anchor each comment to the right line.

36. user (2026-06-19T16:53:45.546Z)

```
From https://github.com/avidullu/badminton-highlight-indexer
 * branch            refs/pull/245/head -> FETCH_HEAD
=== who am I (for review-event rules) ===
avidullu
=== dump PR files for line numbers ===
written
=== head sha ===
1e221d914ee5c65c6ab31f841853c256523c5ec5
```

37. assistant (2026-06-19T16:53:48.450Z)

I'm authenticated as `avidullu` ? the PR author ? so GitHub only allows me a **COMMENT** review (not REQUEST_CHANGES) with inline comments. Let me read both files for exact line numbers.

38. user (2026-06-19T16:53:49.407Z)

1 # Rally Detection Code Review
2
3 ## 1. How Rally Detection Works
4
5 The project identifies rallies through a combination of local trajectory tracking (e.g., TrackNet or WASB) and a series of heuristic gating/trimming rules, culminating in an AI-based refinement (Phase 3). Here's a breakdown of the pipeline:
6
7 ### Phase 2: Candidate Windowing (The Heuristics)
8 1. **Velocity Thresholding:** The trajectory detector outputs `(Frame, Visibility, X, Y)`. For each visible frame, the normalized velocity (distance divided by frame width per second) is calculated. Frames with a velocity exceeding `velocity_thresh` (e.g., 0.02) are considered "active" (meaning the shuttle is in flight).
9 2. **Merging / Grouping:** Active frames within a certain dead time (`merge_gap`, usually 2.0s) are merged into candidate windows.
10 3. **Splitting Over-merged Windows (W1):** If a window exceeds `max_window_duration`, the code splits it at the largest internal gap (dead-time) that is at least `min_split_gap`. If the window lacks gaps (the shuttle never stops moving), a `window_hard_cap` setting (if enabled) forcibly chops it in half at the temporal midpoint.
11 4. **Inactive Tail Trimming (R4):** Even after a rally finishes, a player might pick up the shuttle and toss it, creating a slow-moving trajectory tail. To prevent the window from overextending, the algorithm trims the end back to the last segment of the window that had a

dense cluster of "fast" frames (velocity > `rally_thresh`).

5. **Blob Fallback (R5):** If a window still vastly exceeds the cap after the W1 split and R4 trim (meaning it's an unusually continuous blob of motion), it is chopped down to the maximum cap limit and marked with a `forced_cut` flag. This signals that the clip needs "rally-state" cues (like net-crossings) rather than a simple duration cap.

6. **Rally Gate (Gate A):** In `rally_gate.py`, a heuristic filters out "false" rallies (like a player tossing a shuttle back to their opponent). It estimates the net position (median play axis coordinate) and counts how many times the shuttle crosses the net. A valid rally must have `min_crossings` (e.g., 2).

Phase 3: AI Handoff

Valid candidate windows are padded (e.g., by 2 seconds on each end) and sent to a Generative AI Provider (like Gemini) along with a sports-specific prompt. The AI confirms the semantic boundaries, ensuring the window correctly matches a true rally, and also attempts to count the shuttle exchanges.

---

2. Deep Code Review & Potential Issues Found

I performed a deep review of `tracknet_runner.py`, `trajectory_hybrid.py`, `rally_gate.py`, and the parsing logic in `sports-data-collector`. Here are the findings and hidden issues:

A. Subtle Calculation Artifact in `_spread` (`rally_gate.py`)

Observation: The `_spread` function calculates a robust spread (ignoring the bottom 5% and top 5% of coordinates to eliminate outliers):

`lo = s[max(0, int(0.05 * (len(s) - 1)))]`

`hi = s[min(len(s) - 1, int(0.95 * (len(s) - 1)))]`

Hidden Behavior: Because `int()` truncates rather than rounds, for very short arrays (e.g., `< 20` items), the 95th percentile will simply be the second-to-last item, and for `len(s) == 2`, the spread evaluates to strictly `0.0`.

Impact: This doesn't crash the pipeline, because the play axis is correctly estimated over the *entire* video's thousands of trajectory points. However, if this utility function is reused on small window slices in the future, it will quietly collapse the spread to zero.

B. Sensitivity to Sparse Tracking in `_trim_inactive_tail` (`tracknet_runner.py`)

Observation: The inactivity trimmer identifies the end of a rally by looking for the last frame where the trailing `trail_frames` window has a sufficient density of fast-moving frames (`fast_sum / win_count >= min_density`).

Hidden Issue: `win_count` represents the number of *visible, active frames* in that time boundary, rather than the total elapsed time. If the detector completely loses track of the shuttle for 1.8 seconds and only picks up 1 fast frame right at the end of the window, `win_count` will be very low (e.g., 1 or 2). Consequently, `fast_sum / win_count` could easily evaluate to 1.0 (100% density).

Impact: Sparse tracking points at the very end of a clip might prevent the tail trimmer from cutting off dead-time correctly, artificially inflating the end of the candidate window. This is somewhat mitigated by the `merge_gap` breaking trajectories, but it remains a boundary logic quirk.

C. Graceful Degradation on Incomplete CSV Validation (`rally_cvat.py` in `sports-data-collector`)

Observation: The `parse_cvat_boxes` function searches zip files for an `.xml` ending in `annotations.xml` or, failing that, *any* `.xml`.

Issue: In `load_canonical_csv()`, `safe_int()` is used to catch malformed `rally_number` fields. If a row is entirely malformed (but doesn't break the CSV parsing), it just `continue`s silently. While this provides stability, a severely malformed CSV might silently drop half its rallies without failing the import gate. Logging a warning for skipped rows would improve observability.

D. Safe Slicing in the W1/R5 Blob Splitters
The W1/R5 gap-splitters use a recursive worklist:
`split_i = TrackNetRunner._midpoint_split_index(h)`
`work.append(h[:split_i]); work.append(h[split_i:])`

Because it explicitly enforces `len(h) < 2` before splitting, `split_i` will always be at least 1, meaning the code safely guarantees it will never enter an infinite loop of empty arrays. This is an excellent, defensively written chunk of code.

39. user (2026-06-19T16:53:49.741Z)

<!--
PROJECT DOC TEMPLATE ? copy this file to docs/<YOUR_PLAN>.md and fill it in.
-->

Refactor Rally Detection Edge Cases & Parser Stability

> **Status:** `IN PROGRESS` · **Owner:** `Agent` · **Created:** `2026-06-19` · **Last updated:** `2026-06-19`
> **Lifecycle:** `DRAFT ? IN PROGRESS ? DONE ? archived`
> **Tracking anchors:** §7 progress tracker is the source of truth; indexed in `docs/README.md`.
> **Relation to existing docs:** Extends `RALLY_DETECTION_QUALITY_REPORT.md`.

---

0. TL;DR
The rally detection pipeline is largely robust, but a deep code review highlighted three minor edge cases/bugs: a potential divide-by-zero or zero-spread in `_spread` for tiny sequences, a density calculation vulnerability for sparse tracks in `_trim_inactive_tail`, and silent skipping of malformed rows in the CVAT parsing script. This plan outlines the fixes for these issues to increase stability.

1. Problem & goal
The current implementation of the candidate windowing handles normal inputs extremely well (including handling blobs), but could fail or behave erratically under specific conditions:
- **Zero-spread calculation**: `_spread` in `rally_gate.py` evaluates to `0.0` for arrays of length 2, which could cause a zero deadband if used on short clips.
- **Sparse trailing frames**: `_trim_inactive_tail` calculates frame density based on *visible* frames. A long gap followed by a single active frame gives 100% density, artificially inflating the rally end.
- **Silent failure on malformed CSV**: `load_canonical_csv` in `rally_cvat.py` silently ignores rows without a valid `rally_number`.

Goal: Fix these three issues to harden the rally detection and ingestion pipelines without changing the existing heuristic thresholds.

4. Design / architecture
The changes will be entirely localized to the identified functions:
- `rally_gate.py`: Modify `_spread` to use `math.ceil()` or properly enforce a minimum subset size before trimming percentiles, ensuring it never inadvertently calculates a spread of 0 for non-identical values.
- `tracknet_runner.py`: Update `_trim_inactive_tail` to factor the *temporal length* (or total expected frames in the trailing window) into the density calculation, rather than purely relying on `win_count` (the number of visible frames).
- `rally_cvat.py`: Introduce `logger.warning` before the `continue` statement in `load_canonical_csv` to surface skipped rows during CVAT ingestion.

6. Honest limits ? what this does NOT do
This will NOT change the core AI validation (Phase 3) or the existing `velocity_thresh` values. It solely tightens up boundary calculation edge cases. It will not suddenly "fix" rallies that the tracking model missed entirely due to lighting or occlusion.

```
34      ## 7. Deliverables & progress tracker   ? **source of truth**
35      Legend: ? Todo · ? In progress · ? Done · ? Blocked/gated. **One small PR per row**,
branched from `origin/master`, no stacking. Update the row (status + PR link) as work lands.
36
37      | ID | Deliverable | Depends on | Gated? | Status | PR |
38      |----|-------------|-----------|--------|--------|----|
39      | P0 | Fix `_spread` and `_trim_inactive_tail` edge cases | ? | No | ? | ? |
40      | P1 | Add logging to `load_canonical_csv` for silent skips | ? | No | ? | ? |
41
42      ## 8. Open questions ? owner / external
43      1. Do we want `_trim_inactive_tail` to fall back to the existing behavior if there are
very few frames, or strictly enforce a minimum temporal density? (Assuming strict temporal
density is preferred for accuracy).
44
45      ## 9. Definition of done
46      - [ ] `_spread` bug fixed and tested.
47      - [ ] `_trim_inactive_tail` density calculation corrected.
48      - [ ] Warning log added for malformed rows in `rally_cvat.py`.
49      - [ ] Tests updated and passing.
50
51      ## 10. References
52      **Internal code anchors `[verified]`:**
53      - `backend/pipeline/detectors/rally_gate.py`
54      - `backend/pipeline/detectors/tracknet_runner.py`
55      - `src/parsers/cvat/rally_cvat.py`
56      **Internal docs:**
57      - `docs/RALLY_DETECTION_CODE_REVIEW.md`
58
```

40. assistant (2026-06-19T16:55:00.212Z)

Now I have exact line numbers. I'll submit a **COMMENT** review (the only event GitHub allows me as the author) with inline comments anchored to specific lines ? pushing back where the framing overstates, and crediting the genuine catches.

41. assistant (2026-06-19T16:55:28.729Z)


```
{
  "commit_id": "1e221d914ee5c65c6ab31f841853c256523c5ec5",
  "event": "COMMENT",
  "body": "Inline pass ? pushing back where the framing overstates, and flagging the genuinely
good catches (B is a real catch; D's termination analysis is correct). Net recommendation is
unchanged from my top-level review: drop A (cosmetic non-bug), re-home C to the `sports-data-
collector` repo, and promote B into the measured rally-END track rather than blind-patching it.
Event is COMMENT-only because I am the PR author. Appreciate the depth of the code reading here
? the asks below are about calibration and process, not effort.",
  "comments": [
    {
      "path": "docs/RALLY_DETECTION_CODE_REVIEW.md",
      "line": 13,
      "side": "RIGHT",
      "body": "Worth flagging: `min_crossings=2` is only the bare-function default in
`rally_gate.py`. In the served config `FusionConfig.min_crossings=0` (OFF), and `fusion_hybrid`
is not the default segmenter (`default_segmenter='yolo_hybrid'`), so Gate A does not actually
run in the out-of-box pipeline. The same function-signature-vs-config split bites elsewhere:
`max_window_duration` (sig 0/OFF vs config 6.0/ON) and the R4 thresholds. When describing live
behaviour, quote the config defaults in `backend/config/models.py`, not the function
signatures."
    },
    {
      "path": "docs/RALLY_DETECTION_CODE_REVIEW.md",
      "line": 31,
      "side": "RIGHT",
      "body": "Pushing back: this is a cosmetic non-bug, not something to fix. `_spread` is only
ever called by `estimate_play_axis` over the whole-video trajectory (thousands of points), never
```

on small slices, and it already returns 0.0 for len<2 *by design* (`rally_gate.py:39`). There is also no division by spread anywhere downstream ? its only consumer is `deadband = deadband_frac * span`, a multiply at `rally_gate.py:102` ? so a zero spread is harmless. Recommend dropping fix A; at most a one-line comment documenting the domain assumption. (Aside: the `_spread` that feeds the actual feature schema is a *different* function, `rally_seq.py:43`, using non-truncating np.percentile ? so even the future-reuse worry is already covered.)"
    },
    {
      "path": "docs/RALLY_DETECTION_CODE_REVIEW.md",
      "line": 38,
      "side": "RIGHT",
      "body": "Good catch, and credit where due ? this one is real. I reproduced it through the public API at the production operating point: a rally that truly ends ~2.0s over-extends to 3.87s, because `win_count` genuinely counts active frames, not elapsed time. The important caveat is that `_trim_inactive_tail` *is* R4 ? default-ON and measured (`window_inactivity_end=1.5`, `inactivity_min_density=0.30`; ?tolF1 +0.040, p=0.022, n=13 at `models.py:150-155`) ? and rally-END is the project's #1 quality gap. So this is not a free hardening; see my note on the proposed fix in the plan doc."
    },
    {
      "path": "docs/RALLY_DETECTION_CODE_REVIEW.md",
      "line": 43,
      "side": "RIGHT",
      "body": "Appreciated ? accurate, and a genuine observability gap (the module, and in fact the whole `src/` tree, has no logger at all). One nuance to fold in: a fully-empty CSV is already hard-gated by the CLI (`sys.exit(1)` + `emit_delivery_report`), so only *partial* undercounts slip through silently ? which is exactly what a warning would surface. Two asks: this lives in the sibling `sports-data-collector` repo so it cannot ride this PR, and prefer fitting the existing Issue/summary reporting over a lone logging channel nothing configures."
    },
    {
      "path": "docs/RALLY_DETECTION_CODE_REVIEW.md",
      "line": 50,
      "side": "RIGHT",
      "body": "Nice ? the analysis is correct: `_midpoint_split_index` returns an index in [1, len-1], so both halves are non-empty and strictly smaller and the worklist provably terminates. Two tiny precision notes: the code is an *iterative* worklist (the docstring literally says 'Iterative worklist to bound stack depth'), not recursive; and the `h`-variable snippet is the R5 path specifically (`_blob_fallback_chop`) ? W1's `_split_overlong_groups` uses `g` and only midpoint-chops under `hard_cap`. Substance is spot-on though."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 10,
      "side": "RIGHT",
      "body": "This relation is mis-stated. `RALLY_DETECTION_QUALITY_REPORT.md` is the paper-checkpoint doc; this cleanup neither extends nor relates to it. The genuinely relevant tracker for the rally-END piece (fix B) is `RALLY_DETECTION_GAP_CLOSURE.md` (the G3 rally-END over-extension gap). Please point there instead."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 15,
      "side": "RIGHT",
      "body": "'potential divide-by-zero' is not real ? there is no division by `_spread`/`span` anywhere in the code. Its only consumer is `deadband = deadband_frac * span` (a multiplication, `rally_gate.py:102`), and the deadband is then used only in a comparison (`abs(v) < deadband`). Please drop the divide-by-zero framing."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 19,
      "side": "RIGHT",
      "body": "`_spread` is never called on short clips ? it runs once per video over the entire visible trajectory (`fusion_hybrid.py:106`, cached per-trajectory). And a zero deadband is

harmless: `abs(v) < 0` is always False, so net-crossings are simply counted without jitter suppression. There is no failure mode here to fix."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 23,
      "side": "RIGHT",
      "body": "Fix B cannot honour this line. Swapping `_trim_inactive_tail`'s density denominator from active-frame count to temporal length silently re-defines the calibrated `inactivity_min_density=0.30` ? that is a threshold change in effect, just an implicit one. Simulated against a normal sparse-tracking rally (5-7.5 active-fps) it collapses a real 6s rally end by ~6s. So this work is explicitly NOT threshold-neutral, and the plan should say so."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 27,
      "side": "RIGHT",
      "body": "This would change production behaviour for no measured benefit: `int()`->`ceil()` shifts the p5/p95 indices on the real full-trajectory estimate, moving `span`, the deadband, and the kept/dropped-rally decisions on every video. That is a Launch-Report-gated default flip, not a harmless edge-case fix. Recommend dropping it."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 28,
      "side": "RIGHT",
      "body": "This is the item with real merit, but it must not be a blind edit. `_trim_inactive_tail` is R4 ? default-ON, tuned, and measured (see my note at the code-review B section). Changing the denominator without re-tuning `min_density` regresses normal rallies (~6s in simulation). Route this into the owner-gated rally-END track (`RALLY_DETECTION_GAP_CLOSURE.md`), flag-gated, with a Launch Report (both eval rulers + over-seg guard) + dual-set no-regression, and re-tune `min_density` for the new denominator. The open question in section 8 is exactly the unresolved part ? it needs answering before code, not after."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 29,
      "side": "RIGHT",
      "body": "Fine and useful, but this targets the sibling `sports-data-collector` repo, so per our one-PR-per-item rule it cannot land in this PR. Open it there. Also consider surfacing skipped-row counts through the existing `emit_delivery_report` / Issue summary rather than introducing a logging channel the project does not configure."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 49,
      "side": "RIGHT",
      "body": "'Tests updated and passing' is not a sufficient DoD for fix B. It is a boundary-behaviour change on the #1 quality lever, so the DoD must include a measured Launch Report and dual-set (RAW + golden) no-regression / ?tolF1 ? not just a green suite. A passing suite proves the new math is internally consistent, not that rally-END quality held."
    },
    {
      "path": "docs/RALLY_DETECTION_FIXES_PLAN.md",
      "line": 55,
      "side": "RIGHT",
      "body": "This anchor is marked `[verified]`, but `src/parsers/cvat/rally_cvat.py` does not exist in this repo ? it lives in the sibling `sports-data-collector` repo. Please correct the `[verified]` label and re-home this reference."
    }
  ]
}

42. user (2026-06-19T16:55:28.758Z)

```
File created successfully at: C:\Users\avidu\Projects\badminton-highlight-
indexer\artifacts\pr245_review.json (file state is current in your context ? no need to Read it
back)
```

43. user (2026-06-19T16:55:37.870Z)

```
JSON valid
posted review id=4534316763 state=COMMENTED comments=true
```

44. assistant (2026-06-19T16:55:42.509Z)

```
Posted ? review `4534316763` (COMMENTED) with 14 inline comments. Let me confirm they all
anchored correctly.
```

45. user (2026-06-19T16:55:44.372Z)

```
docs/RALLY_DETECTION_CODE_REVIEW.md:13 ? Worth flagging: `min_crossings=2` is only the bare-
function default in
docs/RALLY_DETECTION_CODE_REVIEW.md:31 ? Pushing back: this is a cosmetic non-bug, not something
to fix. `_spre
docs/RALLY_DETECTION_CODE_REVIEW.md:38 ? Good catch, and credit where due ? this one is real. I
reproduced it t
docs/RALLY_DETECTION_CODE_REVIEW.md:43 ? Appreciated ? accurate, and a genuine observability gap
(the module, a
docs/RALLY_DETECTION_CODE_REVIEW.md:50 ? Nice ? the analysis is correct: `_midpoint_split_index`
returns an ind
docs/RALLY_DETECTION_FIXES_PLAN.md:10 ? This relation is mis-stated.
`RALLY_DETECTION_QUALITY_REPORT.md` is th
docs/RALLY_DETECTION_FIXES_PLAN.md:15 ? 'potential divide-by-zero' is not real ? there is no
division by `_spr
docs/RALLY_DETECTION_FIXES_PLAN.md:19 ? `_spread` is never called on short clips ? it runs once
per video over
docs/RALLY_DETECTION_FIXES_PLAN.md:23 ? Fix B cannot honour this line. Swapping
`_trim_inactive_tail`'s densit
docs/RALLY_DETECTION_FIXES_PLAN.md:27 ? This would change production behaviour for no measured
benefit: `int()
docs/RALLY_DETECTION_FIXES_PLAN.md:28 ? This is the item with real merit, but it must not be a
blind edit. `_t
docs/RALLY_DETECTION_FIXES_PLAN.md:29 ? Fine and useful, but this targets the sibling `sports-
data-collector`
docs/RALLY_DETECTION_FIXES_PLAN.md:49 ? 'Tests updated and passing' is not a sufficient DoD for
fix B. It is a
docs/RALLY_DETECTION_FIXES_PLAN.md:55 ? This anchor is marked `[verified]`, but
`src/parsers/cvat/rally_cvat.p
```

46. assistant (2026-06-19T16:56:04.925Z)

```
Done ? engaged at the line level on the PR, balancing push-back and credit. Two things are now
live:

- **Top-level review comment** ? [the full assessment +
recommendation](https://github.com/avidullu/badminton-highlight-
indexer/pull/245#issuecomment-4753405903)
- **A COMMENT review with 14 inline comments** ([review
4534316763](https://github.com/avidullu/badminton-highlight-
indexer/pull/245#pullrequestreview-4534316763)) ? all anchored correctly

(GitHub only let me use the **COMMENT** event, not a formal *Request changes*, because I'm
authenticated as `avidullu`, the PR author ? noted that in the review body so it's not mistaken
for a sign-off.)

**Where I pushed back** (claims that don't survive contact with the code):
- The TL;DR's **"divide-by-zero"** ? fabricated; `span` is only ever multiplied (`deadband =
deadband_frac * span`).
```

- **Fix A** as a whole ? cosmetic non-bug; the `int()?ceil()` "fix" would silently move the play-axis span on every video for no measured benefit.
- **"without changing the existing heuristic thresholds"** ? fix B *does* change a threshold implicitly (re-defines the calibrated `min_density=0.30`).
- **DoD "tests passing"** ? insufficient for a boundary change on the #1 quality lever.
- The mis-stated **"Extends QUALITY_REPORT"** relation and the `[verified]`-but-nonexistent cvat anchor (sibling repo).

Where I gave credit (genuinely good work):
- **§B** ? a real, reproduced edge case (well spotted; I just flagged that it lands on the measured R4 lever).
- **§D** ? the splitter termination-safety analysis is correct (minor: it's *iterative*, not recursive).
- **§C** ? a legitimate observability gap (just needs to land in the right repo).

My call stands: don't merge as-is, and there's nothing to "fund." Drop A, re-home C to `sports-data-collector`, and promote B into the measured rally-END track (`RALLY_DETECTION_GAP_CLOSURE.md`) with a Launch Report + dual-set regression. The contributor clearly read the code ? the gap is calibration, not effort.

Want me to take either of the follow-ups off your plate? I can (a) open the C logging fix as its own PR in `sports-data-collector`, and/or (b) draft a properly-measured plan for the B / sparse-tail end-rule under the gap-closure tracker. Both are squarely owner-gated areas, so I'll only start if you greenlight.

47. user (2026-06-19T16:56:08.743Z)

So between you and me, I should just ask to close this PR?

48. assistant (2026-06-19T16:56:47.656Z)

Yeah ? honestly, I'd close it. Here's the candid version.

The PR is a *container*, and once you work through the four items there's nothing left inside it worth merging:

- **A** ? drop (non-bug).
- **B** ? real catch, but it can't be done as written; it has to move into the measured rally-END track with a Launch Report. That's a different doc and a different process, not an edit to this plan.
- **C** ? belongs in `sports-data-collector`, not this repo.
- **D** ? praise, no action.

So every actionable thread either gets dropped, re-homed to another repo, or re-scoped into a different tracked effort. Merging the two docs as-is would actively *hurt* you: the plan greenlights a blind change to a measured lever and a fabricated "divide-by-zero," and the code-review doc would inject stale/incorrect claims into `docs/` ? exactly the kind of drift you just spent #242/#244 cleaning up. This isn't "iterate in place," it's "close and redirect."

The "approve, then I'll implement" framing is the tell: it's asking you to pre-bless a plan before the work, and the plan is mostly wrong. Don't approve a plan you'd reject at implementation.

One caveat before you close: don't let the one real signal evaporate. The B sparse-tail quirk is a legitimate, reproduced catch. I'd capture it durably first ? either a one-line entry under `RALLY_DETECTION_GAP_CLOSURE.md` (G3) or a skipped/xfail regression test that pins the current behavior ? so closing the PR doesn't throw away the good part.

So my honest read: **close it, graciously, with a pointer to where B should actually live** ? not a silent close, since the contributor clearly did real reading and you want that energy aimed at the right target next time.

Want me to:
1. Capture the B sparse-tail catch (a line in the gap-closure tracker, or an xfail test), and
2. Draft a short, appreciative closing comment that explains the redirect ?

then you hit close (or tell me to)? I won't close it myself without your explicit go-ahead.

49. user (2026-06-19T17:05:29.950Z)

Done. Closed it. Now can you tell me why is current-state.md so complex? Can you not shard it or simplify it?

50. assistant (2026-06-19T17:05:33.517Z)

I'll take a look at the actual file before judging it.

51. user (2026-06-19T17:05:34.974Z)

File content (315.8KB) exceeds maximum allowed size (256KB). Use offset and limit parameters to read specific portions of the file, or search for specific content instead of reading the whole file.

52. assistant (2026-06-19T17:05:39.298Z)

315 KB ? that's the problem in one number. A file meant to be read first every session has ballooned into something larger than most source files. Let me see how it's structured and why.

53. user (2026-06-19T17:05:41.316Z)

```
=== size + line count ===
  2376 323424 current-state.md
=== all headings (structure) ===
10:# Current state ? updated continuously (see [[working-agreement-checkpoints]])
1854:## DISTILLED MODEL BUILT (2026-06-06) ? user wants this approach; honest result = needs
more data
1878:## ? AUDIO SCAFFOLDING + PRELIMINARY E1 DONE (2026-06-07) ? PR #35 (feat/audio-hit-
detector-e1, OPEN). Verdict = AMBER.
1906:## ? DECISION DOCUMENTED (2026-06-07): AUDIO next, STANCE parked ? PR #32 (docs/audio-plan-
adr007)
1920:## ? RESEARCH DONE (2026-06-06) ? rally-detection methods ?
`docs/RALLY_DETECTION_RESEARCH.md` (full cited report)
1930:## ? VLC RALLY ANNOTATOR ? STANDALONE PUBLIC REPO (2026-06-07): github.com/avidullu/rally-
annotator
1951:## ? GOLDEN DATA via SELF-RATING (confirmed 2026-06-06) ? preferred over Gemini silver
1961:## ?? NEXT-WEEK RESUMPTION RUNBOOK (more teacher videos ? re-run calibrate_local +
distill_local; learned beats baseline once n grows)
1985:## OLD checkpoint context below ? (this session's journey)
1987:## DEMO RUN ? 2nd video "Adarsh and Vasanth.MP4" (2026-06-06, stress + checkpoint test)
1997:## BASELINE vs TUNED on Adarsh (2026-06-06) ? tuned config does NOT transfer + over-fire
pattern
2029:## GEMINI REFINE (2026-06-06) ? WASB candidates -> Gemini verify/refine (in progress)
2048:## PRECISION FINDING (2026-06-06): full-context Gemini > candidate-refine (envelope over-
merges)
2069:## ? TEACHER?STUDENT DISTILLATION (user's framing, 2026-06-06): Gemini=offline teacher,
LOCAL-ONLY WASB+gate=student
2085:## ? RAN local calibration vs Gemini silver-GT (2 videos: TestLargeVideo=7 rallies,
Adarsh=36) ? HONEST result
2098:## CODE_MAP artifact ? PR #27 (`docs/code-map`, OPEN) ? see [[code-map-artifact]]
2104:## ? NEXT-SESSION pickup
2116:## ? (1) EXPORT TUNED SEGMENTS DONE ? PR #26 MERGED to master (`feat/wasb-export-segments`)
2139:## ? PR status: #25 (checkpointing) + #26 (export) BOTH MERGED to master. No open indexer
PRs.
2141:## NEW session 2026-06-06 (late) ? user asked for tasks in order: (4) repo cleanup ? (2)
WASB checkpointing ? (1) export tuned segments
2164:## Merged to master (both repos clean, branches deleted)
2168:## WASB inference API (reverse-engineered 2026-06-06 ? for the wrapper)
2213:## GOAL CORRECTION (user clarified 2026-06-06 ? important)
2225:## Tuning plan = a Workflow (per user's ultracode ask), runs AFTER the wrapper exists
2230:## ? PLAN: checkpoint the heavyweight WASB pass (resume across restarts) ? IMPLEMENT NEXT
```

```
SESSION
2252:## ? STEP 4 ? FIRST CALIBRATION DONE (2026-06-06) ? RESULT
2266:## (historical) STEP 4 IN PROGRESS (2026-06-06)
2295:## NEXT-SESSION RUNBOOK ? step 4 (the actual calibration run)
2317:## Eval/overfit harness BUILT + license soul-check (2026-06-06)
2329:## Step-3 substrate finding (2026-06-06)
2336:## Merged: #6, #7, #19, #8, #20, **#9 (validate-delivery)**
2340:## Open PR (pushed, awaiting merge)
2346:## Pilot status ? DONE verifying
2353:## Decisions locked
2363:## Next steps (sequence 1?2?3, user approved)
2374:## Open design threads
```

54. assistant (2026-06-19T17:05:49.634Z)

The structure tells the whole story: there's a single heading at line 10, then nothing until line 1854 ? so lines ~10?1853 are **one massive ever-growing block**, and lines 1854?2376 are a graveyard of historical checkpoints back to 2026-06-06 that were never pruned. Let me see the top (the part actually meant to be "live") and sample the bloat.

55. user (2026-06-19T17:05:51.622Z)

```
1       ---
2       name: current-state
3       description: "Live micro-checkpoint state for Claude ? updated every checkpoint (not
only on request). What just landed, what's next, open decisions."
4       metadata:
5         node_type: memory
6         type: project
7         originSessionId: b13673cc-f087-43b1-9527-35dc0b67b301
8       ---
9
10      # Current state ? updated continuously (see [[working-agreement-checkpoints]])
11
12      **? HANDOFF (2026-06-19, clean-slate) ? NEXT SESSION START HERE for the prominent-court
/ multi-shuttle DESIGN journey.** Owner is "ramping up on the design mindset"; the prior
session's context filled after a marathon (9 PRs merged today:
#220/#237/#238/#239/#240/#241/#242/#243/#244). Owner asked for an honest call ? handed off here
so the design work continues fresh.
13      - **WHERE WE ARE (prominent-court track, gap-closure G2):** designed + built +
validated. `docs/PROMINENT_COURT_DETECTION.md` (tracked, C0-C8) is the design.
`backend/pipeline/detectors/court_gate.py` (default-OFF, merged #243) has the **window-level**
gate `gate_intervals_to_court` (drop a rally whose shuttle is MAJORITY outside a normalized
`court_polygon`) ? wired into `rally_seg_eval.windows_for` via a per-video `court_polygon`.
Validated on GX010142: cleanly drops the adjacent-court FPs incl the owner's rally #1, zero
central false-drops. §4.5 (#244) captures the owner's **floor/white-line shuttle-floor-contact =
deterministic rally-END** idea (C8, targets G3).
14      - **? VALIDATIONS LOGGED (2026-06-19, `scratch/validate_multishuttle_142.py`) ? answers
the owner's 3 questions:**
15        - **Q1 CAN WASB GIVE ALL SHUTTLES? YES.** Code: `~/models/WASB-
SBDT/src/detectors/postprocessor.py::_detect_blob_concomp` loops ALL connected components (no
top-k/cap) ? detector returns every ?0.5 heatmap blob; the tracker is the sole single-shuttle
reducer. Data: ?2-candidate frames are a STABLE venue property ? GX010142 31.2% / 143 32.3% /
145 26.7% / 147 27.7%, max 6/frame; the 2nd-best blob-mass (median 9.5-11.7) is comparable to
the top (13.7-14.9) ? genuine, not noise. Bound: "all" = all the model fires ?0.5 on (a
small/distant 2nd shuttle it misses won't appear; some ?2 are noise ? selection must
disambiguate).
16        - **Q2 CAN WE PICK THE PROMINENT SHUTTLE? YES.** A court-AGNOSTIC greedy multi-target
linker over the cache found 9 tracks in 21-25s; the **LONGEST/most-continuous (2.4s) is the
ADJACENT one** (xmean-norm 0.90) ? naive "pick longest/most-confident" (? what the single
tracker does) picks WRONG; a court+motion+vertical-excursion objective correctly flags the
prominent fragments (xmean 0.44-0.68, in-court 100%). Caveats: prominent rally came back
FRAGMENTED (~4 tracks ? need same-court fragment stitching; the time-windower mostly does this);
separated cleanly because courts are X-apart HERE (X-overlapping courts need a reliable court
signal ? Q3).
```

17 - **Q3 DO PERSON + FLOOR COALESCE? NOT YET ? blocker is the INPUTS, not the fusion design.** `output/person_GX010142.json` person detection is BLIND on this wide-GoPro footage: clip-wide **max conf 0.532**, only 7/2994 frames ?0.3, **ZERO confident persons in the 21-25 window** (players are visibly there). ? can't anchor court, corroborate rallies, or even test the rally-#1 "players haven't started" insight. Floor: prominent fragment descends only to y-norm 0.64 (no landing in 2.4s) + player-feet floor estimate unreliable ? **floor must come from COURT-LINE detection (§4.5/C8), not feet** (validates that design choice). The shuttle layer (Q1+Q2) stands alone; person needs a stronger detector, floor needs court lines, before they coalesce. ? **NEW work item: fix person extraction (lower threshold / better detector / fisheye) ? it's a prerequisite for person-anchoring + person-state rally-END.**
18 - **? THE LIVE DESIGN CRUX (validated on real data ? pick this up):** the gate works on the single TRACKED shuttle, but the **tracker** (not the detector) collapses to one shuttle/frame. ? **CORRECTION (2026-06-19, `det_raw.jsonl` probe ? refutes the earlier "isn't in the data / can't be saved"):** WASB's **detector already emits MULTIPLE shuttle candidates/frame** (cached in `det_raw.jsonl`; ?2 candidates on **31%** of detected GX010142 frames). It is the single-target **online tracker** (`wasb_infer.py:537 tracker.update`) that discards all-but-one. MEASURED on GX010142 rally #3 (21-25s): the *tracked* shuttle is 94% adjacent-court (x-norm ~0.88), BUT the raw candidates ALSO carry a **prominent-court shuttle in 83% of the window's frames** (x-norm ~0.56), **both shuttles present in 76%**, and a greedy-NN reconstruction yields a **continuous prominent-court rally track** (75% coverage, 431px vertical excursion, 8.6px median per-frame jump, ~2.4s sustained). **So the prominent rally the owner saw IS in the data** ? recoverable by **multi-track selection over the already-cached candidates (zero GPU)**, just not by gate-tuning the single track. ? reframes the whole thread: the ceiling is **track SELECTION, not the eyes**. **NEXT DESIGN THREAD: multi-track + court-aware selection** (a window is valid if ANY sustained track lives in the prominent court) ? this is **recall-positive** (recovers the prominent rally, a G1 miss) AND precision-positive (drops the adjacent one), unlike the current subtractive single-track gate. It also subsumes the **ghost** (a stuck-pixel track has ~0 motion energy ? never selected). Repros: `scratch/analyze_multishuttle_142.py` + `scratch/reconstruct_prominent_track_142.py` + `scratch/overlay_tracks_142.py` (trajectory overlay) + `scratch/extract_multishuttle_frames_142.py`. ? **VISUALLY CONFIRMED (2026-06-19, `artifacts/multishuttle_142/tracks_overlay_crop.png`):** the green prominent track is a textbook **rally arc anchored to a player mid-stroke** (bottom-left) ? second player (center) ? a real foreground-court rally; the yellow tracker output is a tangle over the **far-right adjacent court** (player at right margin). Two shuttles, two courts, detector saw both. ? Corrected mid-investigation: the adjacent track is NOT a static fixture (528px y-span ? a real moving adjacent rally). ? Still n=1 clip; 31% multi-candidate includes detector noise (selection scoring must disambiguate). Owner collecting more two-shuttle examples; **DO NOT PR yet** (owner's call ? this is design).
19 - **OTHER OPEN THREADS:** (C6) **player-anchored polygon auto-derivation** ? person features ALREADY extracted on GX010142 ? `output/person_GX010142.json` (`scratch/extract_person_19june.py`; median 3 players/frame). Raw player COUNT doesn't separate rally #1 (need court-LOCALIZED occupancy via `fusion_features.in_court_counts` + the same polygon ? players inside the prominent court ? 0 during rally #1). Foot-points can auto-derive the polygon. (Analysis DONE) **multicourt quality workflow measured the gate vs golden on the 4 multicourt clips ? `output/multicourt_analysis.md`.** Honest result: the gate is a **REAL precision win ONLY where a genuinely separable adjacent court exists** ? **Boxhill_Doubles ?P +0.096 / ?F1 +0.078, recall FLAT, 20 adjacent-court FPs dropped (all median x>0.80)**. NULL/marginal elsewhere: mahadevpura_2 +0.025 (trims 4 edge FPs, no separable court), **GX010128 + Badminton_BXH_2 = exact NULL** (their over-fire is on-prominent-court over-merge/boundary FPs the court gate can't touch, NOT adjacent-court rallies). ? **2nd detection-layer finding (besides single-shuttle): STUCK-TRACKER artifacts** ? GX010128 has 284 pts at exactly x=1680px, BXH_2 ~41% of pts parked at court-center ? a fixed-pixel WASB false-lock (a data-quality bug, not a court issue). ? NO Gemini preds for these ? "closeness to Gemini" can't be direct; Gemini-on-multicourt run = a follow-up (alias degraded today). **Verdict: the prominent-court gate is a worthwhile precision lever for TRUE multicourt (separable adjacent court), not a cure-all for multicourt over-fire (which has 3 distinct causes: adjacent-court rallies [gate fixes], on-court over-merge [other levers], stuck-tracker [detection fix]).**
20 - **RAMP-UP KIT:** `docs/PROMINENT_COURT_DETECTION.md` (the design + tracker) · `backend/pipeline/detectors/court_gate.py` + `backend/pipeline/detectors/wasb_runner.py` (the single-shuttle constraint) · `tests/test_court_gate.py` · the litmus pattern in this session (windows_for + gate_intervals_to_court on a cached `output/wasb/<clip>_*_wasb.csv`) · `docs/RALLY_DETECTION_GAP_CLOSURE.md` (G1/G2/G3) · scratch: `analyze_19june.py`, `extract_person_19june.py`. **Branch:** currently `docs/court-floor-rally-end` (merged); base new work off `origin/master`. [[paper-coauthor-aim]] [[decision-rigor-norm]] [[working-

agreement-merging]]

Checkpoint: 2026-06-19 (? **PROMINENT-COURT DESIGN+EXPERIMENT + DOC-STALENESS SWEEP ? 3 MORE PRs MERGED (#241/#242/#243)**). Owner directive: "extract person features" + "build up the 'prominent court' concept and make a design" (in multicourt, the foreground court covers the majority of the video area ? keep only shuttle rallies in it; may need a 'pseudo-3D' court model; experiment = `(x,y,visible)` ? `visible-in-prominent-court`) + "fix all staleness in the design docs and code, ultracode on, spin off agents" + **"auto accept, I'll be back, make and merge PRs"** (broad autonomy granted).

- ? **PROMINENT-COURT (gap-closure G2 / multicourt over-fire):** born from owner's **GX010142 rally #1** = a shuttle in the far-right adjacent court (verified: shuttle median X-norm **0.85**) that formed a 2.94s FP. **[#241](https://github.com/avidullu/badminton-highlight-indexer/pull/241) design doc** `docs/PROMINENT_COURT_DETECTION.md` (tracked, C0-C7; 2D floor polygon ? pseudo-3D play-volume; reuse `point_in_polygon`+`FusionConfig.court_polygon`+`PersonDetector.detections_per_frame`). **[#243](https://github.com/avidullu/badminton-highlight-indexer/pull/243) experiment** `backend/pipeline/detectors/court_gate.py` (default-OFF): **WINDOW-level** gate `gate_intervals_to_court` (drop a rally whose shuttle is MAJORITY out-of-court) wired into `rally_seg_eval.windows_for` via a per-video `court_polygon`. **KEY litmus finding:** point-level gating is TOO LEAKY ? rally #1 is 90% adjacent-court but a 10% central tail (12/121 pts) holds the window, surviving every 1D cut 0.66-0.82; the **window gate cleanly drops 6 rallies (all median X-norm 0.85-0.97 incl rally #1) with ZERO central false-drops.** ? confirms the design: 2D region + window-majority gate, NOT a 1D cut. ? Litmus polygon is crude/manual; **product path auto-derives it from PLAYER location (C4/C6)**. Eval n=15 UNCHANGED (no-op without polygon). 22 gate tests + suite 916 green, cov 83.07%.
- ? **PERSON FEATURES (C4) RUNNING (GPU):** `scratch/extract_person_19june.py` runs `PersonDetector.detections_per_frame` on GX010142 proxy ? `output/person_GX010142.json` for (a) player-anchored prominent-court derivation, (b) the rally-#1 "prominent players have not started" check (owner insight ? the PERSON-occupancy signal is the OTHER half of the rally-#1 fix; the court gate handles the shuttle side), (c) `shuttle_player_colocation` held-vs-flight. **shuttle_player_colocation IS BUILT** (`fusion_features.py:168`, default-OFF, unmeasured) ? corrected in the doc sweep.
- ? **DOC-STALENESS SWEEP [#242](https://github.com/avidullu/badminton-highlight-indexer/pull/242):** ran a **multi-agent Workflow** (7 themed auditors ? per-file adversarial verify; 30 candidates ? **25 confirmed**, each verified verbatim vs source). Fixed: signal build-status (Gate B / person cue / `shuttle_player_colocation` / `net_crossings` "[design]/NEW/unbuilt/eval-only" ? built-default-OFF-unmeasured), corpus n=6/9/12/13 ? **n=15**, CODE_MAP #233/#234 module moves, milestone-flip "owner-gated" ? LAUNCHED #204, CLAUDE.md/NEXT_STEPS GPU staleness, splits.py-DONE-#171, R2 `boundary_error_report` signature. ? One audit agent edited `R2_EVAL_METRIC_DESIGN.md` directly (rogue but accurate + in-scope) ? verified + included. Reusable: `output/_staleness_fixes.json`.
- ? **STILL IN-FLIGHT:** #243 CI (1 check, ? merge on green); person extraction (GPU). **NEXT:** analyze person features ? player-anchored polygon auto-derivation (C6) + rally-#1 player-activity check + colocation measurement; consider wiring the court gate into the SERVED fusion path. [[decision-rigor-norm]] [[working-agreement-merging]] [[paper-coauthor-aim]]

Checkpoint: 2026-06-19 (? **BOXHILL 19-JUNE LOCAL >7s EXTRACTION + GEMINI COMPARISON + 4 DUE-DILIGENCE PRs MERGED**). Owner: sync to remote head, verify sane, GPU local-only >7s extraction on 4-5 19-June videos ? new G: dir, Gemini comparison on 2-3 for golden-labeling, learnings on local-vs-Gemini gaps, fix issues via PRs as I go. **This box IS the owner's Windows machine** (CLAUDE.md "dev container lacks GPU/cv2/numpy" is fully STALE ? GPU+WSL+WASB all live here; RTX 2070 Super Max-Q).

- ? **Synced to origin/master** (local `master` is 50 behind + locked by stale `.claude/worktrees/tier1-windowing` worktree ? based work on `origin/master` directly). **Found + fixed a MASTER-BREAKING bug ? [#237](https://github.com/avidullu/badminton-highlight-indexer/pull/237) MERGED:** `test_ablation` called `ablation._is_fusion_schema` (moved to `experiment` by the #215/#235 SSOT refactor) ? **AttributeError at COLLECTION time ? whole local suite red** (green in CI only because the golden glob is empty there ? same green-CI/red-locally class as #236). Fix: import from the SSOT. Suite then 904 pass, cov 82.98%. **Quality NOT regressed** (n=15 tIoU F1@0.5 0.468, R2 tolF1 0.353, identical).
- ? **3 more PRs MERGED** (all verified-green, owner-authorized): [#238](https://github.com/avidullu/badminton-highlight-indexer/pull/238) `--min-rally-duration` CLI flag (surfaces #220's knob); [#239](https://github.com/avidullu/badminton-highlight-indexer/pull/239) re-baseline the nightly regression to n=15 (was red on corpus-growth, NOT a quality regress); [#240](https://github.com/avidullu/badminton-highlight-indexer/pull/240)

gemini surfaces failed chunks LOUDLY (a timeout/503-exhausted chunk silently dropped its time range ? partial reel that looked complete; mirrors the oversize guard).
- ? **INFRA FINDINGS (3):** (1) **`gemini-flash-latest` alias is currently DEGRADED** ? 11.6s for a trivial-text query vs `gemini-2.5-flash` 1.3s ? video chunks blew the 180s timeout ? 0 rallies (this is why the prior Gemini run "hung"). Pinned 2.5-flash for the reference run; did NOT change the repo default (owner-gated model choice ? FLAG for owner). (2) **WASB `timeout_sec=1800` (30min) too tight** for long 59.94fps clips: GX010144 (12.8min, ~43min WASB at ~3.4× RT) + GX010147 timed out ? 0 rallies (surfaced clearly in run.log, not silent). Re-running with 5400s. (3) **GoPro 4K is HEVC Main10/10-bit** ? `h264_nvenc` can't encode ? proxies need `format=yuv420p` (8-bit). GPU-decode proxy (NVDEC?scale_cuda?NVENC) ~4.5× RT.
- ? **LOCAL >7s reels published to `G:\?\Boxhill\2026-06-19 Boxhill - Local no-AI (7s+ rallies)`** (cap=6+R4 milestone config, `--min-rally-duration 7`, fully local): **all 5 reels DONE** ? GX010142 (44 rallies?**3 >7s**), GX010143 (38?**7**), GX010145 (55?**6**), GX010144 (58?**12**), GX010147 (40?**11**) = 39 >7s clips total. 144/147 first TIMED OUT at the 30min WASB default (12.8min + a slow-running 7min clip) ? re-ran at 5400s timeout. Detect on 1080p proxies in `F:/KhelSutra_work/19june_proxies/`.
- ? **GEMINI COMPARISON (3 single-court clips, 142/143/145; reusable `scratch/analyze_19june.py`, gemini intervals `output/gemini_*_19june.json`, local in `output/19june.db`):** **Local OVER-FIRES 2.08× (137 vs Gemini 66 rallies); recall-vs-Gemini 0.85, precision-vs-Gemini 0.41** (IoU0.3) ? the high-recall/low-precision signature, now on FRESH 19-June footage. **>7s counts close (L16/G13) BUT agreement LOW (only 5 match; long-rally precision 0.31/recall 0.38)** ? local's long windows are **over-merges/over-extensions** that don't 1:1 map to Gemini's clean long rallies (e.g. GX010145 has a 43s blob = degenerate merge). ? confirms gap-closure **G2 (over-fire, mostly short) + G3 (rally-END over-extension/merge)**. ? L-only ? all FP (could be Gemini-misses) ? **owner golden-labels 142/143/145 next ? REAL accuracy** (these become gap-closure §3 golden rows). Not committed to gap-closure doc yet (un-golden = agreement, not truth; [[eval-dual-set-regression]] D3). [[paper-coauthor-aim]] [[working-agreement-merging]]

Checkpoint: 2026-06-18 (? **SERVED Gate-A litmus RE-GROUNDED on the present multicourt golden set ? PR [#236](https://github.com/avidullu/badminton-highlight-indexer/pull/236) OPEN, owner-review-gated**). Closes the chip `task_f38284e5` (the 3 pre-existing cv2-gated fails flagged in the P7 checkpoint). **Discovered (corrects the task premise):** the original-6 single-court clips are GONE from the box (label dir `Annotation Setup/Collect/Golden Labelled/` deleted); present golden set = **9 multicourt clips** (labels now in repo `output/*.rallies.csv`), NOT "original-6 + added" ? original-6 corpus no longer measurable locally.
- **3 stale assertions fixed** (`tests/test_served_gate_a_regression.py`): (1) `len==6`?`==len(_SPECS) and >=6`; (2) served no-regression band re-measured **?F1 ?0.034** (was ?0.008), band ?0.06..0.03, kept the 0.5× baseline guard (ratio?0.83); (3) the +0.074 control RE-MEASURED on present corpus ? **?F1 ? ?0.001, over-seg 1.16?0.64** (was +0.074/1.68?1.01) + pinned the durable `seg_on<seg_off` direction-of-effect. Tests renamed for honesty; module docstring re-grounded to the recall-starved multicourt regime (Gate A only DROPS windows ? mild F1-loss when already under-segmented).
- **Owner decisions (AskUserQuestion):** re-measure test 3 on present corpus (not pin-to-original-6, not synthetic fixture); re-ground narrative+values; push+open-PR-for-review (NO self-merge); fold the doc fix into this PR.
- **Doc fix in same PR:** `docs/QUALITY_ITERATIONS.md` ? the two now-stale "standing guard reproduces +0.074 exactly" claims corrected to past-tense + a dated re-grounding note; +0.074/6-of-6/p=0.031 preserved as the real archived original-6 result.
- **Gate:** 5/5 litmus pass, full suite **864 pass / cov 81.23%**, ruff+mypy clean. ? CI green here = SKIPS this litmus (cv2-gated) ? the local run was the real check. Branch `fix/served-gate-a-corpus-regrounding` off origin/master (e67add1), 2 clean commits (e59b2fd test + de4240f doc), explicit-path staged, HEAD re-verified pre-commit. ? Magnitudes are a date-stamped snapshot ? re-measure as the golden set grows. [[eval-dual-set-regression]] [[launch-report-convention]] [[working-agreement-merging]] [[working-agreement-attribution]]

Checkpoint: 2026-06-18 (? **TECH-DEBT SWEEP #2 ? CODE_MAP + 6 deferred refactors MERGED (#229-#235), master `5f555cf`, CI green; 1 HELD, 4 reported-not-doable**). Owner: "do the code map + fix the 11 tech debts with the same rigor, no regression." Approach = **characterization tests FIRST** (pin current behavior ? prove they pass on today's code ? then behavior-preserving refactor), worktree-isolated agent fleet, combined-state gate, merge-on-green.
- ? **#229** docs: CODE_MAP UPLOAD endpoints re-pointed to `backend/api/uploads.py` (post-#228; download_highlight stayed in main).

43 - ? **#230** motion: `process_video`?`_detect_motion_segments`+`_populate_db_with_events` (new char tests mock cv2 module + seed `random` to pin the fabricated segment/events ? the previously-untested 204L path).
44 - ? **#231** rally_seg_eval: extract `_build_preset(args,side)` (11 char tests for the 0%-coverage report/main() paths ? module 0%?98%).
45 - ? **#232** database: `_execute_write` helper routes the 11 bool-returning single-statement writers (6 failure-branch char tests pin each EXACT log message ? byte-identical).
46 - ? **#233** golden_fixtures: split real-fixture subsystem into `golden_real_fixtures.py` with a **PEP-562 lazy `__getattr__` re-export** (handles both import orders; eager bottom-import would've cycled).
47 - ? **#234** main.py: job-worker drain/wake loop ? `backend/api/job_worker.py`; the monkeypatch-seam hazard solved by routing inter-fn calls through `backend.main` (`_main._arm_wake_timer`...) so `monkeypatch.setattr(m,...)` still intercepts (no leaked daemon timer). Added a minimal `mypy.ini` quarantine entry mirroring `[mypy-backend.main]` (same Optional-global debt). I spot-checked this one personally.
48 - ? **#235** eval: one shared parameterized `experiment.load_golden_videos` behind the 3 `load_videos` (column-lists + audio-policy args; char test pins the 3-way audio policy). Coverage 80.87% locally (?0.07pp = pure denominator effect from deleting 33 covered dup lines; absolute misses flat).
49 - **DUE DILIGENCE:** each PR = 1 clean commit + GitHub CI green; **combined-state gate** (all 7 merged into one worktree) = full suite **889 passed**, ruff+mypy(109) clean, **coverage 82.68% ? UP from 80.94%** (char tests net-added coverage); master push-CI = success. Worktrees/branches cleaned.
50 - ? **HELD (owner decision): `base-resolve-provider-dedup`** ? the gemini/yolo `_resolve_provider` bodies are NOT byte-identical (5 real diffs: signature, sport_profile capture-vs-discard, model_name source, an extra log line, return shape). A shared base method = a **behavior-changing redesign**, not a pure move ? agent correctly STOPPED. Needs owner approval + char tests around both segmenters' early-returns + the relocated log line if wanted.
51 - ? **NOT-DOABLE-AS-REFACTOR (reported, not forced):** trajectory_hybrid (already decomposed, hook-based base ? nothing to do); person_detector/tracknet grouping dedup (different tuple shapes ? redesign, real regression surface); `wasb_infer.run()` (GPU/WSL-only, needs owner-side torch-mocked test + real run); main.py routes-vs-CLI split (high blast radius, needs cv2 tests). [[working-agreement-merging]] [[decision-rigor-norm]]
52
53 **Checkpoint:** 2026-06-18 (? **REFACTORING SWEEP + BUG ? ALL 8 PRs (#221-#228) MERGED to master `e67add1`, master CI green**). Owner reviewed-and-approved by saying "merge when all suites + regression tests green." Before merging I did a **combined-state gate**: merged all 8 branches into one worktree (no conflicts, disjoint files) and ran the full suite (857 passed, ruff+mypy clean, cov 80.94%) so the AGGREGATE doesn't regress ? then squash-merged all 8; the push-triggered **master CI run = success**. NOTE: other sessions' PRs **#219 (docs/desktop-app-plan)** and **#220 (feat/long-rally-lens, awaiting owner review of numbers)** were intentionally LEFT OPEN (not mine, #220 owner-gated). Original sweep: a 2-phase discovery workflow (~18 agents): bug-hunt + tech-debt analysis over the grown files (mapped LOC + longest-functions: gemini.process_video=472L/71br worst, golden_fixtures=1113L, main.py=871L, database.py=738L). Produced a risk-rated worklist; the genuinely-risky/under-tested items were correctly DEFERRED (motion's random-heavy path, wasb GPU run(), main.py job-worker seam, golden_fixtures split, _execute_write dedup, rally_seg_eval reporting ? all "needs characterization tests first"). Shipped:
54 - **#221 BUG (`fix/config-loader-non-object-json`):** a valid-but-non-object `config.json` (`[]`/`42`/`"x"`/`null`) crashed startup (`.items()`/unpack), violating load_config's "non-fatal" contract ? main.py startup crash. One isinstance guard ? defaults+warn. Regression test proven to fail without fix. I implemented + verified this one personally.
55 - **7 behavior-PRESERVING refactors** (fanned out to worktree-isolated agents, each ran the full local gate then opened a PR): **#222** tracknet dedup+_active_frames · **#223** database _init_db?per-version helpers · **#224** compiler _trim_one_segment · **#225** gemini process_video?6 helpers (the 472L monster) · **#226** ablation export_pdf?ablation_pdf.py · **#227** yolo_hybrid process_video?phase helpers · **#228** main.py upload island?`backend/api/uploads.py` APIRouter (medium-risk; I spot-checked the global_config seam + lazy-import cycle-break myself).
56 - **VERIFIED INDEPENDENTLY (not just agent self-reports):** all 8 = exactly 1 clean commit off origin/master + **GitHub CI all green** (full suite + ruff + mypy + golden-fixtures cross-OS); local gates ~857 pass / cov ~80.7%. Pure-move refactors backed by existing tests (gemini/yolo have output-equivalence guards). **All OPEN for owner review ? NONE merged** (owner

said "I'll take a look").
- ? **Follow-up (out of scope of #228):** `docs/CODE_MAP.md` still lists the upload symbols under `backend.main` ? update if #228 merges. The 11 DEFERRED tech-debt items (need characterization tests first) are a future sweep. [[working-agreement-merging]] [[decision-rigor-norm]]

Checkpoint: 2026-06-18 (? **P7 LONG-RALLY (>5s) QUALITY LENS + REEL KNOB SHIPPED ? PR [#220](https://github.com/avidullu/badminton-highlight-indexer/pull/220), both default-OFF, awaiting owner review of the NUMBERS**). Owner: "bake in a long-rally lens (eval) + product knob today." Delivered + measured n=15:
- ? **Eval lens** (`backend/eval/rally_seg_eval.py`): `--min-duration <s>` filters BOTH preds + golden GTs to rallies >s before matching (tIoU AND R2 paths consistently; a FILTER not a new metric; default 0=OFF). `--stratify`/`--long-tau`(5s) ? `{all vs long}×{single vs multicourt}` table. Court type = committed `MULTICOURT_CLIPS` set (manifest is **gitignored** ? committed set is the auditable source; manifest `court_type` overrides). `predict_all` windows once, re-scores per stratum.
- ? **Product knob**: `stitching.min_rally_duration` (s, 0=OFF) drops short rally pairs before the reel compile in `tools/generate_highlights.py` ? cleaner than `min_shot_count` (Unknown shot_count). +unit & integration tests.
- ? **MEASURED PAYOFF (n=15, local vs golden, milestone cap6+R4) ? the hypothesis split 3 ways (honest, NUMBERS INFORM):** (1) ? single-court over-fire is **short-FP-dominated** ? restrict to >5s collapses seg_ratio 1.43?1.04 (preds?GTs). (2) ?? **multicourt FAILS the "FPs are short" thesis**: over-fires **1.58× even on long rallies** (adjacent-court games = real long rallies ? long *spurious* windows) ? multicourt precision needs **court-localization (Gate-B/court mask, gap-closure Lever C)**, NOT a duration filter. (3) ?? strict tolP does NOT rise (0.326?0.278 single) ? the >5s pred set is **over-merge-contaminated** (kushagra med 9.9s, mahadevpura_1 63s blob; cap=6 truncation breaks the strict |?end|?1.5 match; tIoU forgiving holds ~0.42). (4) ? product knob still a clean reel win (a reel needs no strict boundary tolerance). **Gemini-long NOT computable in-checkout** (Boxhill gemini intervals owner-side only) ? lens is metric-agnostic, feed Gemini intervals through `filter_by_min_duration`. Pooled: single 608pred/416gt (long 218:210=1.04), multi 288/179 (long 158:100=1.58).
- ? **Gate:** suite 868 pass, cov **80.96%** (?70 floor, ~baseline 80.72%, not regressed), ruff+mypy(104) clean. ? **3 PRE-EXISTING fails** in `test_served_gate_a_regression.py` (cv2-gated ? **skip in CI**; hardcode old n=6 corpus, now 9+ local) ? PROVEN pre-existing (reproduce with my changes stashed; served_gate_a imports none of my funcs). Spawned chip `task_f38284e5` to recalibrate (separate PR).
- ?? **SHARED-CHECKOUT COLLISION HIT + RECOVERED:** mid-work a sibling switched HEAD `feat/long-rally-lens`?`docs/desktop-app-plan` (b9ff235, the #219 work). Caught via a `git stash` WIP message naming the wrong branch; my 7 files were uncommitted+intact and identical between branch tips, so `git checkout feat/long-rally-lens` carried them cleanly; re-verified HEAD=a010667 + `origin/master..HEAD` lists ONLY my commit before commit/push. [[gotcha-shared-checkout-branch-collisions]] held ? do NOT commit while on the wrong branch.
- ? Docs: `QUALITY_ITERATIONS.md` "P7" + `RALLY_DETECTION_GAP_CLOSURE.md` §7 row P7 (? measured). **Defaults stay OFF until owner reviews the numbers** ([[working-agreement-merging]]: touches eval+product config ? owner-gated flip, NOT a self-merge). [[eval-dual-set-regression]] [[decision-rigor-norm]] [[launch-report-convention]]

Checkpoint: 2026-06-18 (? **NEW TRACKED SUBPROJECT SCOPED ? `docs/DESKTOP_APP_PLAN.md`, PR [#219](https://github.com/avidullu/badminton-highlight-indexer/pull/219) OPEN (docs-only, owner-gated, all `[design]` ? NOTHING built)**). Owner-requested exploratory doc for future sessions: package the local pipeline as a **cross-platform desktop app** (Win/mac/Linux) for non-tech recreational players ? plug in the camera USB/SD card, click once, get back ONLY the highlight reels, **fully on-device** (no upload, no original-video bloat). Follows `PROJECT_DOC_TEMPLATE` (status hdr · D1?D6 decisions · §7 tracker = source-of-truth · DoD).
- **Centerpiece = the de-WSL native-compute port** (the real "port to Linux/Mac" subproject): WASB/TrackNet run today in **WSL2+CUDA (Windows-only)**; must run natively per-OS ? Linux/Win **CUDA**, macOS (no CUDA) **MPS/CPU**. Grounded in the `DetectorRunner` ABC (`backend/pipeline/detectors/base.py`) which already documents a `LinuxNativeRunner`/`ONNXRunner` as "the main de-risk goal"; the torch inference core (`wasb_infer.py`) just needs lifting out of WSL (identical `Frame,Visibility,X,Y` CSV ? windowing/stitch unchanged).
- **MAKE-OR-BREAK unknown (the P0 spike):** compute on a typical **NO-GPU enthusiast laptop**. WASB ~**3.4× real-time** on a laptop 2070S (measured, COST_SCALING); CPU far slower; Apple MPS unverified. P0 = measure CPU/MPS for a ~10-min clip ? go/no-go BEFORE any app code. If

CPU-only too slow ? needs a **smaller exported owned model** (OWNED_MODEL export-clean ONNX/CoreML) or a tiered fast/slow design.
70 - **§7 tracker:** P0 compute-feasibility spike ? P1 de-WSL native detector ? P2 app shell (Tauri vs Electron vs CLI+installer over the Python/FastAPI backend) ? P3 packaging/code-signing per OS. Only P0 is un-gated; P1?P3 owner-greenlit after P0. **Licensing locked (D4):** MIT/Apache/BSD bundle only; **ffmpeg = LGPL build** (no GPL x264 ? openh264/hardware encoders); WASB weights MIT ?.

56. assistant (2026-06-19T17:07:48.821Z)

Now I see exactly why. This isn't really "complex" ? it's **unbounded append-only growth**. The file is a reverse-chronological stack of full session checkpoints going back to 2026-06-06 (2,376 lines), and nothing ever gets pruned. Only the top ~30 lines are "live"; the other ~2,340 are history that should have been archived. A file you're supposed to read *first, every session* has become too big to even open in one read.

Let me confirm the split boundary and grab the convention file that caused this, then shard it.

57. user (2026-06-19T17:07:53.033Z)

```
=== total lines ===
2376 current-state.md
=== boundary check (lines 33-35) ===
- ? **GEMINI COMPARISON (3 single-court clips, 142/143/14

**Checkpoint:** 2026-06-18 (? **SERVED Gate-A litmus RE-GR
=== count checkpoint blocks ===
98
=== oldest dates present ===
Checkpoint:** 2026-06-07
Checkpoint:** 2026-06-08
Checkpoint:** 2026-06-09
Checkpoint:** 2026-06-10
Checkpoint:** 2026-06-11
Checkpoint:** 2026-06-12
Checkpoint:** 2026-06-13
Checkpoint:** 2026-06-14
Checkpoint:** 2026-06-15
Checkpoint:** 2026-06-16
Checkpoint:** 2026-06-17
Checkpoint:** 2026-06-18
Checkpoint:** 2026-06-19
```

58. user (2026-06-19T17:07:53.070Z)

<system-reminder>This memory is 12 days old. Memories are point-in-time observations, not live state ? claims about code behavior or file:line citations may be outdated. Verify against current code before asserting as fact.</system-reminder>

```
1       ---
2       name: working-agreement-checkpoints
3       description: "Standing user preference ? check code in frequently as small PRs; keep a Claude-consumable current-state/checkpoint file updated at every micro-checkpoint"
4       metadata:
5         node_type: memory
6         type: feedback
7         originSessionId: b13673cc-f087-43b1-9527-35dc0b67b301
8       ---
9
10      Standing working preference (user confirmed 2026-06-06).
11
12      **The rule:** commit/check-in code **very periodically** as **small PRs** (micro-checkpoints),
13      and at each micro-checkpoint clearly record progress in a **`current-state` file written for
```

14   Claude's consumption** ? so a future session (or the user) can resume from the file, not
from
15   replaying the conversation. The living file is [[current-state]] (memory dir); per-repo
state
16   docs are fine too when a feature spans many commits.
17
18   **Why:** the user runs long multi-step build/verify loops across two repos and wants
frequent,
19   reviewable checkpoints rather than one big PR; the state file is the durable hand-off so
context
20   isn't lost between sessions.
21
22   **Long-running jobs must be observable + resumable (user, 2026-06-06):**
23   - **Periodic flushed progress logging** with ETA ? never a multi-minute (or longer)
process that's
24   a black box until it finishes. (Done for `wasb_infer` per-N-batches; apply the same to
any heavy
25   pass.) Flushed plain lines (not `\r` bars) so they show in background-task logs.
26   - **Checkpoint heavyweight compute** so a restart resumes, not restarts (see the WASB
checkpointing
27   plan in [[current-state]] ? incremental cache of the expensive stage + cheap re-run of
the rest).
28
29   **How to apply:**
30   - Prefer many small, isolated branches/PRs over one large one (see [[session-handoff]]
for the
31   isolation convention). Push + open the PR when a coherent unit is done ? don't sit on
commits.
32   - After each meaningful commit/PR, update the current-state file: what landed, what's
next,
33   any open decisions. Keep it terse and Claude-readable.
34   - **Two docs, both kept fresh automatically** (user upgraded this 2026-06-06; global
CLAUDE.md
35   updated to match):
36   - [[current-state]] ? granular/live, flushed at **every** micro-checkpoint.
37   - [[session-handoff]] ? ~1-page orientation, **auto-refreshed at meaningful
checkpoints** (PR
38   pushed/merged, key decision, phase shift, session wrap-up) ? no longer on-request-
only. The
39   "update/refresh the handoff" phrases still force an immediate full rewrite.
40   - Don't rewrite either every turn ? only when something worth resuming from changed.
41   - **Staleness tripwire (so it can't drift a whole session again ? confirmed by user
2026-06-07):**
42   if ?3 substantive [[current-state]] checkpoints ? or ANY PR-merge / new ADR / phase
shift ? have
43   landed since the handoff's "Last updated" date, refresh the handoff *now*.
Practically: any turn
44   where you update [[current-state]] for a merge/decision/phase-shift, also touch the
handoff.
45   - **Keep the handoff's OWN note consistent with this policy.** Root cause of the
2026-06-07 staleness:
46   the handoff file still said "updated only on request," contradicting this agreement,
and that note
47   was followed. The file's frontmatter/refresh-policy line must always match this
auto-refresh rule.
48

59. assistant (2026-06-19T17:08:03.123Z)

Confirmed: **98 checkpoint blocks** spanning 2026-06-07 ? 06-19 stacked in one file, and the convention doc that drives this ([working-agreement-checkpoints.md](../memory/working-agreement-checkpoints.md)) says to flush a checkpoint *every* micro-checkpoint but **never says to prune** ? that's the root cause.

Here's my fix: shard it into **live** (read-first, ~today only) + **archive** (full history,

preserved verbatim ? nothing lost), and add a prune rule so it can't regrow. Let me do it.

60. assistant (2026-06-19T17:08:07.102Z)

```
---
name: current-state-archive
description: "Chronological checkpoint history split out of current-state.md so the live file
stays read-first-short. Newest first; read ONLY for deep history."
metadata:
  node_type: memory
  type: project
---
```

Current-state ARCHIVE ? checkpoint history

> Split out of [[current-state]] on 2026-06-19 to keep the live file readable (it had grown to ~2.4k lines / 315 KB ? too big to open in one read). This holds the aged-out checkpoint blocks, newest first. Live state ? current handoff, what just landed, next steps, open decisions ? lives in [[current-state]]. When today's checkpoints age out, prepend them here.

61. user (2026-06-19T17:08:07.134Z)

```
File created successfully at: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-
highlight-indexer\memory\current-state-archive.md (file state is current in your context ? no
need to Read it back)
```

62. user (2026-06-19T17:08:15.452Z)

```
=== new live file ===
   34 16628 current-state.md
=== archive file ===
  2355 307513 current-state-archive.md
=== integrity: checkpoint blocks live+archive should = 98 ===
live: 2  archive: 96
=== last line of archive (should be the oldest content, not truncated) ===
```

Open design threads

```
- Per-video few-shot adaptation: calibration first (static camera = go
  (held-out rallies). cf. train/eval guardrail in [[collector-indexer-
```

63. assistant (2026-06-19T17:08:29.804Z)

Split verified ? **2 (live) + 96 (archive) = 98** checkpoints, zero lost, archive ends cleanly at the oldest 2026-06-06 content. Live file is now 34 lines / 16 KB (openable in one read) vs the old 315 KB. Let me add the read-first header note and the prune rule so it stays small, then clean up.

64. user (2026-06-19T17:08:30.454Z)

```
8       ---
9
10      # Current state ? updated continuously (see [[working-agreement-checkpoints]])
11
12      **? HANDOFF (2026-06-19, clean-slate) ? NEXT SESSION START HERE for the prominent-court
/ multi-shuttle DESIGN journey.** Owner is "ramping up on the design mindset"; the prior
session's context filled after a marathon (9 PRs merged today:
#220/#237/#238/#239/#240/#241/#242/#243/#244). Owner asked for an honest call ? handed off here
so the design work continues fresh.
```

65. user (2026-06-19T17:08:39.197Z)

```
The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\memory\current-state.md has been updated successfully. (file state is current in your
context ? no need to Read it back)
```

66. assistant (2026-06-19T17:08:46.203Z)

Now append the archive pointer footer to the live file, and clean up the backup.

67. user (2026-06-19T17:08:49.413Z)

```
=== final live file ===
   39 17164 current-state.md
(was 2376 lines / 315 KB)
```

68. assistant (2026-06-19T17:08:56.135Z)

Live file is now **39 lines / 17 KB** (was 2,376 / 315 KB). Now update the index and add the prune rule to the convention that caused the bloat.

69. user (2026-06-19T17:08:57.362Z)

```
<system-reminder>This memory is 2 days old. Memories are point-in-time observations, not live
state ? claims about code behavior or file:line citations may be outdated. Verify against
current code before asserting as fact.</system-reminder>
1       # Memory Index
2
3       ## ? Start here
4       - [Current state](current-state.md) ? **live micro-checkpoint** (updated every
checkpoint): open PRs, what just landed, next steps, open decisions. **Read this first.**
5       - [Session handoff](session-handoff.md) ? current state · next steps · open decisions ·
ramp-up kit. AUTO-refreshed at meaningful checkpoints (PR merged, key decision, phase shift,
session wrap) ? see working-agreement-checkpoints staleness tripwire.
6       - [Code map artifact](code-map-artifact.md) ? for CODE/architecture ramp-up, read
`docs/CODE_MAP.md` FIRST (dense anchor-rich map; replaces a file crawl). How to regenerate it.
7       - [File-placement conventions](file-placement-conventions.md) ? **where new files go**
(scripts/ vs backend/eval/ vs tools/ vs scratch/ vs training/ vs artifacts/ ?); **read before
committing a new file.** Authoritative table = `docs/CODE_MAP.md` §0 (PR #138); scratch/
gitignored (#136).
8
```

70. user (2026-06-19T17:09:04.266Z)

```
The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\memory\MEMORY.md has been updated successfully. (file state is current in your context ?
no need to Read it back)
```

71. assistant (2026-06-19T17:09:07.612Z)

Now the durable fix ? add the prune rule to the convention doc so it can't regrow:

72. user (2026-06-19T17:09:12.092Z)

```
The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\memory\working-agreement-checkpoints.md has been updated successfully. (file state is
current in your context ? no need to Read it back)
```

73. assistant (2026-06-19T17:09:15.724Z)

Let me do a final consistency check across the affected files.

74. user (2026-06-19T17:09:19.342Z)

```
=== files (no stray .bak/.tmp) ===
current-state-archive.md
current-state.md
=== sizes ===
    39  17164 current-state.md
```

```
  2355 307513 current-state-archive.md
  2394 324677 total
=== live file footer ===
---
**Older checkpoints (2026-06-06 ? 2026-06-18, 96 blocks) ? [[current-state-archive]].*
=== cross-links resolve (archive <-> live) ===
current-state.md
MEMORY.md
links present
```

75. assistant (2026-06-19T17:09:31.917Z)

Done. Here's the answer to both parts of your question.

Why it was so complex

It wasn't really complexity ? it was **unbounded append-only growth**. `current-state.md` is meant to be the file you read *first, every session*, but the checkpoint convention says "flush a checkpoint at every micro-checkpoint" and **never said to prune**. So 98 full session checkpoints (2026-06-06 ? 06-19) piled up in one file: **2,376 lines / 315 KB** ? so big it couldn't even be opened in a single read. Only the top ~30 lines were ever "live"; the other ~2,340 were history that should have been archived long ago.

What I did ? sharded it (nothing lost)

File	Before	After	Role
`current-state.md`	2,376 lines / 315 KB	**39 lines / 17 KB**	**Live only** ? current handoff + today's 2 checkpoints + the prune rule + a pointer
`current-state-archive.md`	?	2,355 lines	**Full history**, verbatim, newest-first (96 aged-out checkpoints)

Integrity checked: **2 live + 96 archived = 98** checkpoints, the archive ends cleanly at the oldest 2026-06-06 content, and no backup/temp files left behind.

I also closed the loop so it **can't regrow**:
- Added a **scope banner** at the top of the live file: live = handoff + *today's* checkpoints only.
- Added the **prune rule** to the convention that caused this ([working-agreement-checkpoints.md](../memory/working-agreement-checkpoints.md)): when checkpoints age out (new clean-slate day), move them to the archive instead of stacking ? checkpoints are still flushed every micro-checkpoint, just relocated once stale.
- Updated [MEMORY.md](../memory/MEMORY.md) "Start here" so the index points at both files and marks the archive as deep-history-only.

Net effect: the read-first file is now an actual ~1-pager, and the detailed measured history (which you value for the paper aim) is fully preserved, just out of the hot path.

One thing I deliberately **didn't** do: `session-handoff.md` overlaps with the live handoff block here, so there's still some duplication between the two "live" files. I left that alone since you asked specifically about `current-state.md` ? but if you want, I can reconcile them so the handoff block points at `session-handoff.md` instead of repeating it. Want me to?

76. user (2026-06-19T17:14:31.085Z)

we can do that later for sure and appreciate the house cleaning! Can you help me cleanup the docs/ folder and identify all the docs which are "partially or completed" projects/ideas/single efforts which we can move to docs/archives now?

77. assistant (2026-06-19T17:15:02.818Z)

Great ? glad the cleanup helps! Before touching anything, the project's archive policy is specific: **only terminal-state (DONE/REJECTED/superseded) work moves to `docs/archives/`; live docs stay at top level.** "Partially complete" is the tricky bucket ? genuinely-active work should *not* be archived even if incomplete, but *abandoned/superseded* efforts should. So this

is really a lifecycle-classification pass per doc, with evidence. Let me inventory first.

78. user (2026-06-19T17:15:08.008Z)

```
=== branch / HEAD ===
docs/court-floor-rally-end
430de76 docs(rally): capture the shuttle-floor-contact rally-END signal (owner note)

=== top-level docs/*.md (count) ===
52
=== top-level docs/*.md ===
docs/ABLATION_LAYER.md
docs/ANALYZERS_AND_RECONCILERS.md
docs/BACKLOG.md
docs/CODE_AUDIT_AND_TEST_HARDENING.md
docs/CODE_MAP.md
docs/COMMERCIALIZATION.md
docs/DEFERRED_HARDENING_PLAN.md
docs/DESKTOP_APP_PLAN.md
docs/EVALUATION.md
docs/EXECUTION_POLICY.md
docs/EXPERIMENT_HARNESS.md
docs/FIELD_VISIT_ANSWER_SHEET.md
docs/FIELD_VISIT_PLAYBOOK.md
docs/GOLDEN_DATA_SHARING.md
docs/GOLDEN_REGRESSION_FIXTURES.md
docs/GOLDEN_SET_IMPLEMENTATION.md
docs/GOLDEN_SET_PHASE1_VIDEOS.md
docs/GOLDEN_SET_VENDORING.md
docs/GOLDEN_VIDEOS.md
docs/HARDENING_LOOP_HANDOFF.md
docs/HOSTING_PLAN.md
docs/HOW_RALLY_DETECTION_WORKS.md
docs/I18N_PLAN.md
docs/KHELSUTRA_PRODUCT_OVERVIEW.md
docs/MULTI_SIGNAL_FUSION_PLAN.md
docs/NEXT_STEPS.md
docs/OWNED_MODEL_IMPLEMENTATION_PLAN.md
docs/PERSONALIZATION_PLAN.md
docs/PER_VIDEO_OUTPUT_LAYOUT.md
docs/PLATFORM_ARCHITECTURE.md
docs/PMF_FIELD_SIGNALS.md
docs/PMF_MARKET_RESEARCH.md
docs/POST_166_WORK_ITEMS_MILESTONES_AND_RISK_ASSESSMENT.md
docs/PRODUCT_FIELD_ASSOCIATE_FLYER.md
docs/PROJECT_DOC_TEMPLATE.md
docs/PROMINENT_COURT_DETECTION.md
docs/QUALITY_ITERATIONS.md
docs/R1_MILESTONE_LAUNCH_REPORT.md
docs/R2_EVAL_METRIC_DESIGN.md
docs/RALLY_DETECTION_GAP_CLOSURE.md
docs/RALLY_DETECTION_QUALITY_REPORT.md
docs/RALLY_QUALITY_RESEARCH.md
docs/README.md
docs/REAL_FOOTAGE_VALIDATION_REPORT.md
docs/ROADMAP.md
docs/ROORA_LABELING_GUIDELINES.md
docs/STORAGE_SHARING_MODEL.md
docs/TRACKNET_WSL_SETUP.md
docs/UI_ALPHA_EXECUTION_PLAN.md
docs/UI_PROPOSAL.md
docs/VIDEO_LOCALITY_MODEL.md
docs/VIDEO_RECORDING_GUIDELINES.md

=== docs/ subdirs ===
```

```
docs/archives/
docs/mockups/

=== existing archives layout ===
docs/archives
docs/archives/checkpoints
docs/archives/checkpoints/2026-06-eval-baseline
docs/archives/checkpoints/2026-06-logging-reporting
docs/archives/checkpoints/2026-06-scaffolding-complete
docs/archives/checkpoints/2026-06-strategy-foundation
docs/archives/decisions
docs/archives/past_projects
docs/archives/past_projects/low-regret-moves
docs/archives/quality-iterations
docs/archives/quality-iterations/2026-06-18-lovo-resweep-n13
docs/archives/quality-iterations/2026-06-first-ab-experiment
docs/archives/quality-iterations/2026-06-human-gt-tuning
docs/archives/quality-iterations/2026-06-initial
docs/archives/quality-iterations/2026-06-multivideo-lovo
docs/archives/research
```

79. assistant (2026-06-19T17:15:21.224Z)

52 top-level docs, and the archive has a clear taxonomy (`past_projects/`, `research/`, `decisions/`, `checkpoints/`, `quality-iterations/`). Let me read the doc index and pull every status header in one pass to seed the classification.

80. user (2026-06-19T17:15:22.178Z)

1 - ROORA_LABELING_GUIDELINES.md ? Latest (v8) vendor labeling guidelines mirrored from the collector's `docs/annotation/`, with the fusion-driven court-calibration / singles-doubles additions marked `? ADDED (fusion)`.
2 - RALLY_QUALITY_RESEARCH.md ? **? Foundation doc** for closing the local-vs-Gemini rally-detection gap: the measured gap + the code-grounded over-merge diagnosis ("shuttle moving ? rally"), an inventory of the eval/experiment harness + its gaps, the strengthened evaluation framework (segmental Edit/F1@k that make over-segmentation first-class, golden-corpus + active learning, Gemini-agreement as a *shadow* signal), an **adversarially-verified, cited external-research synthesis** (TAL/TAS boundary heads, racket-sports rally-state cues, small-data, eval rigor), a **prioritized cheap?durable roadmap with falsifiable success criteria**, and (§7) a **tracked project plan** ? work-item checklist, sequencing, the single quality **number** attributed to the effort, and a **definition of done** (project completes when all tiers land + the number is recorded). The plan that the QUALITY_ITERATIONS.md log executes against; ties together EXPERIMENT_HARNESS / MULTI_SIGNAL_FUSION_PLAN / ANALYZERS_AND_RECONCILERS / OWNED_MODEL_IMPLEMENTATION_PLAN.
3 - RALLY_DETECTION_GAP_CLOSURE.md ? **Tracked execution doc (2026-06-18)** for closing the remaining rally-detection accuracy gap **per video, before launch**, driven by the growing golden corpus. First ground-truth clip (GX010141) reframed it: boundaries are tight (R2/R3/R4 holds) ? the gap is **detection**: (G1) recall (local & Gemini both miss ~half on hard clips, and *different* rallies ? complementary), (G2) precision (local over-fires), (G3) **rally-END over-extension on the quick-return-after-point** (shuttle still moving ? needs a *player-state* signal). Levers (owner-gated): player-kinematics + shuttle-in-hand / sustained-invisibility end-rule with reverse-timeline backtrack, `rally_gate` Gate A wiring, complementary fusion. **Peer** of `RALLY_QUALITY_RESEARCH.md` (execution tracker, not research); §7 is the source of truth.
4 - PROMINENT_COURT_DETECTION.md ? **Design + experiment (2026-06-19, default-OFF, owner-gated flip)** for the **multicourt G2 over-fire**: identify the **prominent (foreground) court** (the one covering the majority of the frame) and keep only rallies whose shuttle lives inside it ? re-defining the shuttle signal `(x,y,visible)` ? `(x,y,visible-AND-inside-prominent-court)` *before* windowing, so a background/adjacent-court shuttle never forms a rally. 2D floor polygon (MVP) ? **pseudo-3D play-volume** (refinement, so high clears aren't clipped). Born from the GX010142 rally-#1 background-court FP; reuses `fusion_features.point_in_polygon` + `FusionConfig.court_polygon`. §8 progress tracker is the source of truth.
5 - RALLY_DETECTION_QUALITY_REPORT.md ? **Technical-

paper checkpoint (2026-06-17)** of rally-detection quality: the honest measured state (over-segmentation milestone tolerance-F1 0.091?0.323 at n=11, p=0.002; merge-rate 0.694?0.150; first real-footage ground-truth at-par with START at Gemini parity), the task-faithful R2 metric, a detailed Gemini comparison, next steps + expected wins, areas yet to be explored, and a publishability / path-to-submission verdict. Renders to a polished PDF; the externally-shareable "toward a paper" summary of `RALLY_QUALITY_RESEARCH.md` + `QUALITY_ITERATIONS.md`.

- EVALUATION.md ? How we measure and improve quality.
- EXPERIMENT_HARNESS.md ? Design (P1 SHIPPED) for A/B + shadow + SxS experimentation on rally-detection variants with **zero production regression** ? experiments are decoupled from deploy (new cues default OFF, promote only on measured LOVO lift, rollback = flip a flag). Composes the existing `backend/eval/` machinery (`calibration`, `regression`, `metrics`); implemented in `backend/eval/experiment.py`.
- QUALITY_ITERATIONS.md ? **Living log** of rally-detection quality iterations (hypothesis ? what shipped ? measured/expected effect ? lesson). The reverse-chronological record; terminal-state snapshots get archived under `archives/quality-iterations/`.
- ANALYZERS_AND_RECONCILERS.md ? Design (owner-gated, NOT started): a **signal-fusion framework** for rally detection ? producers emit confidence-scored **signals** at three temporal **scopes** (frame detectors / window analyzers / session analyzers), a learned **scorer** weights them, and an auditable **report-first reconciliation** pass lints the decision against the evidence. The spine: **no special-casing in the decision path** (signals?scorer, never `if/else`). Worked examples: a service-fault analyzer, a warm-up/practice span detector, and consuming a per-frame shuttle-state detector. A peer to `EXPERIMENT_HARNESS.md` and `MULTI_SIGNAL_FUSION_PLAN.md` (ADR-007 modular rally layer).
- PERSONALIZATION_PLAN.md ? Design (owner-adopted **hybrid**; Phase-0 landing) for the **per-user personalization feature on a generalization-first baseline**: a thin per-user *tuning* layer gated by the honest small-N statistics (a **min-N promotion floor** + a **structural-guard exemption**, never a service gate) + a **contribution flywheel** (free-tier videos pool into the owner-trained baseline; *"a tool that gets better the more you help it"*). Records the **locked owner decisions** (identity, storage, `min_n` semantics, Gate-A default, free-tier pooling). Built so far: per-user promotion gate (#141) + held-out-USER learning curve (#142).

- GOLDEN_REGRESSION_FIXTURES.md ? **Design (owner-gated, NOT started):** video-free, non-attributable regression fixtures so a clean clone runs the rally-seg suite with **no video/GPU/DB**. Two layers ? the post-detector freeze + **dual gates** (characterization + quality-floor) reusing the `backend/eval/` stack, and the **SEALED-FIXTURES** privacy filter (de-attribution-at-source, tiered public/key-gated, shuttle-only public tier). Carries a cited **IN/US/EU legal memo** (derived-data IP/privacy), a threat model, and a **§7 progress tracker + definition-of-done**. Extends GOLDEN_DATA_SHARING.md.

Product / platform plans

- DESKTOP_APP_PLAN.md ? **Tracked scoping doc (2026-06-18, owner-gated, `[design]` ? NOTHING built):** package the local pipeline as a **cross-platform desktop app** for non-technical players ? plug in the camera's USB/SD card, click once, get back **only the highlight reels**, fully on-device (no upload, no original-video bloat). The centerpiece is the **de-WSL native-compute port** (run WASB natively per-OS: Linux/Windows CUDA · macOS MPS/CPU ? drop WSL2), and its **make-or-break unknown**: is CPU acceptable on a typical *no-GPU* enthusiast laptop (WASB is ~3.4× real-time on a laptop 2070S)? §7 tracker = **P0 compute-feasibility spike ? P1 native detector ? P2 app shell ? P3 packaging/installers**. Realizes the **L0 fully-local** tier of VIDEO_LOCALITY_MODEL.md; is the **`LocalDesktop` ComputeBackend** of PLATFORM_ARCHITECTURE.md §3.2; the smaller/export-clean owned model (OWNED_MODEL_IMPLEMENTATION_PLAN.md) is one answer to the compute crux.

Practical / ongoing

- TRACKNET_WSL_SETUP.md
- Golden set docs (`GOLDEN_SET_*`)
- Evaluation harness and calibration drivers (in `backend/eval/`)
- GOLDEN_DATA_SHARING.md ? How the GPU box and CPU dev/agent sessions share golden artifacts (trajectory CSVs, `*_golden_features.json`) via a OneDrive folder; the GPU-extract ? CPU-litmus split. See **#175**.

```
23      ## Conventions & templates
24      - [PROJECT_DOC_TEMPLATE.md](PROJECT_DOC_TEMPLATE.md) ? Reusable skeleton for a **tracked
project doc**: status header · decisions-locked · progress tracker (?/?/?) · definition-of-done
· archived on completion. Standardizes the emergent pattern in `RALLY_QUALITY_RESEARCH.md` §7
and `GOLDEN_REGRESSION_FIXTURES.md`. Use for any substantial design/project work.
25
26      ## Historical documents
27      Completed plans, old research, ADRs, and checkpoints live in [`archives/`](archives/).
28
29      - [CODE_AUDIT_AND_TEST_HARDENING.md](CODE_AUDIT_AND_TEST_HARDENING.md) ? **COMPLETED**
(2026-06-14; T1?T5 + full loop + verification per #162). Living record of the June 2026 plumbing
/ contract / architectural audit, 626-test infrastructure, positive test plans for the 5
identified issues, multi-layer regression protocol (including ablation/experiment layer), and
pre/post-change verification checklist. See `docs/archives/CODE_AUDIT_AND_TEST_HARDENING-
COMPLETED-2026-06-14.md` and `docs/archives/checkpoints/2026-06-14-hardening-loop-complete.md`
for the final archived state. (Top-level is now pointer stub per review; full content in
archives/.)
30
31      See `archives/README.md` for the umbrella policy and `archives/past_projects/README.md`
for classification of small tech debt / starter projects.
```

81. user (2026-06-19T17:15:29.699Z)

```
=== status header lines (first status-ish line per doc) ===
ABLATION_LAYER.md       |       (no status header)
ANALYZERS_AND_RECONCILERS.md    |       **Status:** **DESIGN ? owner-gated, NOT started.** A
peer to
BACKLOG.md      |       (no status header)
CODE_AUDIT_AND_TEST_HARDENING.md        |       Status: COMPLETED** (2026-06-14; all T1?T5
positive tests + verification + loop complete per #162/#163; archived snapshot in
`docs/archives/`)
CODE_MAP.md     |       status:'queued'
COMMERCIALIZATION.md    |       **Status:** ? LIVING ? track this in parallel with
product/quality work.
DEFERRED_HARDENING_PLAN.md      |       STATUS: COMPLETED** (2026-06-14). This planning doc's
three items are all dispatched:
DESKTOP_APP_PLAN.md     |       > **Status:** `DRAFT` ? `owner-gated` (scoping doc only;
**nothing is built**) · **Owner:** avidullu · **Created:** 2026-06-18 · **Last updated:**
2026-06-18
EVALUATION.md   |       (no status header)
EXECUTION_POLICY.md     |       **Status:** **P1 SHIPPED** (the actual hang fix); P2 + P3a
substrate landed; P3b next. Motivated by a
EXPERIMENT_HARNESS.md   |       **Status:** **P1 IMPLEMENTED** (`backend/eval/experiment.py` +
`tests/test_experiment.py`, shipped
FIELD_VISIT_ANSWER_SHEET.md     |       (no status header)
FIELD_VISIT_PLAYBOOK.md |       (no status header)
GOLDEN_DATA_SHARING.md  |       **Status:** convention + plan (2026-06-15). Tracked in **#175**.
GOLDEN_REGRESSION_FIXTURES.md   |       > **Status:** `IN PROGRESS` ? owner-gated · **Owner:**
Avi · **Created:** 2026-06-16 · **Last updated:** 2026-06-16 · **Legal:** owner **accepted the
researched risk** 2026-06-16 (conditions §3 stand; this is the owner's informed acceptance,
**not** a retained-counsel opinion)
GOLDEN_SET_IMPLEMENTATION.md    |       **Status:** In progress · **Date:** 2026-05-30
GOLDEN_SET_PHASE1_VIDEOS.md     |       **Status:** Design v1 · **Date:** 2026-06-02
GOLDEN_SET_VENDORING.md |       **Status:** Design v1 · **Date:** 2026-05-30 · **Vendor
shortlist added 2026-05-30 (§6)**
GOLDEN_VIDEOS.md        |       (no status header)
HARDENING_LOOP_HANDOFF.md       |       (no status header)
HOSTING_PLAN.md |       **Status:** PROPOSAL ? awaiting owner ratification.
HOW_RALLY_DETECTION_WORKS.md    |       (no status header)
I18N_PLAN.md    |       **Status:** Design (P0 deliverable; phases P1+ are owner-gated, not
started).
KHELSUTRA_PRODUCT_OVERVIEW.md   |       (no status header)
MULTI_SIGNAL_FUSION_PLAN.md     |       **Status:** **P1 DONE** (2026-06-13) ? the person cue is
extracted to
NEXT_STEPS.md   |       (no status header)
```

OWNED_MODEL_IMPLEMENTATION_PLAN.md | **Status:** PLAN / roadmap. No owned-model *implementation* starts until the owner greenlights a specific
PERSONALIZATION_PLAN.md | Status: DESIGN ? owner-adopted (hybrid); Phase-0 engineering landing.** This is the authoritative
PER_VIDEO_OUTPUT_LAYOUT.md | **Status:** **DESIGN ? owner-gated, NOT started.** Proposed convention so processing N videos
PLATFORM_ARCHITECTURE.md | **Status:** ? Strategy / research ? for review. No code yet.
PMF_FIELD_SIGNALS.md | **Status:** ACTIVE ? the decision layer for the C13 beachhead validation
PMF_MARKET_RESEARCH.md | (no status header)
POST_166_WORK_ITEMS_MILESTONES_AND_RISK_ASSESSMENT.md | **Status:** Plan for owner review first. Execution only after explicit owner approval ("review first and then execute"; "I want a plan detail as to what are the next items planned, please execute only after I explicitly say so").
PRODUCT_FIELD_ASSOCIATE_FLYER.md | (no status header)
PROJECT_DOC_TEMPLATE.md | > **Status:** `DRAFT` ? `<owner-gated?
PROMINENT_COURT_DETECTION.md | > **Status:** `DRAFT` ? design + experiment **owner-gated for any default flip** · **Owner:** Avi · **Created:** 2026-06-19 · **Last updated:** 2026-06-19
QUALITY_ITERATIONS.md | Status: retained default-OFF**, justified by the real-footage fix + the foundational argument that
R1_MILESTONE_LAUNCH_REPORT.md | **Status:** LAUNCHED 2026-06-18 (PR #204). **Flip:** the local-windowing defaults go ON ?
R2_EVAL_METRIC_DESIGN.md | **Status:** **IMPLEMENTED (augmenting), 2026-06-17** ? owner-reviewed + decisions locked 2026-06-17;
RALLY_DETECTION_GAP_CLOSURE.md | > **Status:** `DRAFT` ? implementation **owner-gated** · **Owner:** Avi · **Created:** 2026-06-18 · **Last updated:** 2026-06-18
RALLY_DETECTION_QUALITY_REPORT.md | (no status header)
RALLY_QUALITY_RESEARCH.md | **Status:** Foundation doc (2026-06-16), run as a **tracked project** (§7) ? *"Bridge the
README.md | (no status header)
REAL_FOOTAGE_VALIDATION_REPORT.md | **Status:** validation complete, roadmap proposed (all flags default-OFF)
ROADMAP.md | (no status header)
ROORA_LABELING_GUIDELINES.md | (no status header)
STORAGE_SHARING_MODEL.md | **Status:** DRAFT ? owner + Claude co-design session pending; nothing here is ratified.
TRACKNET_WSL_SETUP.md | (no status header)
UI_ALPHA_EXECUTION_PLAN.md | **Status:** ACTIVE · updated 2026-06-10 · **Linked:** `UI_PROPOSAL.md` · `HOSTING_PLAN.md` ·
UI_PROPOSAL.md | **Status:** PROPOSAL ? awaiting owner ratification (no implementation starts without explicit approval, per CLAUDE.md).
VIDEO_LOCALITY_MODEL.md | **Status:** THINKING DOC (owner-requested, 2026-06-12) ? a shared mental model for
VIDEO_RECORDING_GUIDELINES.md | **Status:** ? Draft / placeholder · **Date:** 2026-05-30

82. assistant (2026-06-19T17:16:50.025Z)

This gives me a strong first pass. The status headers already separate the obvious **LIVING/ACTIVE/DRAFT-current** docs (keep) from terminal ones, but several need care ? e.g. `DEFERRED_HARDENING_PLAN.md` says "COMPLETED" yet CLAUDE.md cites it as the live source for the *next* platform slice (#16). Moving a doc also breaks every relative link to it, and the owner just ran two doc-staleness sweeps ? so each archive candidate needs an evidence check + inbound-reference check, not just a status-header glance.

Let me read the archive policy itself so I classify by the repo's own rules.

83. user (2026-06-19T17:16:52.618Z)

########## docs/archives/README.md ##########

archives/

Umbrella directory for all **historical and completed** material.

Live vs Historical

- **Live docs** (for ramp-up, constant updates, and current work) live at the top level of `docs/`.
- Everything here is no longer the primary source of truth. It is kept for context, decision history, and regression detection.

Classification inside archives/

- `past_projects/` ? Completed small tech debt, starter projects, or low-regret initiatives (e.g. the original low-regret moves + scaffolding work). These are "past" once they reach a terminal state.
- `decisions/` ? Historical ADRs and major architectural decisions.
- `research/` ? Completed or superseded research notes and plans.
- `checkpoints/` ? Snapshots of major project states (especially useful for detecting regressions in quality, priorities, or baselines over time as we do quality iterations).
- Other one-off audits, etc.

Archival Rule

Only move something into `archives/` (or its subdirectories) when the work is in a clear **terminal state**:
- DONE
- COMPLETED
- REJECTED

Never archive in-flight or pending work.

When archiving a project/phase:
1. Ensure the main document has an explicit `**STATUS: COMPLETED**` (or DONE/REJECTED) header at the top.
2. Link back to the current living documents in `../` (e.g. `COMMERCIALIZATION.md`, `BACKLOG.md`).
3. For regression detection value, consider adding a dated entry under `checkpoints/`.

See `past_projects/README.md` for the specific guidance on that category.

########## docs/archives/past_projects/README.md ##########

past_projects/

This subdirectory is for **completed small tech debt, starter projects, or low-regret initiatives**.

Criteria for past_projects/

- The work was a bounded "project" (e.g. low-regret moves, scaffolding maturation, specific tech debt cleanup).
- It has reached a terminal state (STATUS: DONE / COMPLETED / REJECTED).
- It is no longer the active focus (new work happens via the live docs + BACKLOG).

Example

- `low-regret-moves/LOW_REGRET_MOVES_PLAN.md` ? the 2026 low-regret + scaffolding phase. Marked COMPLETED after the three scaffolding PRs were merged.

Future small completed starter projects or tech debt cleanups that are similar in nature should land here (with their own subdir and clear STATUS header).

Larger ongoing quality iterations or major feature work will typically use dated entries under `../checkpoints/` or `../quality-iterations/` instead.

########## existing archived files (what's already there) ##########
docs/archives/CODE_AUDIT_AND_TEST_HARDENING-COMPLETED-2026-06-14.md
docs/archives/COMMERCIAL_READINESS_REVIEW.md
docs/archives/DEFERRED_HARDENING_PLAN-COMPLETED-2026-06-14.md
docs/archives/HARDENING_LOOP_HANDOFF-COMPLETED-2026-06-14.md
docs/archives/MONETIZATION_AUDIT.md

```
docs/archives/README.md
docs/archives/TEST_RUN_SUMMARY.md
docs/archives/checkpoints/2026-06-14-hardening-loop-complete.md
docs/archives/decisions/DECISIONS.md
docs/archives/past_projects/README.md
docs/archives/quality-iterations/README.md
docs/archives/research/AUDIO_E1_FINDINGS.md
docs/archives/research/AUDIO_RALLY_DETECTION_PLAN.md
docs/archives/research/OWNED_MODEL_TRAINING_STUDY.md
docs/archives/research/RALLY_DETECTION_RESEARCH.md
```

84. assistant (2026-06-19T17:17:30.536Z)

The policy is now unambiguous: **archive only terminal-state (DONE/COMPLETED/REJECTED); never in-flight work.** And a key finding already ? three docs (`CODE_AUDIT`, `DEFERRED_HARDENING_PLAN`, `HARDENING_LOOP_HANDOFF`) **already have archived `-COMPLETED-` copies**, so their top-level versions should be stubs/duplicates, not fresh moves. I'll fan out a classification pass over all 52 docs ? each with evidence, a duplicate-check, a "cited-as-live" cross-check, and inbound-link risk.

85. user (2026-06-19T17:18:10.887Z)

Workflow launched in background. Task ID: wvgslr8mh
Summary: Classify every docs/*.md by lifecycle state against the repo archive policy (terminal-only), flag archive-now candidates with destination + duplicate-check + cited-as-live cross-check + inbound-link risk
Transcript dir: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\subagents\workflows\wf_4e32c8b8-725
Script file: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\docs-archive-classification-wf_4e32c8b8-725.js
(Edit this file with Write/Edit and re-invoke Workflow with {scriptPath: "C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\docs-archive-classification-wf_4e32c8b8-725.js"} to iterate without resending the script.)
Run ID: wf_4e32c8b8-725
To resume after editing the script: Workflow({scriptPath: "C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\docs-archive-classification-wf_4e32c8b8-725.js", resumeFromRunId: "wf_4e32c8b8-725"}) ? completed agents return cached results.

You will be notified when it completes. Use /workflows to watch live progress.

86. assistant (2026-06-19T17:18:16.991Z)

While the classification runs, let me independently verify the highest-value finding: whether the three docs that already have archived `-COMPLETED-` copies are thin stubs (fine) or full duplicates (cleanup needed).

87. user (2026-06-19T17:18:20.267Z)

```
=== CODE_AUDIT_AND_TEST_HARDENING ===
  top-level: 13 lines
  archived : 493 lines (docs/archives/CODE_AUDIT_AND_TEST_HARDENING-COMPLETED-2026-06-14.md)
=== DEFERRED_HARDENING_PLAN ===
  top-level: 11 lines
  archived : 130 lines (docs/archives/DEFERRED_HARDENING_PLAN-COMPLETED-2026-06-14.md)
=== HARDENING_LOOP_HANDOFF ===
  top-level: 15 lines
  archived : 259 lines (docs/archives/HARDENING_LOOP_HANDOFF-COMPLETED-2026-06-14.md)
```

88. assistant (2026-06-19T17:18:29.439Z)

All three are already thin **pointer stubs** (11?15 lines) pointing to their full archived copies ? so they're correctly handled, not duplicates to clean up. That removes them from the candidate list. The classification workflow is sweeping all 52 docs in parallel; I'll synthesize the move-list once it returns.

89. user (2026-06-19T17:18:31.551Z)

There is too much clutter and I want to only maintain active items and want to get a prioritized list of pending items (UI or desktop application, Setting up a domain now that i have bought khelsutra.guru for 3 years! and its a very dumb page so I want to add some good html there etc.)

90. assistant (2026-06-19T17:19:19.981Z)

The reframe makes sense ? and the running classification gives us the "active vs archive" split for the decluttering half. But your real ask is a **prioritized pending-items list** that includes new product work (the domain + landing page), not just doc hygiene. Let me read the owner's own pickup doc and the product overview (the landing-page source material) to ground it.

91. user (2026-06-19T17:19:20.878Z)

```
1       # Next steps (live) ? owned-model + multisport
2
3       **Updated:** 2026-06-13 (**rally-detection quality track**: fusion + experiment-harness
design docs merged,
4       owner-gated ? see the dedicated section below, the desktop-agent pick-up). Prior:
2026-06-10 (audit?M0?Gemini-battery session; later same day the **UI-alpha track opened**:
5       ADR-009 ratified ? KhelSutra Guru / khelsutra.guru, private `khelsutra` repo, local-
first delivery ? see
6       `docs/UI_ALPHA_EXECUTION_PLAN.md`; **PR-1 async jobs (#16) MERGED 2026-06-11, PR #87**;
**PR-2
7       upload/download MERGED 2026-06-11, PR #99** ? live-E2E-verified (3-part chunked upload
MD5 byte-identity +
8       compile?browser-download); next: PR-3 identity. Same day: the **8?9 quality sweep
#90?#98 landed** ? ruff/mypy/
9       coverage-floor CI lanes, SQLite conn-close fix, config unknown-key warning, hybrid-twin
collapse into
10      `trajectory_hybrid.py`, WSL tilde+abspath live-path fixes; suite 388 green; LOVO 0.626
re-verified exact).
11      **Authoritative roadmap:**
12      `docs/OWNED_MODEL_IMPLEMENTATION_PLAN.md` (+ ADR-008 for topology). **Latest blow-by-
blow:** memory
13      `current-state.md` / `session-handoff.md`. This file = the prioritized "what to do
next," refreshed each session.
14
15      ## ? Next-session / GPU-box pickup (2026-06-15)
16      **M0 in-container quick wins are COMPLETE:** ship-gate provisional blocker (**#170**),
train/eval split guard (**#171**), the **doc-consistency sweep** (**#174** ? corrected 16
drifted docs vs the code), and the **P3 person-feature fusion litmus machinery is already
present + suite-green** (`backend/eval/fusion_compare.py` + `ablation.py`: person-driven ?F1 /
bootstrap CI / paired sign test; the real-golden `skipif` litmus waits only on GPU-extracted
data). ? CI is red repo-wide due to a **GitHub Actions minutes/billing outage** (account-level,
not code) ? work is local-gated + `--admin`-merged until Actions is restored (Settings ? Billing
? Actions minutes / spending limit).
17
18      **Immediate next step ? the real P3 LOVO measurement** (this box has a GPU +
`torch+cu124` + `cv2`; videos and trajectory CSVs are present, but the GPU-extracted golden
features + `.rallies.csv` labels are not yet here):
19      1. **[owner / GPU box]** `python -m backend.eval.fusion_golden --manifest
output/human_lovo_manifest.json --out-dir "%USERPROFILE%\OneDrive\rally-annotator\golden-
data\features\2026-06-15-n6"` ? produces `*_golden_features.json` for the 6 golden videos. (GPU
+ WSL + videos + trajectory CSVs + `.rallies.csv` labels are owner-side only.)
20      2. Artifacts land in the **shared cloud folder** `C:\Users\avidu\OneDrive\rally-
annotator\golden-data\` (OneDrive auto-syncs both ways). Convention + workflow:
**[`GOLDEN_DATA_SHARING.md`](GOLDEN_DATA_SHARING.md)**; tracked in **#175**.
```

3. **[CPU dev/agent session]** `python -m backend.eval.fusion_compare --features-dir <shared>\features\2026-06-15-n6 --record` + `python -m backend.eval.ablation --features-dir <shared>\... --granularity group` ? real person-feature ?F1; copy the JSONs into `output/` to also pass the `skipif` litmus tests (`tests/test_ablation.py::test_litmus_reproduces_gate_a`).

Also queued (owner-gated / pending input):
- Ship-gate target calibration ? **#172** (owner decision; needs corpus growth).
- Tests reorganization ? blocked on the owner's grouping metric (provide it ? an agent will do it).
- Restore CI ? the Actions outage is account-level (billing/minutes), not a repo/config bug.

Leaving (bulky / M2 ? no start without owner go): ActionFormer (M0.4), event-index DB migration (M0.1), Gemini-silver purge (M0.3), cost-to-Gen-2 benchmark + ByteTrack deprecation, multisport, repo rename, PMF execution.

Where we are (one paragraph)

ADR-008 ratified ("repo split driven by productionization, not isolation"): training lives in-repo (`training/`) behind a CI-enforced import wall; the featurizer is single-sourced (`backend/features/rally_seq.py`, train==serve); the **Gen-0 harness ships** (`python -m training.gen0.harness --manifest output/human_lovo_manifest.json --floor eval_baselines/heuristic_lovo_n6.json --version <semver>`) and produced artifact v0.1.0 (LOVO **0.626**, floor passed, sign test 5/6 wins p=0.109 ? honestly not significant, ship=False). The **n=6 owned corpus is in the collector** (262 rallies, Proprietary, commercial export = owned data only after the ShuttleSet reclassification). The **Gemini battery is done** (see table) ? the moat is quantified.

The quality table that now drives strategy (IoU 0.5 vs human golden labels)

Path	Mahadevpura (clean-ish)	Kushagra (multi-court)
Heuristic (velocity windowing)	0.657	0.157
Gen-0 owned head (LOVO, n=6)	0.744	0.561
Two-stage WASB+Gemini refine	0.83	?
Gemini wrapper (`gemini-flash-latest`)	**0.93**	**0.66**

Reading: clean footage is commoditized (serve it with Gemini for cents). **Multi-court drops the commodity wrapper to 0.66 ? that is the owned model's ship target** (its FPs are background-game pickups + dead-time merges, the exact contrast-metric confound). Gen-0 is 0.10 behind with only n=6 ? a corpus-growth-sized gap, not an architecture-sized one. Two-stage refine = the cost-efficient serving path for LONG tapes (uploads candidate clips only); wrapper 503-flakiness (observed all session) makes the provider-fallback registry load-bearing. Baselines tracked in `eval_baselines/` (heuristic_lovo_n6.json, gemini_wrapper_2026-06-10.json).

Rally-detection quality track (design merged 2026-06-13, owner-gated; implementation in progress)

This is the track a desktop/local agent is slated to pick up. Design docs landed (PR #112 + session). All implementation is **owner-gated** (per explicit in the plan); each step = ONE PR off `master`.

Hardening loop (T1-T5 positive tests + audit + finale) COMPLETE (2026-06-14, PRs #156/#157/#160/#161/#163 + #165 + final records). See archived `docs/archives/CODE_AUDIT_AND_TEST_HARDENING-COMPLETED-2026-06-14.md` (Living Log with every pre/post/review) and `docs/archives/HARDENING_LOOP_HANDOFF-COMPLETED-2026-06-14.md` (full rules + STATE). Top-level now pointer stubs per #165 review. All 5 items + verification + archive done; clean slate.

- **Docs:** `docs/MULTI_SIGNAL_FUSION_PLAN.md` (shuttle + person + audio, one fusion layer) ·
  `docs/EXPERIMENT_HARNESS.md` (A/B + shadow + promotion gate, zero-regression) ·
  `docs/ROORA_LABELING_GUIDELINES.md` (v8) · `docs/HOW_RALLY_DETECTION_WORKS.md` (2-layer mental model).
- **Key finding (no re-investigation needed):** the owner's "held-shuttle counted as a rally" bug is **NOT a missing capability** ? `backend/pipeline/detectors/rally_gate.py` already implements the net-crossing **"Gate A"** (`apply_gate(..., min_crossings=2)`) that kills service-feed/held-shuttle windows, but it is **only called from eval drivers** (and there was no `min_crossings` knob in `IndexingConfig`). **Wiring Gate A into the live segmenter (now via fusion_hybrid) was the single cheapest, highest-confidence quality win.** (P2 FusionSegmenter specifics: registry `fusion_hybrid`, **default-OFF**, seeds from substrate, persists `net_crossings` + optional person features, optional served Gate A/B via config knobs; adversarially reviewed; suite green. See `tests/test_served_gate_a_regression.py` for the honest finding that on production windowing Gate A is neutral/negative in some cases, hence kept opt-in with knob for safety, and the leaderboard revert note.)
- **Current status (2026-06-14):** P1 (person cue extraction to reusable `PersonDetector`) + P2 (FusionSegmenter + Gate A signal + served handoff gating) + experiment harness P1 DONE (adversarial review + NO-REGRESSION; suite green). **P3 next (this PR + follow-ups):** learned fusion (person features ? harness `compare` LOVO lift on the 6 golden videos; eager-person-compute optimisation); then P4 audio via hit_detector.
- **Pick-up order (each = ONE PR off `master`, owner approval first, flag-gated default-OFF, promote only on measured LOVO):**
  1. (P2 first-step, largely done via fusion_hybrid) Wire Gate A into production served path + config knob (min_crossings); apply in hybrids at runtime. (Now available opt-in; served neutral on production windowing per litmus ? kept opt-in with knob for safety.)
  2. P3 first step (this work item): harness `compare` LOVO litmus for person-on variant (players_in_court + two_sided + colocation etc. from `fusion_features`) on the 6 golden; add the eager compute optimisation note/landing. Pure + harness (testable here).
  3. P3/P4 continuation: audio cue integration + full learned promotion if lift proven.
- **Discipline (from EXPERIMENT_HARNESS.md):** every new cue **flag-gated, default OFF**; promote to default **only on measured LOVO lift** through `run_regression.py` (exit 0/1/3); pinned `CONTROL` ? rollback = config flip. Harness re-uses ~80% existing code.
- ? Gate A, person cue logic, harness, and pure fusion paths are testable here; the heavier real-video / golden LOVO / ablation confirmation waits on the GPU-extracted golden features + `.rallies.csv` labels, which are not yet present in this checkout.
- **External due-diligence review inputs (2026-06-17) ? PICK UP AFTER current ideas are exhausted; full details + citations in [`RALLY_QUALITY_RESEARCH.md`](RALLY_QUALITY_RESEARCH.md) §5.6.** An owner-commissioned 31-reference review of the quality report *validated* the methodology and surfaced these **future** levers (research inputs, not active work, not new claims; verify citations before any submission):
  1. **"First-mile" paper framing (highest-value):** stroke-classifiers (BST/DSTA) + the big datasets (ShuttleSet ~36k strokes, Fine-Badminton ~27k actions) all **assume a pre-trimmed rally clip already exists** ? we solve the ignored prerequisite. Pitch the paper as *"the missing untrimmed-video preprocessing pipeline that lets BST/MonoTrack run on raw footage."*
  2. **Ground R2 in Action-Spotting / Precise-Event-Spotting** (SoccerNet/TenniSet/T-DEED) ? turns R2 into a defensible publishable contribution.
  3. **Add domain baselines for any submission:** MonoTrack + BST on ShuttleSet / Fine-Badminton (we have only heuristic + Gemini today).
  4. **Prefer TemporalMaxer over ActionFormer** for the M0.4 scorer swap (parameter-free max-pool, ~2.8× fewer GMACs, small-N robust) ? see "M0.4 next increment" below.
  5. **Court polygons ? pose ? the rally-END lever:** masking on-court players can resolve *occluded end-of-rally states* (the #1 remaining gap vs Gemini), not just spectator noise.
  6. **Name the eval?serve lesson "pipeline distribution skew"** ? a citable paper "lesson."

Discipline reminder: owner approval before starting any owner-gated item; one PR per step; babysit the PR (poll/fix/reply) until merge observed, *then* advance.

Prioritized next steps

1. **[owner, in progress] GROW THE GOLDEN CORPUS** ? the single highest-leverage move; every lever below feeds on it.
   Priority order for new videos: **(a) multi-court badminton** (the target distribution AND where the ship target

lives), **(b) ?3 matches per new sport ? table tennis first** (WASB transfers at F1 0.909; pickleball labels are
corpus-gold but not trainable until a clean MIT/Apache ball detector exists), (c) capture variety per
`VIDEO_RECORDING_GUIDELINES.md`; **adult sessions only for the training pool** (consent guardrail). Per video:
label ? `curate import-local --normalize --license Proprietary --fps <true fps>` ? re-run the Gen-0 harness.
The paired sign test needs n: at n=10 with 9 wins, p?0.011 ? significance is reachable this month.
2. **[owner decision] Set the ship-gate target** ? now informed: floor = heuristic 0.480 (necessary), **multi-court
reference = wrapper 0.66** (the number that makes the owned model worth serving), clean-footage ceiling = 0.93
(not the target ? Gemini serves that tier). Suggested shape: "LOVO F1@tIoU0.5 ? 0.70 on multi-court folds with
paired-sign p<0.05 vs heuristic" ? owner to ratify/adjust.
3. **M0.4 next increment ($0/local):** swap the Gen-0 LogReg for a lightweight temporal head (MIT, minutes on the
2070S ? NOT the heavyweight Gen-2 stack) inside the existing harness; per-video threshold calibration (the
LOVO-vs-oracle gap is visible per fold in the harness output). **Prefer TemporalMaxer** (parameter-free
max-pool, ~2.8× fewer GMACs, small-N robust) over ActionFormer/ASFormer per the 2026-06-17 external review
(`RALLY_QUALITY_RESEARCH.md` §5.6) ? better fit for the cheap/local/fast mandate.
4. **Multisport thread:** ingest Roora pickleball/TT CSVs (`import-local --normalize`); merged-prototype LOVO across
badminton+TT once TT has ?3 matches; identify a clean pickleball ball detector. ? TT-on-WASB serving stays gated
by C3 (provenance multiplies per sport).
5. **PMF validation (cheaper than any engineering month):** C13 academy interviews (+ the two added questions:
"multi-court footage nobody watches?" / "pay per-match to index it?"); camera pilot at the warmest 2-3 academies
(consent at capture; demo served via the Gemini path or human-polish ? the reel pattern exists:
`output/MahadevpuraSingles_highlights.mp4`). **Field kit COMPLETE (2026-06-12, owner ratified the
physical-visit approach; reframed product-feedback-first per owner spec):**
`docs/PRODUCT_FIELD_ASSOCIATE_FLYER.md` (the recruiting flyer ? owner fills ?/contact placeholders and recruits) +
`docs/FIELD_VISIT_PLAYBOOK.md` (the hired associate's visit playbook: coach + student script, demo-after-Q6
rule, ground rules, Phase-2 record-and-show-back stretch goal) + `docs/FIELD_VISIT_ANSWER_SHEET.md` (per-visit
capture form incl. demo-reaction + student blocks) + `docs/PMF_FIELD_SIGNALS.md` (G/A/R anchors + **pre-registered
decision rules** ? PROCEED / PIVOT A-C / FALSIFIED thresholds fixed before the first visit).
6. **Housekeeping:** ~~rotate the Gemini API key~~ **DONE 2026-06-10** (owner-verified working; confirm old-key
deletion in AI Studio); **rename repo off "badminton"** (owner action,
in BACKLOG); collector=SoR confirmation; off-peak niceties: none pending ? the battery completed.

Pointers / artifacts

- Roadmap: `docs/OWNED_MODEL_IMPLEMENTATION_PLAN.md` · Topology: ADR-008 (`docs/archives/decisions/DECISIONS.md`) ·
  Study: `docs/archives/research/OWNED_MODEL_TRAINING_STUDY.md` · Quality history:
  `docs/archives/quality-iterations/2026-06-multivideo-lovo/`.
- Harness: `training/gen0/harness.py` (one command ? train ? LOVO ? gate ? artifact bundle; weights gitignored,
  manifest reviewable). Featurizer: `backend/features/rally_seq.py` (FEATURE_SCHEMA v1).
- Corpus: collector DB (system-of-record, 6 matches/262 rallies) · eval manifest

```
`output/human_lovo_manifest.json`
121       (gitignored) · durable trajectories + Roora files: `C:\Users\avidu\Projects\Annotation
Setup\Collect\Trajectories\` ·
122       labels: `...\Collect\Golden Labelled\`.
123     - Gemini ops learnings: serial uploads (`indexing.ai_max_workers=1`) + chunk re-encode
124       (`indexing.ai_chunk_scale_width=1280`) + 6-attempt exponential generate retry ? a
cold/prepaid project burst-throttles
125       with a MISLEADING "credits depleted" 429; ramp gradually.
126     - ? No owned-model implementation beyond greenlit M0.x until owner greenlights the next
milestone. Committed next
127       PLATFORM slice = async job model (single-tenant).
128
```

92. user (2026-06-19T17:19:21.262Z)

```
1       # KhelSutra ? Product Overview & Early-User Guide
2
3       *What it does today, how to get the best out of it, and where it's going.*
4       *For field associates and for invited alpha / beta academies, coaches, and players.*
5
6       ---
7
8       ## In one line
9
10      **KhelSutra turns ordinary badminton video into useful pieces ? automatically.** Put a
phone or GoPro
11      on a tripod behind the court, record a session, and KhelSutra automatically finds the
rallies, stitches
12      a **highlight reel**, and gives you a **simple summary** (how many rallies, how much
real play time).
13      No special hardware, no editing, no manual scrubbing.
14
15      ---
16
17      ## The problem we're solving
18
19      Academies and serious players already record a lot of video ? and then almost never
watch it. Hours of
20      footage pile up on phones and memory cards because cutting it into useful clips by hand
is tedious and
21      nobody has the time. The good rallies, the moments worth reviewing, stay buried.
22
23      Most academy halls also run **several courts at once**, so a single recording is a busy,
noisy scene
24      that ordinary video tools choke on. **Pulling the real rallies out of busy, multi-court
footage is the
25      hard problem KhelSutra is built to crack** ? it already works very well on clean,
single-court
26      recordings, and the multi-court case is exactly what we're sharpening now.
27
28      ---
29
30      ## Who it's for
31
32      - **Coaching academies & coaches** ? turn raw session footage into review-ready clips
and reels for
33        teaching, without an editor.
34      - **Serious players** ? get clips of your own rallies to review, keep, and share.
35      - Today's focus is **badminton academies in Bengaluru** ? that's who we're building with
first.
36
37      ---
38
39      ## What it does **today**
40
```

This is what's working right now, in the private web app invited users can try:

- **Upload a full session or match video** (from a phone on a tripod, or a GoPro). Long videos are fine.
- **Automatic rally detection** ? it finds the rallies and where they start and end, on its own.
- **Highlight reel** ? the rallies stitched into one MP4 you can watch, download, and share, with a small
  "Generated by Khelsutra.guru" mark.
- **A simple summary** ? how many rallies were found and how much actual play time there was.
- **Re-use & download** ? process once, come back to it later, download the reel.

That's the core loop: **record ? upload ? get back a reel + summary.** No editing skill needed.

> Like any automatic tool, it can occasionally miss a real rally or include one that isn't ? telling us
> when that happens is genuinely useful feedback, not a sign it's broken (more on that below).

How to use it (alpha)

1. **Your KhelSutra associate helps you get access** to the private web app (it's invite-only; today the
   app is in English).
2. **Upload your session video.** Large files upload in pieces and can take a while on Indian uplinks ?
   that's normal.
3. **It processes in the background.** You can close the tab and come back later; you don't have to sit
   and watch it.
4. **When it's done**, you get the **highlight reel + summary** to watch and download.
5. **Re-process anytime.** (Re-running the same video replaces its earlier result rather than keeping both.)

How to get the best out of it (capture matters)

KhelSutra works best on a **steady, full-court recording** ? the same way a coach would set a camera to
review play. A few simple habits make a big difference to the results:

- **Put the camera behind the court, slightly elevated**, with the **whole court in frame**.
- **Keep it still** ? a fixed tripod beats a handheld or panning shot every time.
- **Shoot landscape**, with reasonable lighting.
- **Record the whole session** and upload it as-is ? KhelSutra does the cutting.

Footage that fights these habits (shaky handheld, heavy zooming/panning, broadcast-style cuts, very dark
halls) is the **hardest case** and where results suffer most. When in doubt: fixed camera, full court,
don't touch it. *(Your field contact provides a one-page capture checklist at setup.)*

The part that makes KhelSutra different: **it gets sharper from real academy footage**

This is the heart of the product, and what sets it apart from a generic "AI clipper":

The footage that real academies contribute is what makes KhelSutra better over time ? especially on the busy multi-court halls that are the hard case. When you choose to share your footage, and when coaches confirm or correct what it found, that growing, labelled library is what we use to sharpen the tool. It's designed to improve in new versions as that library grows, rather than stay frozen ? which is why early users aren't just customers, they're co-builders.

Your data stays yours ? and sharing is your choice. We will **never** use your video to improve the tool without your **clear, explicit agreement** ? it is opt-in, never automatic. **Footage involving minors is never used to improve the tool and is never published.** We never publish your footage; you decide what, if anything, you share.

How your video is handled (plainly): these are large videos that need real processing, so your video is handled securely on our infrastructure to produce your reel. We don't publish it, and we don't use it to improve the tool without your opt-in.

> **Read this honestly:** today the tool gives you rally detection + reels + summaries. It does **not** retrain itself as you use it, and there is no per-person tracking. The "gets sharper over time" above is the model we're building toward, powered by contributed footage ? not a switch that's already on.

Where it's going ? toward an **AI assistant coach**

Today KhelSutra finds rallies and makes reels. The roadmap ? fed by the growing video library above ? turns it into a genuine training companion:

- **Shot detection** ? recognising smashes, drops, net shots, serves, and clears within each rally.
- **Deeper per-rally and per-match analysis** ? patterns, tendencies, what's working and what isn't.
- **Per-match history** ? track progress over weeks and months, not just one clip.
- **Coaching insights** ? an **AI assistant coach** that supports (never replaces) the real coach, surfacing what to work on next.
- **More sports** ? table tennis next, then other racket/court sports.
- **Your language** ? Kannada and Hindi, alongside English.

> **Read this honestly:** the items in this section are the **direction**, not today's buttons. What you can use *today* is rally detection + highlight reels + the play-time summary. The deeper analysis arrives as the video library grows ? which is exactly why your footage and feedback matter so much.

What feedback helps us most (for alpha users)

You don't need to be technical ? just tell us, in your words:

- **When it missed a real rally**, or **grabbed something that wasn't a rally.**
- Whether the **reel is the right length and order** for how you'd actually use it.

- **What numbers or details you wish the summary showed.**
- Your real **capture conditions** ? how many courts were running, camera/phone, lighting.
- Anything that would make this **genuinely useful for your academy** (or a reason it wouldn't be).

(Send feedback however is easiest ? your KhelSutra associate will help you with that.)

What we're focused on over the next ~4 weeks

1. **Growing the library of real academy footage** ? especially busy **multi-court** halls (the hard, valuable case). This is the single biggest lever on quality.
2. **Field conversations with academies** ? understanding how coaches and players actually work with video, and what would make this genuinely useful.
3. **Recording sessions and showing academies their own footage** as reels ? *"this is what it looks like for you."*
4. **Sharpening multi-court accuracy** and inviting a few more academies into the private trial.

What KhelSutra is **not** (so expectations stay honest)

- It's an **early, invite-only trial** ? useful, but still improving; you'll hit rough edges.
- It is **not a line-calling, scoring, or umpiring system**, and **not player tracking / identification**.
- **Shot detection and the deeper "AI coach" analytics are on the roadmap, not available today.**
- It is **not a replacement for a coach** ? it's a tool that saves coaches time and adds insight.
- **Pricing isn't set** ? we're still learning what's fair; today's invited use is about getting it right.

Quick answers to common questions

"How accurate is it?"
Very good on a clean, single-court recording. A busy multi-court hall is harder ? and is exactly what we're improving. The best answer is to *show* a sample reel rather than quote a number.

"Do I need a special camera or hardware?"
No. A phone or GoPro on a tripod, behind the court, whole court in frame. We provide a one-page checklist.

"Is my video private? What happens to it?"
We treat your footage as private: we never publish it, and we never use it to improve the tool without your explicit opt-in. These are large videos that need real processing, so your video is handled securely on our infrastructure to make your reel. **Footage involving minors is never used to improve the tool and is never published.**

"Can it tell me the score / who won / call lines?"
Not today. Today it finds rallies and builds a highlight reel + a play-time summary.

"Does it work for doubles and tournaments?"

```
185     It finds rallies in doubles too. Tournament footage varies: a clean, fixed, full-court
recording works
186     well; crowded multi-court tournament halls are the hard case we're specifically working
on.
187
188     **"Will it get better?"**
189     Yes ? that's the whole idea. The more real footage runs through it (with permission) and
the more coaches
190     correct it, the sharper future versions get, and the sooner shot detection and deeper
analysis arrive.
191
192     **"Which sports?"**
193     Badminton today; table tennis and other racket sports are on the roadmap.
194
195     **"When can I use it / how do I join?"**
196     It's an invite-only private trial. If you're interested, the best next step is to let us
record a session
197     and show you your own court and players as a reel.
198
199     ---
200
201     *KhelSutra · khelsutra.guru · Internal & invited-trial overview, June 2026.*
202     *Field associates: the role description is in the Field Associate flyer, and the
conversation script and
203     visit rules are in the separate Field Visit Playbook.*
204
```

93. assistant (2026-06-19T17:21:25.348Z)

This is a clear pivot toward **product/launch**, and the product overview I just read is excellent landing-page source material. Let me capture the domain purchase as a project fact, then give you the prioritized list.

94. assistant (2026-06-19T17:21:34.923Z)

```
---
name: khelsutra-domain-launch
description: "Owner bought khelsutra.guru (3-yr reg, ~2026-06) and wants a real landing page + a
product/launch push (UI, desktop app, domain) alongside the eng/quality moat."
metadata:
  node_type: memory
  type: project
---

Owner purchased **khelsutra.guru** for **3 years** (~2026-06-19) and wants to replace the
registrar's "dumb" parking page with **good HTML / a real landing page**. This marks a
deliberate pivot of attention toward **product & launch** ? domain/landing page, the UI-alpha
web app (`docs/UI_ALPHA_EXECUTION_PLAN.md`), and the cross-platform desktop app
(`docs/DESKTOP_APP_PLAN.md`) ? run **in parallel** with the engineering/quality moat (owned
model + corpus + rally detection), not instead of it.

**Brand facts (for the landing page):** product = **KhelSutra**, domain **khelsutra.guru**,
brand ratified in ADR-009; reels already carry a *"Generated by Khelsutra.guru"* mark. The app
itself is **invite-only / private / local-first** ? so the PUBLIC page is marketing + an honest
"invite-only alpha" status + a "request access" CTA, NOT the app. Best source copy =
`docs/KHELSUTRA_PRODUCT_OVERVIEW.md` (what it does today, who it's for, the "gets sharper from
real footage" angle, honest "not-today" caveats).

**Why:** the owner is "ramping up on the design/product mindset" and wants active items + a
prioritized pending list, with launch surface (domain) as a near-term quick win.

**How to apply:** treat the landing page as a Tier-0 quick win (static single-page site,
Cloudflare/GitHub Pages, adapt the product-overview copy). Keep app + landing-page concerns
separate (private app stays gated). Ties [[monetization-plan]] · [[current-state]].
```

95. user (2026-06-19T17:21:34.960Z)

File created successfully at: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\khelsutra-domain-launch.md (file state is current in your context ? no need to Read it back)

96. assistant (2026-06-19T17:22:49.115Z)

The classification finished, and it flips the premise in a useful way:

Decluttering: there's almost nothing to archive

Under the repo's own policy (**archive only terminal DONE/REJECTED**), the 8-agent sweep across all 52 docs found **zero clean archive-now candidates** and **zero duplicates** (the 3 "COMPLETED" docs are already thin pointer stubs). Only **2 borderline cases**, both with a catch:

- **`R1_MILESTONE_LAUNCH_REPORT.md`** ? a launched-flip posterity record (terminal by convention), but it's *1 day old* and documents the *current* default config. ? Leave it top-level until the next quality iteration supersedes those defaults, then archive.
- **`REAL_FOOTAGE_VALIDATION_REPORT.md`** ? the validation *event* is done, but it carries a live R1?R8 roadmap still cited by active docs. ? Don't archive until that roadmap is folded into `RALLY_QUALITY_RESEARCH`/`GAP_CLOSURE`.

So the "clutter" isn't unarchived history ? it's breadth of *live* scope. You have ~49 simultaneously-live docs, and **11 of them are "design-not-started"** parked futures (`ANALYZERS_AND_RECONCILERS`, `PER_VIDEO_OUTPUT_LAYOUT`, `I18N_PLAN`, `PLATFORM_ARCHITECTURE`, `HOSTING_PLAN`, `UI_PROPOSAL`, `STORAGE_SHARING_MODEL`, `VIDEO_LOCALITY_MODEL`, `OWNED_MODEL_IMPLEMENTATION_PLAN`, `PERSONALIZATION_PLAN`, `DESKTOP_APP_PLAN`). The fix for clutter is **separating "actively pursued" from "parked someday"** (e.g. a `docs/parked/` folder or a PARKED section in `BACKLOG.md`) ? not archiving. The prioritized list below is what makes that call.

Prioritized pending items

#	Item	Tier	Who	Effort	Status / source
1	**khelsutra.guru landing page** ? replace the parking page with a clean, honest one-pager	? Now	**Me**	S	NEW; copy ready in `KHELSUTRA_PRODUCT_OVERVIEW.md`
2	**Domain/DNS wiring** ? point .guru at the host, HTTPS, optional `hello@` forwarding	? Now	Me + you (registrar login)	S	NEW
3	**UI alpha ? PR-3 identity** (invite-only app login)	? Near	Me (your review)	M	`UI_ALPHA_EXECUTION_PLAN` (PR-1/2 merged)
4	**Desktop app ? P0 compute spike** (CPU/MPS on a no-GPU laptop ? go/no-go)	? Near	You (your machine)	M	`DESKTOP_APP_PLAN` (only un-gated step)
5	**Grow golden corpus** (multi-court + table tennis)	? Moat	You	L (ongoing)	NEXT_STEPS #1 lever
6	**Rally quality: prominent-court / multi-shuttle**	? Moat	Me + you (gated flips)	L	`PROMINENT_COURT_DETECTION` / `GAP_CLOSURE`
7	**M0.4 light temporal head (TemporalMaxer)** ? $0/local model bump	? Moat	Me (your GPU box)	M	NEXT_STEPS #3
8	**Set the ship-gate target**	? Moat	You (decision)	S	NEXT_STEPS #2
9	**PMF field visits** ? kit complete, recruit associate + visit academies	? GTM	You	M	`FIELD_VISIT_*`, `PMF_FIELD_SIGNALS`
10	**Repo rename** off "badminton"	? Housekeeping	You	S	BACKLOG

My recommendation: start with **#1 the landing page** ? it's the thing you're most excited about, it's fully agent-doable right now, and it gives khelsutra.guru a real face today. I'd build a single self-contained `index.html` (the "local-first / honest alpha / request-access" framing from the product overview ? marketing only, app stays private), deploy-ready for Cloudflare or GitHub Pages.

Want me to build it now and show you a live preview? If yes, one quick fork on hosting so I wire it right:

97. user (2026-06-19T17:23:47.367Z)

Your questions have been answered: "Where do you want to host the khelsutra.guru landing page? (Affects deploy instructions + any config files I add.)"="Cloudflare Pages (Recommended)", "What should the landing page's primary call-to-action be?"="Request alpha access (Recommended)". You can now continue with these answers in mind.

98. user (2026-06-19T17:24:30.782Z)

Imagine ? Visual Creation Suite

Modules

Call read_me again with the modules parameter to load detailed guidance:
- `diagram` ? SVG flowcharts, structural diagrams, illustrative diagrams
- `mockup` ? UI mockups, forms, cards, dashboards
- `interactive` ? interactive explainers with controls
- `chart` ? charts, data analysis, geographic maps (Chart.js, D3 choropleth)
- `art` ? illustration and generative art

Pick the closest fit. The module includes all relevant design guidance.

Complexity budget ? hard limits:
- Box subtitles: ?5 words. Detail goes in click-through (`sendPrompt`) or the prose below ? not the box.
- Colors: ?2 ramps per diagram. If colors encode meaning (states, tiers), add a 1-line legend. Otherwise use one neutral ramp.
- Horizontal tier: ?4 boxes at full width (~140px each). 5+ boxes ? shrink to ?110px OR wrap to 2 rows OR split into overview + detail diagrams.

If you catch yourself writing "click to learn more" in prose, the diagram itself must ACTUALLY be sparse. Don't promise brevity then front-load everything.

Accessibility: For HTML widgets, begin with a visually-hidden `<h2 class="sr-only">` containing a one-sentence summary of the visualization for screen-reader users. (SVG widgets use `role="img"` with `<title>` and `<desc>` instead ? see SVG setup.)

You create rich visual content ? SVG diagrams/illustrations and HTML interactive widgets ? that renders inline in conversation. The best output feels like a natural extension of the chat.

Core Design System

These rules apply to ALL use cases.

Philosophy

- **Seamless**: Users shouldn't notice where claude.ai ends and your widget begins.
- **Flat**: No gradients, mesh backgrounds, noise textures, or decorative effects. Clean flat surfaces.
- **Compact**: Show the essential inline. Explain the rest in text.
- **Text goes in your response, visuals go in the tool** ? All explanatory text, descriptions, introductions, and summaries must be written as normal response text OUTSIDE the tool call. The tool output should contain ONLY the visual element (diagram, chart, interactive widget). Never put paragraphs of explanation, section headings, or descriptive prose inside the HTML/SVG. If the user asks "explain X", write the explanation in your response and use the tool only for the visual that accompanies it. The user's font settings only apply to your response text, not to text inside the widget.

Streaming

Output streams token-by-token. Structure code so useful content appears early.
- **HTML**: `<style>` (short) ? content HTML ? `<script>` last.
- **SVG**: `<defs>` (markers) ? visual elements immediately.
- Prefer inline `style="..."` over `<style>` blocks ? inputs/controls must look correct mid-stream.
- Keep `<style>` under ~15 lines. Interactive widgets with inputs and sliders need more style rules ? that's fine, but don't bloat with decorative CSS.

- Gradients, shadows, and blur flash during streaming DOM diffs. Use solid flat fills instead.

Rules

- No `<!-- comments -->` or `/* comments */` (waste tokens, break streaming)
- No font-size below 11px
- No emoji. Icons = Tabler **outline** webfont (5800+, already loaded): `<i class="ti ti-home"></i>`. Outline only ? never use `-filled` suffixes (`ti-heart-filled` etc. are not loaded and will render blank). Inherits color + font-size from parent. Decorative icons get `aria-hidden="true"`; icon-only buttons get `aria-label`. Common: ti-home ti-settings ti-user ti-search ti-x ti-check ti-plus ti-trash ti-edit ti-download ti-upload ti-file ti-folder ti-chart-bar ti-calendar ti-clock ti-arrow-right ti-arrow-left ti-chevron-down ti-external-link ti-copy ti-refresh ti-player-play ti-player-pause ti-heart ti-star ti-bell ti-mail ti-lock ti-eye ti-menu-2. Don't hand-draw icon SVG paths.
- No gradients, drop shadows, blur, glow, or neon effects
- No dark/colored backgrounds on outer containers (transparent only ? host provides the bg)
- **Typography**: The default font is Anthropic Sans. For the rare editorial/blockquote moment, use `font-family: var(--font-serif)`.
- **Headings**: h1 = 22px, h2 = 18px, h3 = 16px ? all `font-weight: 500`. Heading color is pre-set to `var(--color-text-primary)` ? don't override it. Body text = 16px, weight 400, `line-height: 1.7`. **Two weights only: 400 regular, 500 bold.** Never use 600 or 700 ? they look heavy against the host UI.
- **Sentence case** always. Never Title Case, never ALL CAPS. This applies everywhere including SVG text labels and diagram headings.
- **No mid-sentence bolding**, including in your response text around the tool call. Entity names, class names, function names go in `code style` not **bold**. Bold is for headings and labels only.
- The widget container is `display: block; width: 100%`. Your HTML fills it naturally ? no wrapper div needed. Just start with your content directly. If you want vertical breathing room, add `padding: 1rem 0` on your first element.
- Never use `position: fixed` ? the iframe viewport sizes itself to your in-flow content height, so fixed-positioned elements (modals, overlays, tooltips) collapse it to `min-height: 100px`. For modal/overlay mockups: wrap everything in a normal-flow `<div style="min-height: 400px; background: rgba(0,0,0,0.45); display: flex; align-items: center; justify-content: center;">` and put the modal inside ? it's a faux viewport that actually contributes layout height.
- No DOCTYPE, `<html>`, `<head>`, or `<body>` ? just content fragments.
- When placing text on a colored background (badges, pills, cards, tags), use the darkest shade from that same color family for the text ? never plain black or generic gray.
- **Corners**: use `border-radius: var(--border-radius-md)` (or `-lg` for cards) in HTML. In SVG, `rx="4"` is the default ? larger values make pills, use only when you mean a pill.
- **No rounded corners on single-sided borders** ? if using `border-left` or `border-top` accents, set `border-radius: 0`. Rounded corners only work with full borders on all sides.
- **No titles or prose inside the tool output** ? see Philosophy above.
- **Icon sizing**: Tabler `<i class="ti ?">` sizes with `font-size` ? 16?20px inline, 24px max decorative. For one-off inline SVG icons, set `width`/`height` explicitly (same limits).
- No tabs, carousels, or `display: none` sections during streaming ? hidden content streams invisibly. Show all content stacked vertically. (Post-streaming JS-driven steppers are fine ? see Illustrative/Interactive sections.)
- No nested scrolling ? auto-fit height.
- Scripts execute after streaming ? load libraries via `<script src="https://cdnjs.cloudflare.com/ajax/libs/...">` (UMD globals), then use the global in a plain `<script>` that follows.
- **CDN allowlist (CSP-enforced)**: external resources may ONLY load from `cdnjs.cloudflare.com`, `esm.sh`, `cdn.jsdelivr.net`, `unpkg.com`, `fonts.googleapis.com`, `fonts.gstatic.com`. All other origins are blocked by the sandbox ? the request silently fails.

CSS Variables

Backgrounds: `--color-background-primary` (white), `-secondary` (surfaces), `-tertiary` (page bg), `-info`, `-danger`, `-success`, `-warning`
Text: `--color-text-primary` (black), `-secondary` (muted), `-tertiary` (hints), `-info`, `-danger`, `-success`, `-warning`
Borders: `--color-border-tertiary` (0.15?, default), `-secondary` (0.3?, hover), `-primary` (0.4?), semantic `-info/-danger/-success/-warning`
Typography: `--font-sans`, `--font-serif`, `--font-mono`
Layout: `--border-radius-md` (8px), `--border-radius-lg` (12px ? preferred for most components), `--border-radius-xl` (16px)

All auto-adapt to light/dark mode. For custom colors in HTML, use CSS variables.

Dark mode is mandatory ? every color must work in both modes:
- In SVG: use the pre-built color classes (`c-blue`, `c-teal`, `c-amber`, etc.) for colored nodes ? they handle light/dark mode automatically. Never write `<style>` blocks for colors.
- In SVG: every `<text>` element needs a class (`t`, `ts`, `th`) ? never omit fill or use `fill="inherit"`. Inside a `c-{color}` parent, text classes auto-adjust to the ramp.
- In HTML: always use CSS variables (--color-text-primary, --color-text-secondary) for text. Never hardcode colors like color: #333 ? invisible in dark mode.
- Mental test: if the background were near-black, would every text element still be readable?

sendPrompt(text)

A global function that sends a message to chat as if the user typed it. Use it when the user's next step benefits from Claude thinking. Handle filtering, sorting, toggling, and calculations in JS instead.

Links

`<a href="https://...">` just works ? clicks are intercepted and open the host's link-confirmation dialog. Or call `openLink(url)` directly.

When nothing fits

Pick the closest use case below and adapt. When nothing fits cleanly:
- Default to editorial layout if the content is explanatory
- Default to card layout if the content is a bounded object
- All core design system rules still apply
- Use `sendPrompt()` for any action that benefits from Claude thinking

UI components

Layout width

The widget container is 680px wide. Use `repeat(auto-fit, minmax(160px, 1fr))` for responsive columns ? auto-fit lets the grid pick column count by available width.

Aesthetic

Flat, clean, white surfaces. Minimal 0.5px borders. Generous whitespace. No gradients, no shadows (except functional focus rings). Everything should feel native to claude.ai ? like it belongs on the page, not embedded from somewhere else.

Tokens

- Borders: always `0.5px solid var(--color-border-tertiary)` (or `-secondary` for emphasis)
- Corner radius: `var(--border-radius-md)` for most elements, `var(--border-radius-lg)` for cards
- Cards: white bg (`var(--color-background-primary)`), 0.5px border, radius-lg, padding 1rem 1.25rem
- Form elements (input, select, textarea, button, range slider) are pre-styled ? write bare tags. Text inputs are 36px with hover/focus built in; range sliders have 4px track + 18px thumb; buttons have outline style with hover/active. Only add inline styles to override (e.g., different width).
- Buttons: pre-styled with transparent bg, 0.5px border-secondary, hover bg-secondary, active scale(0.98). If it triggers sendPrompt, append a ? arrow.
- **Round every displayed number.** JS float math leaks artifacts ? `0.1 + 0.2` gives `0.30000000000000004`, `7 * 1.1` gives `7.700000000000001`. Any number that reaches the screen (slider readouts, stat card values, axis labels, data-point labels, tooltips, computed totals) must go through `Math.round()`, `.toFixed(n)`, or `Intl.NumberFormat`. Pick the precision that makes sense for the context ? integers for counts, 1?2 decimals for percentages, `toLocaleString()` for currency. For range sliders, also set `step="1"` (or step="0.1" etc.) so the input itself emits round values.
- Spacing: use rem for vertical rhythm (1rem, 1.5rem, 2rem), px for component-internal gaps (8px, 12px, 16px)
- Box-shadows: none, except `box-shadow: 0 0 0 Npx` focus rings on inputs

Metric cards

For summary numbers (revenue, count, percentage) ? surface card with muted 13px label above, 24px/500 number below. `background: var(--color-background-secondary)`, no border, `border-radius: var(--border-radius-md)`, padding 1rem. Use in grids of 2-4 with `gap: 12px`. Distinct from raised cards (which have white bg + border).

Layout

- Editorial (explanatory content): no card wrapper, prose flows naturally
- Card (bounded objects like a contact record, receipt): single raised card wraps the whole thing
- Don't put tables here ? output them as markdown in your response text

Grid overflow: `grid-template-columns: 1fr` has `min-width: auto` by default ? children with large min-content push the column past the container. Use `minmax(0, 1fr)` to clamp.

Table overflow: Tables with many columns auto-expand past `width: 100%` if cell contents exceed it. In constrained layouts (?700px), use `table-layout: fixed` and set explicit column widths, or reduce columns, or allow horizontal scroll on a wrapper.

Mockup presentation

Contained mockups ? mobile screens, chat threads, single cards, modals, small UI components ? should sit on a background surface (`var(--color-background-secondary)` container with `border-radius: var(--border-radius-lg)` and padding, or a device frame) so they don't float naked on the widget canvas. Full-width mockups like dashboards, settings pages, or data tables that naturally fill the viewport do not need an extra wrapper.

1. Interactive explainer ? learn how something works

"Explain how compound interest works" / "Teach me about sorting algorithms"

Use HTML for the interactive controls ? sliders, buttons, live state displays, charts. Keep prose explanations in your normal response text (outside the tool call), not embedded in the HTML. No card wrapper. Whitespace is the container.

```html
<div style="display: flex; align-items: center; gap: 12px; margin: 0 0 1.5rem;">
  <label style="font-size: 14px; color: var(--color-text-secondary);">Years</label>
  <input type="range" min="1" max="40" value="20" id="years" style="flex: 1;" />
  <span style="font-size: 14px; font-weight: 500; min-width: 24px;" id="years-out">20</span>
</div>

<div style="display: flex; align-items: baseline; gap: 8px; margin: 0 0 1.5rem;">
  <span style="font-size: 14px; color: var(--color-text-secondary);">£1,000 ?</span>
  <span style="font-size: 24px; font-weight: 500;" id="result">£3,870</span>
</div>

<div style="margin: 2rem 0; position: relative; height: 240px;">
  <canvas id="chart"></canvas>
</div>
```

Use `sendPrompt()` to let users ask follow-ups: `sendPrompt('What if I increase the rate to 10%?')`

2. Compare options ? decision making

"Compare pricing and features of these products" / "Help me choose between React and Vue"

Use HTML. Side-by-side card grid for options. Highlight differences with semantic colors. Interactive elements for filtering or weighting.

- Each option in a card. Use badges for key differentiators. A leading Tabler icon (`<i class="ti ti-NAME">` at 20px, `aria-hidden`) anchors each option visually ? pick the most apt name per option.
- Add `sendPrompt()` buttons: `sendPrompt('Tell me more about the Pro plan')`
- Don't put comparison tables inside this tool ? output them as regular markdown tables in your response text instead. The tool is for the visual card grid only.

- When one option is recommended or "most popular", accent its card with `border: 2px solid var(--color-border-info)` only (2px is deliberate ? the only exception to the 0.5px rule, used to accent featured items) ? keep the same background and border as the other cards. Add a small badge (e.g. "Most popular") above or inside the card header using `background: var(--color-background-info); color: var(--color-text-info); font-size: 12px; padding: 4px 12px; border-radius: var(--border-radius-md)`.

3. Data record ? bounded UI object

"Show me a Salesforce contact card" / "Create a receipt for this order"

Use HTML. Wrap the entire thing in a single raised card. All content is sans-serif since it's pure UI. Use an avatar/initials circle for people (see example below).

```html
<div style="background: var(--color-background-primary); border-radius: var(--border-radius-lg); border: 0.5px solid var(--color-border-tertiary); padding: 1rem 1.25rem;">
  <div style="display: flex; align-items: center; gap: 12px; margin-bottom: 16px;">
    <div style="width: 44px; height: 44px; border-radius: 50%; background: var(--color-background-info); display: flex; align-items: center; justify-content: center; font-weight: 500; font-size: 14px; color: var(--color-text-info);">MR</div>
    <div>
      <p style="font-weight: 500; font-size: 15px; margin: 0;">Maya Rodriguez</p>
      <p style="font-size: 13px; color: var(--color-text-secondary); margin: 0;">VP of Engineering</p>
    </div>
  </div>
  <div style="border-top: 0.5px solid var(--color-border-tertiary); padding-top: 12px;">
    <table style="width: 100%; font-size: 13px;">
      <tr><td style="color: var(--color-text-secondary); padding: 4px 0;"><i class="ti ti-mail" style="font-size:16px; vertical-align:-2px; margin-right:6px" aria-hidden="true"></i>Email</td><td style="text-align: right; padding: 4px 0; color: var(--color-text-info);">m.rodriguez@acme.com</td></tr>
      <tr><td style="color: var(--color-text-secondary); padding: 4px 0;"><i class="ti ti-phone" style="font-size:16px; vertical-align:-2px; margin-right:6px" aria-hidden="true"></i>Phone</td><td style="text-align: right; padding: 4px 0;">+1 (415) 555-0172</td></tr>
    </table>
  </div>
</div>
```

Color palette

9 color ramps, each with 7 stops from lightest to darkest. 50 = lightest fill, 100-200 = light fills, 400 = mid tones, 600 = strong/border, 800-900 = text on light fills.

Class	Ramp	50 (lightest)	100	200	400	600	800	900 (darkest)
`c-purple`	Purple	#EEEDFE	#CECBF6	#AFA9EC	#7F77DD	#534AB7	#3C3489	#26215C
`c-teal`	Teal	#E1F5EE	#9FE1CB	#5DCAA5	#1D9E75	#0F6E56	#085041	#04342C
`c-coral`	Coral	#FAECE7	#F5C4B3	#F0997B	#D85A30	#993C1D	#712B13	#4A1B0C
`c-pink`	Pink	#FBEAF0	#F4C0D1	#ED93B1	#D4537E	#993556	#72243E	#4B1528
`c-gray`	Gray	#F1EFE8	#D3D1C7	#B4B2A9	#888780	#5F5E5A	#444441	#2C2C2A
`c-blue`	Blue	#E6F1FB	#B5D4F4	#85B7EB	#378ADD	#185FA5	#0C447C	#042C53
`c-green`	Green	#EAF3DE	#C0DD97	#97C459	#639922	#3B6D11	#27500A	#173404
`c-amber`	Amber	#FAEEDA	#FAC775	#EF9F27	#BA7517	#854F0B	#633806	#412402
`c-red`	Red	#FCEBEB	#F7C1C1	#F09595	#E24B4A	#A32D2D	#791F1F	#501313

How to assign colors: Color should encode meaning, not sequence. Don't cycle through colors like a rainbow (step 1 = blue, step 2 = amber, step 3 = red...). Instead:
- Group nodes by **category** ? all nodes of the same type share one color. E.g. in a vaccine diagram: all immune cells = purple, all pathogens = coral, all outcomes = teal.
- For illustrative diagrams, map colors to **physical properties** ? warm ramps for heat/energy,

cool for cold/calm, green for organic, gray for structural/inert.
- Use **gray for neutral/structural** nodes (start, end, generic steps).
- Use **2-3 colors per diagram**, not 6+. More colors = more visual noise. A diagram with gray + purple + teal is cleaner than one using every ramp.
- **Prefer purple, teal, coral, pink** for general diagram categories. Reserve blue, green, amber, and red for cases where the node genuinely represents an informational, success, warning, or error concept ? those colors carry strong semantic connotations from UI conventions. (Exception: illustrative diagrams may use blue/amber/red freely when they map to physical properties like temperature or pressure.)

Text on colored backgrounds: Always use the 800 or 900 stop from the same ramp as the fill. Never use black, gray, or --color-text-primary on colored fills. **When a box has both a title and a subtitle, they must be two different stops** ? title darker (800 in light mode, 100 in dark), subtitle lighter (600 in light, 200 in dark). Same stop for both reads flat; the weight difference alone isn't enough. For example, text on Blue 50 (#E6F1FB) must use Blue 800 (#0C447C) or 900 (#042C53), not black. This applies to SVG text elements inside colored rects, and to HTML badges, pills, and labels with colored backgrounds.

Light/dark mode quick pick ? use only stops from the table, never off-table hex values:
- **Light mode**: 50 fill + 600 stroke + **800 title / 600 subtitle**
- **Dark mode**: 800 fill + 200 stroke + **100 title / 200 subtitle**
- Apply `c-{ramp}` to a `<g>` wrapping shape+text, or directly to a `<rect>`/`<circle>`/`<ellipse>`. Never to `<path>` ? paths don't get ramp fill. For colored connector strokes use inline `stroke="#..."` (any mid-ramp hex works in both modes). Dark mode is automatic for ramp classes. Available: c-gray, c-blue, c-red, c-amber, c-green, c-teal, c-purple, c-coral, c-pink.

For status/semantic meaning in UI (success, warning, danger) use CSS variables. For categorical coloring in both diagrams and UI, use these ramps.

99. assistant (2026-06-19T17:25:31.058Z)

Both calls came back the way I'd hoped ? **Cloudflare Pages** + a **"Request alpha access"** email CTA. That's the cleanest setup for a freshly-bought `.guru` domain. Let me build a real, deployable landing page (its own brand styling, not a throwaway) and then show you a live preview.

100. assistant (2026-06-19T17:26:55.506Z)

```
<!doctype html>
<html lang="en">
<head>
<meta charset="utf-8" />
<meta name="viewport" content="width=device-width, initial-scale=1" />
<title>KhelSutra ? automatic badminton highlight reels from your session video</title>
<meta name="description" content="KhelSutra turns ordinary badminton session video into rallies,
a highlight reel, and a simple play-time summary ? automatically. Invite-only alpha for
academies and players in Bengaluru." />
<meta property="og:title" content="KhelSutra ? automatic badminton highlights" />
<meta property="og:description" content="Record a session, get back the rallies and a highlight
reel. No editing, no scrubbing. Built to crack the hard multi-court case." />
<meta property="og:type" content="website" />
<meta property="og:url" content="https://khelsutra.guru" />
<meta name="theme-color" content="#0C261F" />
<link rel="preconnect" href="https://fonts.googleapis.com" />
<link rel="preconnect" href="https://fonts.gstatic.com" crossorigin />
<link href="https://fonts.googleapis.com/css2?family=Plus+Jakarta+Sans:wght@400;500;600;700;800&
display=swap" rel="stylesheet" />
<style>
  :root{
    --ink:#0C261F; --ink2:#123A2E; --green:#19B36B; --green-d:#149A5C;
    --lime:#C7F2B0; --paper:#F6F8F4; --card:#fff;
    --tx:#14271F; --mut:#566B61; --line:#E4E9E2;
    --on-dark:#EAF3EE; --on-dark-mut:#A7C3B7;
    --r:14px; --rs:10px; --maxw:1080px;
```

```
    --font:'Plus Jakarta Sans',system-ui,-apple-system,Segoe UI,Roboto,Helvetica,Arial,sans-
serif;
  }
  *{box-sizing:border-box}
  html{scroll-behavior:smooth}
  body{margin:0;font-family:var(--font);color:var(--tx);background:var(--paper);line-
height:1.65;-webkit-font-smoothing:antialiased}
  a{color:inherit}
  .wrap{max-width:var(--maxw);margin:0 auto;padding:0 22px}
  h1,h2,h3{line-height:1.15;margin:0;letter-spacing:-.02em}
  h1{font-size:clamp(2rem,5vw,3.4rem);font-weight:800}
  h2{font-size:clamp(1.5rem,3.2vw,2.2rem);font-weight:700}
  h3{font-size:1.15rem;font-weight:700}
  p{margin:.6rem 0}
  .btn{display:inline-flex;align-items:center;gap:.5rem;font-weight:600;font-size:1rem;border-
radius:999px;padding:.78rem 1.4rem;text-decoration:none;border:1px solid
transparent;cursor:pointer;transition:.15s}
  .btn-primary{background:var(--green);color:#04140D}
  .btn-primary:hover{background:var(--green-d)}
  .btn-ghost{background:transparent;border-color:rgba(255,255,255,.28);color:var(--on-dark)}
  .btn-ghost:hover{border-color:rgba(255,255,255,.6)}
  .btn-dark{background:var(--ink);color:#fff}
  .btn-dark:hover{background:var(--ink2)}
  .eyebrow{display:inline-flex;align-items:center;gap:.5rem;font-size:.82rem;font-
weight:600;letter-spacing:.02em;
    background:rgba(25,179,107,.14);color:#0a6c45;border:1px solid
rgba(25,179,107,.3);padding:.3rem .8rem;border-radius:999px}

  /* header */
  header{position:sticky;top:0;z-index:20;background:rgba(246,248,244,.86);backdrop-
filter:blur(10px);border-bottom:1px solid var(--line)}
  .nav{display:flex;align-items:center;justify-content:space-between;height:66px}
  .brand{display:flex;align-items:center;gap:.6rem;font-weight:800;font-size:1.18rem;letter-
spacing:-.02em;text-decoration:none}
  .brand .dot{color:var(--green)}
  .nav-links{display:flex;align-items:center;gap:1.6rem}
  .nav-links a{font-size:.95rem;font-weight:500;color:var(--mut);text-decoration:none}
  .nav-links a:hover{color:var(--tx)}
  .logo-mark{width:30px;height:30px;flex:none}

  /* hero */
  .hero{background:linear-gradient(160deg,var(--ink) 0%,var(--ink2) 100%);color:var(--on-
dark);position:relative;overflow:hidden}
  .hero .wrap{padding-top:74px;padding-bottom:88px;position:relative;z-index:2}
  .hero .eyebrow{background:rgba(199,242,176,.12);color:var(--lime);border-
color:rgba(199,242,176,.25)}
  .hero h1{margin:1.1rem 0 0;max-width:15ch}
  .hero h1 .hl{color:var(--lime)}
  .hero .lead{font-size:clamp(1.05rem,1.6vw,1.28rem);color:var(--on-dark-mut);max-
width:54ch;margin:1.2rem 0 0}
  .hero-cta{display:flex;flex-wrap:wrap;gap:.8rem;margin-top:2rem}
  .hero-note{margin-top:1.1rem;font-size:.86rem;color:var(--on-dark-mut)}
  .shuttle{position:absolute;right:-60px;top:-40px;width:520px;max-
width:55%;opacity:.16;z-index:1;pointer-events:none}
  .loop{display:flex;flex-wrap:wrap;gap:.6rem;margin-top:2.4rem}
  .loop .chip{display:flex;align-
items:center;gap:.5rem;background:rgba(255,255,255,.06);border:1px solid rgba(255,255,255,.12);
    border-radius:999px;padding:.5rem 1rem;font-weight:500;font-size:.92rem;color:var(--on-
dark)}
  .loop .chip b{color:var(--lime);font-weight:700}
  .loop .arrow{color:var(--on-dark-mut);align-self:center}

  /* sections */
  section{padding:72px 0}
  .sec-head{max-width:60ch}
```

```
  .sec-head .kicker{font-size:.82rem;font-weight:700;letter-spacing:.08em;text-
transform:uppercase;color:var(--green-d)}
  .sec-head h2{margin:.5rem 0 0}
  .sec-head p{color:var(--mut);font-size:1.08rem;margin-top:.7rem}

  .grid{display:grid;gap:18px;margin-top:2.4rem}
  .g4{grid-template-columns:repeat(4,1fr)}
  .g3{grid-template-columns:repeat(3,1fr)}
  .g2{grid-template-columns:repeat(2,1fr)}
  .card{background:var(--card);border:1px solid var(--line);border-
radius:var(--r);padding:1.4rem}
  .card .ic{width:42px;height:42px;border-radius:11px;display:flex;align-items:center;justify-
content:center;
    background:rgba(25,179,107,.12);color:var(--green-d);font-size:1.3rem;margin-bottom:.9rem}
  .card h3{font-size:1.05rem}
  .card p{color:var(--mut);font-size:.96rem;margin:.4rem 0 0}

  .split{display:grid;grid-template-columns:1.1fr .9fr;gap:42px;align-items:center}
  .panel{background:var(--ink);color:var(--on-dark);border-radius:18px;padding:2.2rem}
  .panel h2{color:#fff}
  .panel p{color:var(--on-dark-mut)}
  .panel .pill{display:inline-block;background:rgba(199,242,176,.14);color:var(--lime);font-
size:.8rem;font-weight:600;
    padding:.3rem .7rem;border-radius:999px;margin-bottom:1rem}

  .road{display:grid;grid-template-columns:repeat(5,1fr);gap:14px;margin-top:2.2rem}
  .road .step{background:var(--card);border:1px solid var(--line);border-
radius:var(--rs);padding:1.1rem}
  .road .step .n{font-size:1.4rem;color:var(--green-d)}
  .road .step h3{font-size:1rem;margin-top:.5rem}
  .road .step p{font-size:.88rem;color:var(--mut);margin-top:.3rem}

  .honest{background:#fff;border:1px solid var(--line);border-left:4px solid
var(--green);border-radius:0 var(--rs) var(--rs) 0;padding:1.2rem 1.4rem;margin-top:1.6rem}
  .honest b{color:var(--tx)}
  .honest p{color:var(--mut);margin:.3rem 0 0}

  .nots{list-style:none;padding:0;margin:2rem 0 0;display:grid;grid-template-
columns:repeat(2,1fr);gap:12px}
  .nots li{display:flex;gap:.6rem;background:var(--card);border:1px solid var(--line);border-
radius:var(--rs);padding:.9rem 1.1rem;color:var(--mut);font-size:.95rem}
  .nots li i{color:var(--green-d);margin-top:.15rem}

  details{background:var(--card);border:1px solid var(--line);border-radius:var(--rs);padding:0
1.2rem;margin-top:10px}
  details summary{cursor:pointer;font-weight:600;padding:1rem 0;list-
style:none;display:flex;justify-content:space-between;align-items:center;gap:1rem}
  details summary::-webkit-details-marker{display:none}
  details summary::after{content:"+";color:var(--green-d);font-size:1.3rem;font-weight:400}
  details[open] summary::after{content:"?"}
  details p{color:var(--mut);margin:0 0 1.1rem}

  .cta{background:linear-gradient(160deg,var(--ink) 0%,var(--ink2) 100%);color:var(--on-
dark);border-radius:20px;padding:3rem 2.4rem;text-align:center;margin:0 0 0}
  .cta h2{color:#fff;max-width:20ch;margin:0 auto}
  .cta p{color:var(--on-dark-mut);max-width:46ch;margin:.8rem auto 0}
  .cta .hero-cta{justify-content:center}

  footer{border-top:1px solid var(--line);padding:42px 0;color:var(--mut);font-size:.9rem}
  .foot{display:flex;flex-wrap:wrap;justify-content:space-between;gap:1.2rem;align-items:center}
  .foot a{color:var(--green-d);text-decoration:none}
  .foot a:hover{text-decoration:underline}
  .stamp{font-size:.82rem;color:var(--mut)}

  @media (max-width:860px){
```

```
    .nav-links{display:none}
    .g4{grid-template-columns:repeat(2,1fr)}
    .g3,.g2{grid-template-columns:1fr}
    .split{grid-template-columns:1fr;gap:24px}
    .road{grid-template-columns:repeat(2,1fr)}
    .nots{grid-template-columns:1fr}
    .shuttle{opacity:.1}
  }
</style>
</head>
<body>

<header>
  <div class="wrap nav">
    <a class="brand" href="#top">
      <svg class="logo-mark" viewBox="0 0 32 32" fill="none" aria-hidden="true">
        <circle cx="16" cy="22" r="6" fill="#19B36B"/>
        <path d="M16 16 L9 3 M16 16 L16 2 M16 16 L23 3" stroke="#0C261F" stroke-width="2"
stroke-linecap="round"/>
        <path d="M9 3 L16 6 L23 3" stroke="#0C261F" stroke-width="1.6" stroke-linecap="round"
fill="none"/>
      </svg>
      Khel<span class="dot">Sutra</span>
    </a>
    <nav class="nav-links" aria-label="Primary">
      <a href="#how">How it works</a>
      <a href="#diff">Why it's hard</a>
      <a href="#roadmap">Roadmap</a>
      <a href="#faq">FAQ</a>
    </nav>
    <a class="btn btn-dark" href="mailto:hello@khelsutra.guru?subject=KhelSutra%20alpha%20access
&amp;body=Hi%20KhelSutra%20team%2C%20I'd%20like%20to%20try%20the%20alpha.%0A%0AAcademy%20%2F%20n
ame%3A%20%0ACity%3A%20%0ANumber%20of%20courts%3A%20%0AHow%20you%20record%20today%3A%20">Request
access</a>
  </div>
</header>

<a id="top"></a>
<section class="hero" style="padding:0">
  <svg class="shuttle" viewBox="0 0 200 200" fill="none" aria-hidden="true">
    <circle cx="100" cy="150" r="34" fill="#C7F2B0"/>
    <path d="M100 120 L55 25 M100 120 L100 18 M100 120 L145 25 M100 120 L74 30 M100 120 L126 30"
stroke="#C7F2B0" stroke-width="3" stroke-linecap="round"/>
  </svg>
  <div class="wrap">
    <span class="eyebrow">Invite-only alpha · Bengaluru</span>
    <h1>Turn ordinary badminton video into <span class="hl">highlight reels</span> ?
automatically.</h1>
    <p class="lead">Put a phone or GoPro on a tripod behind the court and record a session.
KhelSutra finds the rallies, stitches a highlight reel, and gives you a simple play-time
summary. No editing, no scrubbing.</p>
    <div class="hero-cta">
      <a class="btn btn-primary" href="mailto:hello@khelsutra.guru?subject=KhelSutra%20alpha%20a
ccess&amp;body=Hi%20KhelSutra%20team%2C%20I'd%20like%20to%20try%20the%20alpha.%0A%0AAcademy%20%2
F%20name%3A%20%0ACity%3A%20%0ANumber%20of%20courts%3A%20%0AHow%20you%20record%20today%3A%20">Req
uest alpha access</a>
      <a class="btn btn-ghost" href="#how">See how it works</a>
    </div>
    <p class="hero-note">A phone or GoPro on a tripod is all you need. No special hardware.</p>
    <div class="loop">
      <span class="chip"><b>1.</b> Record a session</span>
      <span class="arrow">?</span>
      <span class="chip"><b>2.</b> Upload the video</span>
      <span class="arrow">?</span>
      <span class="chip"><b>3.</b> Get a reel + summary</span>
```

```
    </div>
  </div>
</section>

<section id="how">
  <div class="wrap">
    <div class="sec-head">
      <div class="kicker">What it does today</div>
      <h2>Record ? upload ? get back a reel + summary</h2>
      <p>This is what's working right now, in the private web app invited users can try. No
editing skill needed.</p>
    </div>
    <div class="grid g4">
      <div class="card">
        <div class="ic" aria-hidden="true">?</div>
        <h3>Automatic rally detection</h3>
        <p>It finds the rallies ? where each one starts and ends ? on its own. Long full-session
videos are fine.</p>
      </div>
      <div class="card">
        <div class="ic" aria-hidden="true">?</div>
        <h3>Highlight reel</h3>
        <p>The rallies stitched into one MP4 you can watch, download, and share ? with a small
"Generated by Khelsutra.guru" mark.</p>
      </div>
      <div class="card">
        <div class="ic" aria-hidden="true">?</div>
        <h3>A simple summary</h3>
        <p>How many rallies were found and how much actual play time there was ? at a
glance.</p>
      </div>
      <div class="card">
        <div class="ic" aria-hidden="true">?</div>
        <h3>Process once, reuse</h3>
        <p>Come back later and download the reel. Re-running a video replaces its earlier result
rather than piling up copies.</p>
      </div>
    </div>
  </div>
</section>

<section id="diff" style="padding-top:0">
  <div class="wrap">
    <div class="split">
      <div>
        <div class="kicker" style="font-size:.82rem;font-weight:700;letter-spacing:.08em;text-
transform:uppercase;color:var(--green-d)">The hard problem we crack</div>
        <h2 style="margin-top:.5rem">Most academy halls run several courts at once</h2>
        <p style="color:var(--mut);font-size:1.05rem">Academies already record a lot of video ?
and then almost never watch it, because cutting it into useful clips by hand is tedious. The
good rallies stay buried.</p>
        <p style="color:var(--mut);font-size:1.05rem">A single recording of a busy, multi-court
hall is exactly where ordinary video tools choke. Pulling the <b
style="color:var(--tx)">real</b> rallies out of busy, multi-court footage is the hard problem
KhelSutra is built for. It already works very well on clean, single-court recordings ? and the
multi-court case is what we're sharpening now.</p>
      </div>
      <div class="panel">
        <span class="pill">Capture tip</span>
        <h2 style="font-size:1.4rem">Best results: a steady, full-court shot</h2>
        <p>Camera behind the court, slightly elevated, whole court in frame. Keep it still ? a
fixed tripod beats handheld every time. Shoot landscape, record the whole session, upload as-is.
KhelSutra does the cutting.</p>
      </div>
    </div>
```

```
  </div>
</section>

<section id="sharper" style="background:#EFF4EC">
  <div class="wrap">
    <div class="sec-head">
      <div class="kicker">What makes it different</div>
      <h2>It gets sharper from real academy footage</h2>
      <p>The footage real academies contribute ? especially the busy multi-court halls ? is what
makes KhelSutra better over time. Early users aren't just customers; they're co-builders.</p>
    </div>
    <div class="grid g3">
      <div class="card">
        <div class="ic" aria-hidden="true">?</div>
        <h3>Improves with the library</h3>
        <p>When you choose to share footage and coaches confirm what it found, that growing
labelled library is what sharpens future versions ? rather than staying frozen.</p>
      </div>
      <div class="card">
        <div class="ic" aria-hidden="true">?</div>
        <h3>Your data stays yours</h3>
        <p>We never use your video to improve the tool without your clear, explicit opt-in. We
never publish your footage. You decide what, if anything, you share.</p>
      </div>
      <div class="card">
        <div class="ic" aria-hidden="true">?</div>
        <h3>Minors are protected</h3>
        <p>Footage involving minors is never used to improve the tool and is never published ?
full stop.</p>
      </div>
    </div>
    <div class="honest">
      <b>Read this honestly:</b>
      <p>Today the tool gives you rally detection, reels, and summaries. It does not retrain
itself as you use it, and there's no per-person tracking. "Gets sharper over time" is the model
we're building toward, powered by contributed footage ? not a switch that's already on.</p>
    </div>
  </div>
</section>

<section id="roadmap">
  <div class="wrap">
    <div class="sec-head">
      <div class="kicker">Where it's going</div>
      <h2>Toward an AI assistant coach</h2>
      <p>Fed by the growing video library, KhelSutra is heading from "finds rallies" to a
genuine training companion. This is the direction ? not today's buttons.</p>
    </div>
    <div class="road">
      <div class="step"><div class="n">?</div><h3>Shot detection</h3><p>Smashes, drops, net
shots, serves and clears within each rally.</p></div>
      <div class="step"><div class="n">?</div><h3>Deeper analysis</h3><p>Per-rally and per-match
patterns ? what's working and what isn't.</p></div>
      <div class="step"><div class="n">?</div><h3>Per-match history</h3><p>Track progress over
weeks and months, not just one clip.</p></div>
      <div class="step"><div class="n">?</div><h3>Assistant coach</h3><p>Surfaces what to work
on next ? supports the real coach, never replaces.</p></div>
      <div class="step"><div class="n">?</div><h3>More sports & languages</h3><p>Table tennis
next; Kannada and Hindi alongside English.</p></div>
    </div>
  </div>
</section>

<section id="not" style="padding-top:0">
  <div class="wrap">
```

```
    <div class="sec-head">
      <div class="kicker">So expectations stay honest</div>
      <h2>What KhelSutra is not</h2>
    </div>
    <ul class="nots">
      <li><i aria-hidden="true">?</i><span>An early, invite-only trial ? useful, but still
improving. You'll hit rough edges.</span></li>
      <li><i aria-hidden="true">?</i><span>Not a line-calling, scoring, or umpiring
system.</span></li>
      <li><i aria-hidden="true">?</i><span>Not player tracking or identification.</span></li>
      <li><i aria-hidden="true">?</i><span>Not a replacement for a coach ? a tool that saves
time and adds insight.</span></li>
      <li><i aria-hidden="true">?</i><span>Shot detection and deeper analytics are on the
roadmap, not available today.</span></li>
      <li><i aria-hidden="true">?</i><span>Pricing isn't set ? today's invited use is about
getting it right.</span></li>
    </ul>
  </div>
</section>

<section id="faq" style="background:#EFF4EC">
  <div class="wrap">
    <div class="sec-head">
      <div class="kicker">Quick answers</div>
      <h2>Common questions</h2>
    </div>
    <div style="margin-top:1.6rem">
      <details><summary>How accurate is it?</summary><p>Very good on a clean, single-court
recording. A busy multi-court hall is harder ? and is exactly what we're improving. The best
answer is to show you a sample reel of your own court rather than quote a number.</p></details>
      <details><summary>Do I need a special camera or hardware?</summary><p>No. A phone or GoPro
on a tripod, behind the court, whole court in frame. We provide a one-page capture checklist at
setup.</p></details>
      <details><summary>Is my video private? What happens to it?</summary><p>We treat your
footage as private: we never publish it, and we never use it to improve the tool without your
explicit opt-in. These are large videos that need real processing, so your video is handled
securely on our infrastructure to make your reel. Footage involving minors is never used and
never published.</p></details>
      <details><summary>Does it work for doubles and tournaments?</summary><p>It finds rallies
in doubles too. Tournament footage varies ? a clean, fixed, full-court recording works well;
crowded multi-court tournament halls are the hard case we're specifically working
on.</p></details>
      <details><summary>Which sports?</summary><p>Badminton today. Table tennis and other racket
sports are on the roadmap.</p></details>
      <details><summary>How do I join?</summary><p>It's an invite-only private trial. The best
next step is to let us record a session and show you your own court and players as a reel. Email
<a href="mailto:hello@khelsutra.guru">hello@khelsutra.guru</a>.</p></details>
    </div>
  </div>
</section>

<section id="join" style="padding-bottom:80px">
  <div class="wrap">
    <div class="cta">
      <h2>Want to see your academy's own footage as a reel?</h2>
      <p>That's the best way to get a feel for KhelSutra. Tell us a little about your setup and
we'll get you in.</p>
      <div class="hero-cta">
        <a class="btn btn-primary" href="mailto:hello@khelsutra.guru?subject=KhelSutra%20alpha%2
0access&amp;body=Hi%20KhelSutra%20team%2C%20I'd%20like%20to%20try%20the%20alpha.%0A%0AAcademy%20
%2F%20name%3A%20%0ACity%3A%20%0ANumber%20of%20courts%3A%20%0AHow%20you%20record%20today%3A%20">R
equest alpha access</a>
        <a class="btn btn-ghost" href="mailto:hello@khelsutra.guru">hello@khelsutra.guru</a>
      </div>
    </div>
```

```
  </div>
</section>

<footer>
  <div class="wrap foot">
    <div>
      <div class="brand" style="font-size:1.05rem">Khel<span class="dot">Sutra</span></div>
      <div class="stamp" style="margin-top:.4rem">Automatic badminton highlights · Bengaluru,
India</div>
    </div>
    <div style="text-align:right">
      <div><a href="mailto:hello@khelsutra.guru">hello@khelsutra.guru</a> · <a
href="https://khelsutra.guru">khelsutra.guru</a></div>
      <div class="stamp" style="margin-top:.4rem">© 2026 KhelSutra · Invite-only alpha · Reels
carry a "Generated by Khelsutra.guru" mark</div>
    </div>
  </div>
</footer>

</body>
</html>
```

101. user (2026-06-19T17:26:55.563Z)

File created successfully at: C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html (file state is current in your context ? no need to Read it back)

102. assistant (2026-06-19T17:27:39.574Z)

Done ? the landing page is built and rendering live in the **Launch preview panel** (and you can open `artifacts/landing/index.html` in any browser). It's a single self-contained file: brand styling, one web font, inline SVG logo, no build step.

What's on it ? adapted straight from your product overview, keeping its honest tone:
- Hero: *"Turn ordinary badminton video into highlight reels ? automatically"* + the record ? upload ? reel loop + **Request alpha access** button (? `hello@khelsutra.guru` with a prefilled academy/city/courts template).
- What it does today (4 cards) · the multi-court "hard problem" differentiator · "gets sharper from real footage" + the privacy/minors guardrail and the honest *"not a switch that's already on"* note · the roadmap (shot detection ? assistant coach) clearly marked as direction · "what it's not" · a 6-question FAQ · footer with the *"Generated by Khelsutra.guru"* mark.

Ship it to khelsutra.guru on Cloudflare Pages

1 · Deploy the page (free, no git needed)
- Cloudflare dashboard ? **Workers & Pages ? Create ? Pages ? Upload assets**.
- Name it `khelsutra-site`, drag in the `landing/` folder (the one holding `index.html`), Deploy ? you get a live `khelsutra-site.pages.dev` URL instantly.

2 · Point the domain at it
- Add `khelsutra.guru` to Cloudflare (dashboard ? **Add a site**). It gives you **2 nameservers** ? set those at the registrar where you bought it. (This hands DNS to Cloudflare and makes the next steps one-click.)
- In the Pages project ? **Custom domains ? Set up a domain** ? `khelsutra.guru` (and `www`). Cloudflare auto-creates the records and provisions **free HTTPS**.

3 · Make the CTA actually receive mail (since the CTA is email)
- Domain ? **Email ? Email Routing ? Enable** ? add address `hello@khelsutra.guru` ? forward to `avi.dullu@gmail.com`. It auto-adds the MX/TXT records; confirm once in Gmail. Now every "Request access" click lands in your inbox. (Free.)

That's the whole path: **upload ? nameservers ? custom domain ? email routing.** ~15 minutes, all free tier.

Two notes

- **Where the file lives:** I put it in `artifacts/landing/index.html` (untracked scratch area of the indexer repo). It really belongs in your private `khelsutra` brand repo, or its own tiny static-site repo if you want git-based auto-deploys. Say where and I'll move/commit it properly.
- **Doc declutter:** parked for now ? as the audit showed, there's essentially nothing terminal to archive, so the real lever there is deciding which not-started design docs to keep active. Happy to do that whenever; the landing page was the higher-value move.

Want me to also add a **favicon + social-share (OG) image**, a privacy-friendly **Cloudflare Web Analytics** snippet, or swap the `mailto:` for a proper **form** (Cloudflare Pages Functions / Formspree) so submissions are structured? Any of those I can do right now.

103. user (2026-06-19T17:27:42.004Z)

I want to see the list of incomplete/partial docs so I can take them to completion
thanks for understanding my excitement and pleae make this an amzing page and run vidual validations or gemini calls if you need!
I will be back with you to do godaddy login and do it tonightCofounder! Lets go LIVE!!

104. assistant (2026-06-19T17:29:07.269Z)

Love the energy, cofounder ? let's make it genuinely shippable tonight. Two things: I'll pull the **incomplete-docs list** for you to drive to done, and I'll **polish the page and actually validate it visually** (render + screenshot + iterate) like you authorized. Let me wire up the preview tooling first.

105. user (2026-06-19T17:29:13.290Z)

[]

106. assistant (2026-06-19T17:29:27.660Z)

No server yet ? I'll stand up a tiny static server for the page. Let me check for an existing launch config first so I don't clobber it.

107. user (2026-06-19T17:29:27.889Z)

No files found

108. assistant (2026-06-19T17:29:32.052Z)

```
{
  "version": "0.0.1",
  "configurations": [
    {
      "name": "landing",
      "runtimeExecutable": "python",
      "runtimeArgs": ["-m", "http.server", "8099", "--directory", "artifacts/landing"],
      "port": 8099
    }
  ]
}
```

109. user (2026-06-19T17:29:32.070Z)

File created successfully at: C:\Users\avidu\Projects\badminton-highlight-indexer\.claude\launch.json (file state is current in your context ? no need to Read it back)

110. user (2026-06-19T17:29:40.651Z)

```
{
  "serverId": "151c90b2-9747-4c46-90c8-2d99d79ecd0a",
  "port": 8099,
```

```
  "name": "landing",
  "reused": false
}
Server started successfully on port 8099.
```

111. assistant (2026-06-19T17:29:43.805Z)

Server's up. Let me capture the page (desktop) and check for console errors.

112. assistant (2026-06-19T17:30:15.735Z)

Strong start ? clean, on-brand, great hero type. Now I'll validate the rest and elevate it. Let me check for console errors and capture more of the page at true desktop width.

113. user (2026-06-19T17:30:16.693Z)

No console logs.

114. user (2026-06-19T17:30:17.345Z)

Viewport set to 1280x2700.

115. assistant (2026-06-19T17:31:53.195Z)

No console errors, and the full layout reads clean and professional. The one weak spot: the hero's right side is empty space with a faint graphic. The highest-impact upgrade is a **real product visual** there ? a highlight-reel mockup showing exactly what KhelSutra outputs (detected-rally timeline + play-time stats). Let me add that, plus a favicon, then re-validate.

116. user (2026-06-19T17:31:56.056Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

117. user (2026-06-19T17:32:05.955Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

118. user (2026-06-19T17:32:07.383Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

119. user (2026-06-19T17:32:21.179Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

120. assistant (2026-06-19T17:32:25.514Z)

Edits applied. Let me re-validate the new hero at desktop width, then check mobile responsiveness.

121. user (2026-06-19T17:32:26.595Z)

Viewport set to 1280x820.

122. assistant (2026-06-19T17:32:47.857Z)

The screenshot is showing the pre-edit page ? the loop chips still read the old text and there's no reel card, so the browser tab hasn't reloaded the updated file. Let me grab the eval tool to force a reload.

123. user (2026-06-19T17:32:53.458Z)

"reloading"

124. user (2026-06-19T17:33:00.438Z)

{"tagName":"div","text":"session.mp4 ? highlights.mp4\n? REEL\n?\nGenerated by Khelsutra.guru\nRALLIES DETECTED\n44\nrallies found\n18 min\nreal play time","className":"reel-card","id":"","styles":{"box-shadow":"rgba(0, 0, 0, 0.55) 0px 26px 60px -24px"},"boundingBox":{"x":676.296875,"y":233.578125,"width":473.1953125,"height":476.5234375}}

125. assistant (2026-06-19T17:33:05.654Z)

The reel card is live. Let me capture the new hero.

126. assistant (2026-06-19T17:33:16.004Z)

That's a genuine product hero now ? the reel mockup (detected-rally timeline + "44 rallies / 18 min" stats + the *"Generated by Khelsutra.guru"* watermark) shows exactly what the product outputs. Let me verify mobile, then capture the full page end-to-end.

127. user (2026-06-19T17:33:16.607Z)

Viewport set to 375x812 (mobile).

128. assistant (2026-06-19T17:33:40.411Z)

Mobile is clean ? hero stacks, big tappable buttons, type scales well. Now let me validate the lower sections (roadmap, FAQ, the final CTA band, footer) at desktop width.

129. user (2026-06-19T17:33:41.837Z)

Viewport set to 1280x920.

130. user (2026-06-19T17:33:45.551Z)

0

131. assistant (2026-06-19T17:33:56.595Z)

Smooth-scroll is deferring the jump. Let me force an instant scroll and confirm the offset.

132. user (2026-06-19T17:33:57.967Z)

[
  2652,
  5008
]

133. assistant (2026-06-19T17:34:08.715Z)

Roadmap and the honest "what it's not" section read cleanly. Last check ? the final CTA band and footer.

134. user (2026-06-19T17:34:09.490Z)

4089

135. assistant (2026-06-19T17:35:17.001Z)

Validated end-to-end ? desktop, mobile, and every section, **zero console errors**. The page is genuinely shippable now.

The page is amazing-ready ?

What I added and verified visually (live in your preview panel):
- A **product hero**: a highlight-reel mockup (detected-rally timeline, a play button on a court, **44 rallies / 18 min** stats, and the *"Generated by Khelsutra.guru"* watermark) ? it shows what KhelSutra actually outputs instead of empty space.
- Favicon (shuttlecock), and I confirmed the hero stacks cleanly on **mobile (375px)**, the roadmap/FAQ/CTA/footer all compose well, and the `hello@khelsutra.guru` CTA is wired with a prefilled academy/city/courts email template.

Incomplete / partial docs ? your "take to completion" list

The audit found **14 genuinely partial (in-progress)** docs. These are the ones with real work left (not the 11 *not-started* designs, which are a separate "start or park" decision):

Doc	Where it is	To finish	Who
`UI_ALPHA_EXECUTION_PLAN`	PR-1/PR-2 merged	**PR-3 identity** ? invite flow ? alpha-ready	Me + you
`RALLY_DETECTION_GAP_CLOSURE`	G1/G2 addressed	close **G3 rally-END** on golden set, measured	Me + you (gated)
`PROMINENT_COURT_DETECTION`	design+experiment, OFF	build multi-track court-aware selection ? measure ? flip	Me + you (gated)
`MULTI_SIGNAL_FUSION_PLAN`	P1 (person) done	P2 fusion / P3 learned / P4 audio	Me (gated)
`EXPERIMENT_HARNESS`	P1 implemented	remaining promotion/shadow phases	Me
`RALLY_QUALITY_RESEARCH`	tracked §7	finish R-tiers + record the final quality number	Me + you
`R2_EVAL_METRIC_DESIGN`	implemented, locked	confirm remaining items or mark done	Me
`PERSONALIZATION_PLAN`	Phase-0 (#141/#142)	the per-user tuning layer + flywheel	you decide ? me
`GOLDEN_SET_VENDORING`	GS-1/2/4 landed	**GS-3 + gold key** pending	you + me
`GOLDEN_SET_IMPLEMENTATION`	older (May)	reconcile w/ built pipeline or fold into `GOLDEN_DATA_SHARING`	Me
`GOLDEN_REGRESSION_FIXTURES`	in-progress, legal OK	implement freeze + dual gates + sealed-fixtures filter	Me (gated)
`EXECUTION_POLICY`	P1/P2/P3a done	ship **P3b**, mark done	Me
`VIDEO_RECORDING_GUIDELINES`	**placeholder**	write the real 1-page capture checklist	**Me ? quick win**
`POST_166_WORK_ITEMS?`	pre-#166 plan	likely **superseded by NEXT_STEPS** ? reconcile/close	you decide

Two are fast, high-value wins I can knock out anytime: **`VIDEO_RECORDING_GUIDELINES`** (the landing page literally promises a "one-page capture checklist" ? closing this doc *is* that asset), and reconciling **`POST_166`** (we're at #244 now; it's probably just archivable). Pick any row and I'll produce a per-doc completion plan or just drive it to done.

Tonight's go-live runbook (GoDaddy ? Cloudflare)

When you're back, ~15 min, all free:
1. **Cloudflare** ? Add site `khelsutra.guru` (Free plan) ? it gives you **2 nameservers**.
2. **GoDaddy** ? your domain ? **Nameservers ? Change ? "I'll use my own"** ? paste Cloudflare's 2 NS ? save. (This is the one GoDaddy step; propagates in minutes?hours.)
3. **Cloudflare ? Workers & Pages ? Create ? Pages ? Upload assets** ? drag the `landing/` folder ? Deploy.
4. Pages project ? **Custom domains** ? add `khelsutra.guru` + `www` ? auto records + free HTTPS.
5. Cloudflare ? **Email ? Email Routing** ? `hello@khelsutra.guru` ? forward to your Gmail. Now the CTA lands in your inbox.

I'll be on hand to walk each click and fix anything live.

Want me to also, before tonight: generate a proper **social-share (OG) image** (so links to khelsutra.guru preview nicely), drop in privacy-friendly **Cloudflare Web Analytics**, and knock out the **`VIDEO_RECORDING_GUIDELINES`** one-pager so the page's "capture checklist" promise is real? I can do all three now.

136. user (2026-06-19T18:43:39.249Z)

Can you refine this to show a feature that "we can take an input match and give you segments of the video indexed as rallies and killing all the "dead time", you can arrange and stictch them using a conversational input form like give me a reel of all videos which lasted over 10 seconds etc." .. Do you see the PMF for this use case so maybe make that the index upon?

137. assistant (2026-06-19T18:46:27.443Z)

Yes ? I see real PMF here, and it's worth locating precisely (I'll be honest, not hype):

The PMF is in the *searchable rally index*, and conversational input is the right interface for it. Your validated pain is "hours of footage nobody watches; hand-cutting is tedious." The wedge isn't "an auto highlight reel" (every AI clipper claims that) ? it's **turning a match into a queryable index of rallies with the dead time stripped, then composing any reel on demand.** "Give me every rally over 10 seconds" is a killer demo because it's *self-serve editing without an editor*. Two honest caveats I'd bake into the page so we don't overpromise:
1. **The moat is the index, not the NL parser.** Natural language is cheap to build over a known schema; the durable value is rally segmentation quality (your actual hard-won asset). So I'll lead with "searchable rallies" and treat "just ask" as the delightful front door.
2. **Scope the queries to metadata you actually have today** ? length, count, order, which game. "Rallies where X smashed" needs shot/player detection, which is your roadmap. I'll keep the page honest about that line.

That's coherent with your architecture too ? you already persist candidate segments, so a "rally index ? query ? stitch" product *is* the natural surface over the segmentation engine. Good instinct to make it the anchor.

Let me re-cut the page around it and re-validate visually.

138. user (2026-06-19T18:46:32.179Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

139. user (2026-06-19T18:46:34.942Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

140. user (2026-06-19T18:46:37.881Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

141. user (2026-06-19T18:46:46.741Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

142. user (2026-06-19T18:46:54.259Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

143. user (2026-06-19T18:47:07.916Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

144. user (2026-06-19T18:47:10.725Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has

been updated successfully. (file state is current in your context ? no need to Read it back)

145. assistant (2026-06-19T18:47:14.515Z)

All edits applied. Let me reload and validate the new hero + the new "ask" centerpiece at desktop width.

146. user (2026-06-19T18:47:15.523Z)

Server not found. No running servers for this workspace.

147. assistant (2026-06-19T18:47:19.939Z)

The preview server stopped. Let me restart it and re-validate.

148. user (2026-06-19T18:47:23.245Z)


```
{
  "serverId": "22b93f63-371e-4f6b-9cc8-8ebd20f0891c",
  "port": 8099,
  "name": "landing",
  "reused": false
}
```

Server started successfully on port 8099.

149. user (2026-06-19T18:47:27.665Z)

Viewport set to 1280x860.

150. assistant (2026-06-19T18:47:39.999Z)

Page loaded but scrolled to the bottom. Let me jump to the top and capture the new hero.

151. user (2026-06-19T18:47:41.067Z)


```
[
  "KhelSutra ? index every rally, then build any reel just by asking",
  true,
  true
]
```


152. assistant (2026-06-19T18:48:10.083Z)

Content's all correct ? new headline, the conversational reel-builder card, the ask?matched-timeline?reel result ? but that capture rendered at an odd scale. Let me reset the viewport cleanly and re-shoot.

153. user (2026-06-19T18:48:10.243Z)

Viewport reset to native size (desktop).

154. assistant (2026-06-19T18:48:29.532Z)

The "desktop" preset reset to a ~800px native width, which stacks the hero (below my 860px breakpoint). Let me set an explicit wide viewport to see the two-column layout.

155. user (2026-06-19T18:48:30.184Z)

Viewport set to 1280x880.

156. assistant (2026-06-19T18:48:57.510Z)

Two-column hero confirmed: headline + the conversational reel-builder card (ask ? matched/dim

timeline ? "Your reel · 6 rallies · 1:32" result). The wide viewport renders small in the capture, so let me read the new **#ask centerpiece** and reframed cards at native width.

157. user (2026-06-19T18:48:57.998Z)

Viewport reset to native size (desktop).

158. user (2026-06-19T18:49:01.842Z)

1095

159. assistant (2026-06-19T18:49:22.706Z)

Scroll math landed on the FAQ. Let me target the #ask heading directly.

160. user (2026-06-19T18:49:23.751Z)

```
[
  1087,
  "Ask for the reel you want ? in plain English"
]
```

161. assistant (2026-06-19T18:50:01.857Z)

The **#ask centerpiece** renders beautifully ? *"The part people love ? Ask for the reel you want ? in plain English"* with the query?result example cards. The refinement is done and validated. Let me do a final console check and capture the positioning decision in memory.

162. user (2026-06-19T18:50:02.810Z)

No console logs.

163. user (2026-06-19T18:50:08.612Z)

The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\khelsutra-domain-launch.md has been updated successfully. (file state is current in your context ? no need to Read it back)

164. assistant (2026-06-19T18:50:33.765Z)

Done ? the page is now anchored on that use case, validated visually end-to-end, **zero console errors**.

What changed (all live in your preview)

- **New headline:** *"Turn a full match into searchable rallies ? and build any reel just by asking."*
- **Hero visual is now a conversational reel-builder:** a query bar (*"Give me a reel of every rally over 10 seconds"*) ? the rally index timeline with the **6 matches lit green and the rest dimmed** ? a composed **"Your reel · 6 rallies · 1:32 · dead time removed"** result bar. It tells the whole story in one glance.
- **New centerpiece section** ("The part people love ? *Ask for the reel you want ? in plain English*") with three example query?result cards, plus an **honest scope line**: today you filter by length / count / order / which game; shot-type and player queries are roadmap.
- **Reframed "what it does today"** cards: *Every rally indexed · Dead time removed · Ask in plain English · Stitched reel + summary*, and a new FAQ entry "What can I ask for?".

My PMF read, as your cofounder

Yes ? and this is the sharpest framing of the product yet. Three things make me believe it:

1. **It productizes your actual moat.** You already persist candidate rally segments ? a *queryable index ? stitch* surface is the natural product over the segmentation engine you've spent months sharpening. You're selling your strength, not a generic "AI clipper."
2. **The index is the wedge; language is the door.** The durable value is rally-detection quality (especially multi-court); NL is the friendly front-end that makes it self-serve for a

non-technical coach. So: invest in the index, treat the chat as delight ? and never let the copy promise a query the index can't answer yet.
3. **It strengthens your GTM loop.** "Ask for your reel" is a far better field-visit demo than "here's an auto reel," *and* every query result is a natural moment for a coach to say "that one's not a rally" ? which is exactly the labelled-correction flywheel your whole strategy depends on.

The one discipline I'd hold: keep the alpha's queries to **length / count / order / game** (what the index actually knows) until shot/player detection lands. That keeps the magic honest.

This also tells you what to build next in the UI alpha: the **query ? filter ? stitch** surface over the rally index ? the data's already there.

Tonight stays simple

The go-live runbook is unchanged (GoDaddy ? move nameservers to Cloudflare ? Pages upload ? custom domain ? email routing). I've noted GoDaddy as your registrar and saved the positioning so we don't lose it. The `index.html` is ready to drag in.

Want me to knock out the three pre-tonight extras now ? **OG share image**, **Cloudflare Web Analytics**, and the **one-page capture checklist** (`VIDEO_RECORDING_GUIDELINES`, which the page references)? I can do all three before you're back for the GoDaddy login. Let's go live ?

165. user (2026-06-19T18:54:26.178Z)

Can you load the website tab in this window?

166. assistant (2026-06-19T18:54:53.060Z)

Let me check the preview server is still up and then make sure it's loaded.

167. user (2026-06-19T18:54:53.143Z)


```
[
  {
    "serverId": "22b93f63-371e-4f6b-9cc8-8ebd20f0891c",
    "name": "landing",
    "port": 8099,
    "status": "running",
    "startedAt": "2026-06-19T18:47:20.214Z",
    "cwd": "C:\\Users\\avidu\\Projects\\badminton-highlight-indexer"
  }
]
```


168. assistant (2026-06-19T18:55:00.906Z)

Server's running. Let me reset it to the top (it was scrolled down) so it opens on the hero.

169. user (2026-06-19T18:55:01.689Z)


```
{
  "title": "KhelSutra ? index every rally, then build any reel just by asking",
  "url": "http://localhost:8099/"
}
```


170. assistant (2026-06-19T18:55:13.927Z)

The website is live and loaded ? it's being served at **http://localhost:8099** and should be showing in the **Launch preview panel** in this window (the pane that opened when I started the preview server). I just reset it to the top, so it'll open on the hero.

If the panel isn't visible:
- Look for the **preview/Launch panel tab** in this Claude Code window ? the `landing` server (port 8099) is running and rendering `artifacts/landing/index.html`.

- Or open **http://localhost:8099** directly in any browser on this machine.

You're looking at the live, refined page ? searchable-rally-index hero, the conversational reel-builder card, and the "ask for the reel you want" centerpiece. Scroll through and tell me what you'd tweak ? copy, colors, the example queries, anything. And whenever you're ready for the GoDaddy login tonight, I'm here to walk it click-by-click. ?

171. user (2026-06-19T18:55:15.871Z)

And yes please do all the three and tell me when do we login to my godaddy and do your thing?

172. assistant (2026-06-19T18:56:30.652Z)

Love it ? let me knock out all three now, then give you the exact GoDaddy timing. First, checking what's available for generating a real OG image and reading the checklist doc I'll complete.

173. user (2026-06-19T18:56:33.038Z)

```
PIL 12.2.0
--- fonts ---
/c/Windows/Fonts/arialbd.ttf
/c/Windows/Fonts/ariblk.ttf
/c/Windows/Fonts/segoeui.ttf
/c/Windows/Fonts/segoeuib.ttf
/c/Windows/Fonts/segoeuii.ttf
/c/Windows/Fonts/segoeuil.ttf
/c/Windows/Fonts/segoeuisl.ttf
/c/Windows/Fonts/segoeuiz.ttf
```

174. user (2026-06-19T18:56:33.107Z)

```
1       # Video Recording Guidelines (DRAFT ? thresholds pending empirical data)
2
3       **Status:** ? Draft / placeholder · **Date:** 2026-05-30
4       **Why this exists:** highlight quality is bounded by *input* quality. To get reliably
good auto-generated highlights, the source video must be recorded a certain way (camera height,
court position, resolution, frame rate, stability, framing). We want to (1) **derive** the
minimum acceptable capture spec **empirically** from the precision/recall of the solutions we
build, and (2) **publish user-facing best practices** so a user knows how to record footage that
the product can index well.
5
6       > **This is a tracked thought, not a finished standard.** The exact numeric thresholds
below are **placeholders** to be replaced with values measured against our golden set. See "How
we'll set the real numbers" (§3). Captured now (per the team's working agreement) so it isn't
lost; to be revisited.
7       >
8       > **Update (2026-06-02):** the **frame-rate** and **camera-geometry** rows now carry
**literature-backed starting values** (cited inline). They remain *starting points* to confirm
empirically per §3. The golden-set **corpus** diversity spec (what footage to collect, not just
how users should record) is now in [GOLDEN_SET_PHASE1_VIDEOS.md](GOLDEN_SET_PHASE1_VIDEOS.md).
9       >
10      > **Update (2026-06-07):** **AUDIO is now a first-class capture requirement** (new row
below), per ADR-007 ? we will use the shuttle "thwack" as a rally cue (audio+visual fusion lifts
recall). This is a hard "record with audio on, mic unobstructed" ask, not a placeholder;
efficacy on real GoPro audio is being validated (experiment E1 in
[AUDIO_RALLY_DETECTION_PLAN.md](archives/research/AUDIO_RALLY_DETECTION_PLAN.md)).
11
12      ---
13
14      ## ?? THE #1 CAPTURE RULE: ONE MATCH PER FRAME (single court)
15
16      > **For the best auto-highlights, record a SINGLE match on a SINGLE court ? with no
other matches
```

> playing in the background or on adjacent courts.** This is the happy path the product is tuned for,
> and as of **2026-06-18 it is MEASURED, not a guess:**
>
> - On a golden **multi-court** clip, the rally detector fired **87 "rallies" against 43 real ones ?
> precision ? 0.20 (~80% false positives)**, almost all of them background-court games it mistook
> for play.
> - On **single-court** clips, precision is markedly higher and rally boundaries are tight; local even
> **out-recalls Gemini** on normal single-match play.
>
> **Multi-court is a KNOWN BOTTLENECK we are actively working to fix** (player-state + court-occupancy
> signals ? see RALLY_DETECTION_GAP_CLOSURE.md). Until that lands,
> **single-match capture is strongly recommended** ? a reel made from busy multi-court footage *will*
> include false "rallies" from the neighbouring games. **North star (for now): consistently beat Gemini
> on the single-match happy path.**

---

1. Why input capture matters (grounded in what we already know)

ADR-004 found a concrete, capture-dependent failure mode: on **broadcast** footage the camera *tracks the players*, so player pixel-motion dips during rallies and spikes on pans/replays ? making player-motion a poor rally proxy. The lesson generalizes: **the pipeline's accuracy is conditional on how the video was shot.** A **static, full-court** camera (typical of a fixed tripod recording) behaves very differently from broadcast ? and is likely much friendlier to both the motion/detector stages and the shuttle-tracking substrate (WASB/TrackNet, ADR-001).

So capture guidance isn't cosmetic ? it directly moves precision/recall.

---

2. The dimensions that likely matter (hypotheses to test)

Dimension	Why it plausibly affects P/R	Placeholder guidance (UNVERIFIED)
Resolution	Shuttle is tiny & fast; too few pixels ? detector misses it	? **1080p** preferred; **720p** likely floor; 2K/4K may help shuttle tracking but ? compute/cost
Frame rate	Motion blur; boundary precision; fast-ball detection	? **30 fps** badminton/pickleball; **? 60 fps table tennis** ? 30 fps collapses fast-ball detection (**22.5% ? 93.2%** at 60 fps; 120 ? 60, [Nature Sci Rep](https://www.nature.com/articles/s41598-023-42966-6))
Camera position	Static full-court view avoids the broadcast camera-tracking trap (ADR-004)	**Fixed tripod**, elevated, behind/above one baseline, whole court in frame
Camera height & angle	Occlusion of players/shuttle; court keystoning	**3?5 m** elevated (~4.5 m), **~30° down-tilt**, behind one baseline, whole court + net headroom ([badminton-CV setup](https://pmc.ncbi.nlm.nih.gov/articles/PMC10747833/))
Framing / margins	Cropping the court or players breaks detection & relevance check	Entire court + a margin; players never leave frame
Stability	Handheld jitter inflates motion signals	Tripod / fixed mount; no panning or zoom during play
Lighting / exposure	Blur & contrast affect detection and the existing blur/relevance validators	Even indoor lighting; avoid backlight/glare; minimize motion blur
Single match in frame (no background games) ? ? MEASURED, the #1 rule	Adjacent/background-court games get detected as false rallies ? this is the **dominant precision failure**	**One match per frame, single court.** Multi-court precision measured at **?0.20** (vs ?0.31 single, tighter boundaries); multi-court fix in progress ? see RALLY_DETECTION_GAP_CLOSURE.md

54 | **Court contrast / background** | Relevance check uses court color; busy crowds add false motion | Clear court markings; minimal moving background where possible |
55 | **Codec / container** | Pipeline already validates MP4 + decodes H.264/HEVC | MP4 (H.264/HEVC) as today; align with `archives/MONETIZATION_AUDIT.md` codec notes |
56 | **Audio (REQUIRED ? ADR-007)** | Shuttle "thwack" hits are a strong rally cue; audio+visual fusion lifts rally *recall* (the WASB bottleneck) ? see ADR-007 / archives/research/AUDIO_RALLY_DETECTION_PLAN.md | **Record with AUDIO ENABLED** (GoPro default ? do not mute). Keep the **mic unobstructed**; **avoid aggressive wind/noise reduction** that flattens the sharp impact transient. Camera reasonably close to the court helps SNR (shuttle hits are quiet). |
57
58   These map onto checks that **already exist** in `backend/pipeline/validators/` (resolution, fps, blur, scene-cut, court relevance) ? so several can be enforced/measured programmatically rather than only advised.
59
60   ---
61
62   ## 3. How we'll set the REAL numbers (empirical plan)
63
64   Replace every placeholder above with a value derived from data, using the golden-set harness:
65
66   1. Take golden-set clips (GOLDEN_SET_VENDORING.md) spanning a **range** of each dimension ? e.g. the same match re-encoded at 480p/720p/1080p/4K, or clips at different camera heights.
67   2. Run the production pipeline; score with `backend/eval/metrics.py`.
68   3. Plot **precision/recall vs. each dimension** ? find the knee where quality collapses.
69   4. Set the **published minimum** at (knee + safety margin); set a **"recommended"** tier above it.
70   5. Re-derive when the pipeline changes materially (new detector/segmenter shifts the knees).
71
72   This turns "record in 1080p" from an assertion into a measured threshold ? and lets us tell users *why*.
73
74   ---
75
76   ## 4. Two surfaces for the output
77
78   1. **User-facing best-practices guide** ? short, friendly, illustrated: "Put the camera here, this high, record at ?X, keep it still." The goal is that a user can record once and get great results.
79   2. **Programmatic pre-flight check** ? extend `backend/pipeline/validators/` to *reject or warn* on out-of-spec uploads at ingest (resolution/fps/blur/framing), so the product fails fast with actionable feedback instead of producing poor highlights silently.
80
81   ---
82
83   ## 5. Open questions
84   1. Which dimension dominates ? resolution, camera staticness, or height? (Run the ablations in §3 to rank them.)
85   2. Do we **enforce** minimums (reject uploads) or only **warn**? Likely warn first, enforce once thresholds are trustworthy.
86   3. Single fixed-camera spec, or several supported "recording profiles" (e.g. baseline-elevated vs. side-elevated)?
87   4. How does the owned/licensed golden corpus need to grow to *cover* the capture conditions real users will actually use? (Feeds GOLDEN_SET_VENDORING.md §8.)
88

175. assistant (2026-06-19T18:57:49.389Z)

PIL 12.2 and Segoe UI fonts are available ? I can generate a real OG PNG. Important catch: `VIDEO_RECORDING_GUIDELINES.md` isn't a placeholder ? it's a solid *internal* empirical-derivation doc. The landing page promises a **user-facing** one-pager, which is a different (missing) asset. So I'll build that printable checklist rather than clobber the good internal

doc. Doing all three now:

176. assistant (2026-06-19T18:58:02.956Z)

```
"""Generate og.png (1200x630) ? the social-share card for khelsutra.guru.
Re-run after copy/brand tweaks:  python artifacts/landing/make_og.py
"""
from PIL import Image, ImageDraw, ImageFont

W, H = 1200, 630
INK, INK2 = (12, 38, 31), (20, 62, 47)
GREEN, LIME = (25, 179, 107), (199, 242, 176)
WHITE, MUT = (234, 243, 238), (167, 195, 183)
F = "C:/Windows/Fonts/"

img = Image.new("RGB", (W, H))
d = ImageDraw.Draw(img)
```

vertical gradient background

```
for y in range(H):
    t = y / H
    d.line([(0, y), (W, y)], fill=tuple(int(a + (b - a) * t) for a, b in zip(INK, INK2)))

def font(name, size):
    return ImageFont.truetype(F + name, size)

bold = lambda s: font("segoeuib.ttf", s)
reg = lambda s: font("segoeui.ttf", s)

M = 80  # margin
```

--- wordmark + shuttle mark ---

```
cx, cy = M + 16, 112
d.ellipse([cx - 17, cy - 1, cx + 17, cy + 33], fill=LIME)        # cork
for dx in (-22, -8, 8, 22):
    d.line([(cx, cy), (cx + dx, cy - 46)], fill=LIME, width=5)   # feathers
d.text((M + 52, 92), "Khel", font=bold(40), fill=WHITE)
khel_w = d.textlength("Khel", font=bold(40))
d.text((M + 52 + khel_w, 92), "Sutra", font=bold(40), fill=LIME)
```

--- headline ---

```
d.text((M, 232), "Search your rallies.", font=bold(78), fill=WHITE)
d.text((M, 322), "Build reels by asking.", font=bold(78), fill=LIME)
```

--- subline ---

```
d.text((M, 436), "Turn a full match into searchable rallies ? the dead time stripped out.",
       font=reg(30), fill=MUT)
```

--- badge pill ---

```
pill = "Invite-only alpha · Bengaluru"
pf = bold(24)
pw = d.textlength(pill, font=pf)
d.rounded_rectangle([M, 506, M + pw + 44, 552], radius=23, fill=(199, 242, 176, 255),
                    outline=None)
```

darker lime text on lime pill

```
d.text((M + 22, 512), pill, font=pf, fill=(10, 70, 45))
```

--- domain bottom-right ---

```
dom = "khelsutra.guru"
df = bold(28)
```

```
dw = d.textlength(dom, font=df)
d.text((W - M - dw, 520), dom, font=df, fill=LIME)
```

--- rally-index timeline motif (top-right) ---

```
bars = [(0, 70, GREEN), (84, 52, (70, 92, 82)), (150, 96, GREEN),
        (262, 44, (70, 92, 82)), (322, 78, GREEN)]
bx, by = W - M - 400, 150
for off, wbar, col in bars:
    d.rounded_rectangle([bx + off, by, bx + off + wbar, by + 16], radius=8, fill=col)

img.save("artifacts/landing/og.png")
print("wrote artifacts/landing/og.png", img.size)
```

177. user (2026-06-19T18:58:03.014Z)

```
File created successfully at: C:\Users\avidu\Projects\badminton-highlight-
indexer\artifacts\landing\make_og.py (file state is current in your context ? no need to Read it
back)
```

178. user (2026-06-19T18:58:08.608Z)

```
wrote artifacts/landing/og.png (1200, 630)
-rw-r--r-- 1 avidu 197609 43813 Jun 20 00:28 artifacts/landing/og.png
```

179. assistant (2026-06-19T18:58:13.301Z)

```
OG image generated. Let me eyeball it.
```

180. assistant (2026-06-19T18:58:39.249Z)

```
OG card looks sharp ? wordmark + shuttle mark, the rally-index bars, "Search your rallies. /
Build reels by asking." Now the printable capture checklist, then I'll wire OG tags + analytics
into the page.
```

181. assistant (2026-06-19T18:59:10.936Z)

```
<!doctype html>
<html lang="en">
<head>
<meta charset="utf-8" />
<meta name="viewport" content="width=device-width, initial-scale=1" />
<title>KhelSutra ? 1-page capture checklist</title>
<meta name="description" content="How to record a badminton session so KhelSutra gives you great
highlight reels: camera placement, settings, stability, lighting." />
<link rel="icon" href="data:image/svg+xml,%3Csvg xmlns='http://www.w3.org/2000/svg' viewBox='0 0
32 32'%3E%3Ccircle cx='16' cy='22' r='6' fill='%2319B36B'/%3E%3Cpath d='M16 16 L9 3 M16 16 L16 2
M16 16 L23 3' stroke='%230C261F' stroke-width='2.6' stroke-linecap='round'/%3E%3C/svg%3E" />
<link rel="preconnect" href="https://fonts.googleapis.com" />
<link rel="preconnect" href="https://fonts.gstatic.com" crossorigin />
<link href="https://fonts.googleapis.com/css2?family=Plus+Jakarta+Sans:wght@400;500;600;700;800&
display=swap" rel="stylesheet" />
<style>
  :root{--ink:#0C261F;--green:#19B36B;--green-d:#137a4c;--paper:#F6F8F4;--tx:#14271F;--
mut:#566B61;--line:#E1E7DF;--font:'Plus Jakarta Sans',system-ui,-apple-system,Segoe
UI,Roboto,sans-serif}
  *{box-sizing:border-box}
  body{margin:0;font-family:var(--font);color:var(--tx);background:#E9EDE6;line-
height:1.5;-webkit-font-smoothing:antialiased}
  .sheet{max-width:820px;margin:18px auto;background:#fff;border-
radius:16px;overflow:hidden;box-shadow:0 18px 50px -24px rgba(0,0,0,.35)}
  .head{background:var(--ink);color:#EAF3EE;padding:22px 30px;display:flex;align-
items:center;justify-content:space-between;gap:1rem;flex-wrap:wrap}
  .brand{display:flex;align-items:center;gap:.55rem;font-weight:800;font-size:1.2rem}
```

```
  .brand .dot{color:var(--green)}
  .brand svg{width:26px;height:26px}
  .head h1{font-size:1.05rem;font-weight:600;margin:0;color:#C7F2B0}
  .body{padding:24px 30px 8px}
  .intro{color:var(--mut);margin:0 0 16px;font-size:.98rem}
  .rule{display:flex;gap:.7rem;background:#EAF6EF;border:1px solid #BfE6D2;border-
radius:12px;padding:14px 16px;margin-bottom:20px}
  .rule .star{color:var(--green-d);font-size:1.3rem;line-height:1}
  .rule b{color:var(--tx)}
  .rule p{margin:.25rem 0 0;color:var(--mut);font-size:.92rem}
  .grid{display:grid;grid-template-columns:1fr 1fr;gap:16px 26px}
  .step h3{display:flex;align-items:center;gap:.5rem;font-size:1rem;font-weight:700;margin:0 0
.4rem}
  .step h3 .n{width:24px;height:24px;border-
radius:7px;background:var(--green);color:#04140D;display:flex;align-items:center;justify-
content:center;font-size:.85rem;flex:none}
  .step ul{margin:0;padding-left:1.1rem}
  .step li{margin:.18rem 0;font-size:.92rem;color:#34453d}
  .step li b{color:var(--tx)}
  .cols{display:grid;grid-template-columns:1fr 1fr;gap:0;border:1px solid var(--line);border-
radius:12px;overflow:hidden;margin:22px 0 8px}
  .cols div{padding:14px 16px}
  .cols .do{background:#EAF6EF}
  .cols .dont{background:#Fbeeeb;border-left:1px solid var(--line)}
  .cols h4{margin:0 0 .5rem;font-size:.92rem}
  .cols .do h4{color:var(--green-d)}
  .cols .dont h4{color:#9a3b22}
  .cols ul{margin:0;padding-left:1.1rem}
  .cols li{font-size:.88rem;margin:.16rem 0;color:#3a4b43}
  .consent{font-size:.86rem;color:var(--mut);border-top:1px solid var(--line);margin-
top:6px;padding:14px 0 4px}
  .foot{display:flex;justify-content:space-between;flex-wrap:wrap;gap:.5rem;padding:14px 30px
24px;color:var(--mut);font-size:.85rem}
  .foot a{color:var(--green-d);text-decoration:none}
  @media (max-width:680px){.grid,.cols{grid-template-columns:1fr}.cols .dont{border-
left:0;border-top:1px solid var(--line)}}
  @media print{
    body{background:#fff}
    .sheet{box-shadow:none;margin:0;max-width:none;border-radius:0}
    @page{margin:12mm}
    .noprint{display:none}
  }
</style>
</head>
<body>
<div class="sheet">
  <div class="head">
    <span class="brand">
      <svg viewBox="0 0 32 32" fill="none" aria-hidden="true"><circle cx="16" cy="22" r="6"
fill="#19B36B"/><path d="M16 16 L9 3 M16 16 L16 2 M16 16 L23 3" stroke="#C7F2B0" stroke-
width="2" stroke-linecap="round"/></svg>
      Khel<span class="dot">Sutra</span>
    </span>
    <h1>1-page capture checklist</h1>
  </div>
  <div class="body">
    <p class="intro">Great highlights start with a steady, full-court recording. Five minutes of
setup makes a big difference to your reels.</p>

    <div class="rule">
      <span class="star" aria-hidden="true">?</span>
      <div>
        <b>The golden rule: one match, one court, in full frame ? and don't move the camera.</b>
        <p>Multi-court footage is the hard case we're still sharpening; a single court in frame
gives the best results today.</p>
```

```
      </div>
    </div>

    <div class="grid">
      <div class="step">
        <h3><span class="n">1</span>Place the camera</h3>
        <ul>
          <li>Behind <b>one baseline</b>, elevated ? a high tripod, a few metres up, or a
balcony.</li>
          <li>Tilt down about <b>30°</b>; keep the <b>whole court</b> in frame with a little
margin and net headroom.</li>
          <li>Players should never leave the frame.</li>
        </ul>
      </div>
      <div class="step">
        <h3><span class="n">2</span>Settings</h3>
        <ul>
          <li><b>1080p or higher</b>, shot in <b>landscape</b>.</li>
          <li><b>30 fps</b> for badminton/pickleball; <b>60 fps</b> for table tennis.</li>
          <li><b>Audio ON</b> (don't mute) ? keep the mic unobstructed, wind/noise reduction
off.</li>
        </ul>
      </div>
      <div class="step">
        <h3><span class="n">3</span>Keep it still</h3>
        <ul>
          <li>Fixed tripod or mount.</li>
          <li><b>No panning, no zooming</b> during play.</li>
        </ul>
      </div>
      <div class="step">
        <h3><span class="n">4</span>Light &amp; record</h3>
        <ul>
          <li>Even indoor lighting; avoid backlight, glare and very dark halls.</li>
          <li>Record the <b>whole session</b> and upload it as-is ? KhelSutra does the
cutting.</li>
        </ul>
      </div>
    </div>

    <div class="cols">
      <div class="do">
        <h4>Do</h4>
        <ul>
          <li>Fixed camera, full court, behind the baseline</li>
          <li>1080p+, landscape, audio on</li>
          <li>One match on one court</li>
        </ul>
      </div>
      <div class="dont">
        <h4>Avoid</h4>
        <ul>
          <li>Handheld or phone-in-hand</li>
          <li>Broadcast-style pans &amp; zooming</li>
          <li>Several courts in one frame; very dark halls</li>
        </ul>
      </div>
    </div>

    <p class="consent">Please record adults with their okay. Footage involving minors is never
used to improve the tool and is never published.</p>
  </div>
  <div class="foot">
    <span>Questions? <a href="mailto:hello@khelsutra.guru">hello@khelsutra.guru</a> · <a
href="https://khelsutra.guru">khelsutra.guru</a></span>
```

```
    <span class="noprint">Tip: press Ctrl/Cmd + P to print or save as PDF.</span>
  </div>
</div>
</body>
</html>
```

182. user (2026-06-19T18:59:11.010Z)

```
File created successfully at: C:\Users\avidu\Projects\badminton-highlight-
indexer\artifacts\landing\capture.html (file state is current in your context ? no need to Read
it back)
```

183. assistant (2026-06-19T18:59:16.414Z)

```
Now wiring the OG image tags, Cloudflare analytics, and a footer link to the checklist into the
page.
```

184. user (2026-06-19T18:59:19.766Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

185. user (2026-06-19T18:59:22.136Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

186. user (2026-06-19T18:59:24.972Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

187. assistant (2026-06-19T18:59:42.200Z)

```
All wired in. Let me verify the new assets are served and capture the checklist page.
```

188. user (2026-06-19T18:59:44.714Z)

```
GET /og.png:
  HTTP/1.0 200 OK
  Server: SimpleHTTP/0.6 Python/3.13.13
  Content-type: image/png
  Content-Length: 43813
GET /capture.html:
  HTTP/1.0 200 OK
  Server: SimpleHTTP/0.6 Python/3.13.13
  Content-type: text/html
  Content-Length: 6951
GET /index.html:
  HTTP/1.0 200 OK
  Server: SimpleHTTP/0.6 Python/3.13.13
  Content-type: text/html
  Content-Length: 29527
```

189. user (2026-06-19T18:59:45.465Z)

```
"navigating"
```

190. assistant (2026-06-19T18:59:49.897Z)

```
All three assets serve cleanly. Let me capture the checklist page.
```

191. user (2026-06-19T19:00:20.660Z)

preview_screenshot timed out after 30s. The preview window may be stuck (modal dialog, navigation hang, or unresponsive renderer). Check preview_console_logs for errors.

192. assistant (2026-06-19T19:00:28.440Z)

Screenshot timed out ? the renderer may be mid-navigation. Let me check its state.

193. user (2026-06-19T19:00:29.224Z)

```
[
  "complete",
  "/capture.html",
  "KhelSutra ? 1-page capture checklist",
  true
]
```

194. assistant (2026-06-19T19:00:35.977Z)

The checklist page loaded fine (eval responds, `.sheet` present) ? the screenshot was just a transient timeout. Retrying.

195. user (2026-06-19T19:01:06.367Z)

preview_screenshot timed out after 30s. The preview window may be stuck (modal dialog, navigation hang, or unresponsive renderer). Check preview_console_logs for errors.

196. assistant (2026-06-19T19:01:25.578Z)

The checklist page is verified loaded and served (200, correct title, `.sheet` present) ? the screenshot tool is just hanging after the cross-page navigation, not a page problem. Let me send the preview back to the landing page for you and confirm the index changes are intact via DOM check rather than fight the flaky screenshot.

197. user (2026-06-19T19:01:26.513Z)

"back-to-home"

198. user (2026-06-19T19:01:33.574Z)

```
{
  "captureLink": true,
  "cfBeacon": true,
  "heroH1": "Turn a full match into searchable rallies ? and build any reel just by asking.",
  "ogImage": "https://khelsutra.guru/og.png",
  "path": "/index.html",
  "twitterCard": "summary_large_image"
}
```

199. assistant (2026-06-19T19:02:17.194Z)

All three are done and verified live on the page:

- ? **OG share image** ? `og.png` (1200×630, serves 200): the wordmark + shuttle mark, the rally-index bars, and *"Search your rallies. / Build reels by asking."* (you saw it above). Wired up `og:image` + Twitter `summary_large_image` tags, so links to khelsutra.guru will preview nicely.
- ? **Cloudflare Web Analytics** ? privacy-friendly beacon added before `</body>` with a clearly-marked `REPLACE_WITH_CF_TOKEN` placeholder + instructions (or you can flip "Automatic setup" in Cloudflare and we delete the snippet).
- ? **Capture checklist** ? `capture.html`: a clean, **printable** one-pager (golden rule ? place the camera ? settings ? keep it still ? light & record ? do/avoid ? consent), linked from the footer. This makes the page's "capture checklist" promise real. (I left your internal `VIDEO_RECORDING_GUIDELINES.md` untouched ? it's the good empirical source doc, this is its

user-facing surface.)

Confirmed on the live page: `og:image` ?, twitter card ?, footer "Capture checklist" link ?, analytics beacon ?.

When do we log into GoDaddy? ? exactly one step, in the middle

Here's the full go-live sequence so you can see where GoDaddy fits (~15 min, we do it together):

#	Where	What
0	Cloudflare	Create a free account (if you don't have one)
1	Cloudflare	**Add a site** ? `khelsutra.guru` (Free plan) ? it gives you **2 nameservers**
2	**? GoDaddy**	**? THIS is the only GoDaddy login.** Your domain ? **Nameservers ? Change ? "I'll use my own"** ? paste Cloudflare's 2 ? Save
3	Cloudflare	**Workers & Pages ? Create ? Pages ? Upload assets** ? drag in `index.html`, `capture.html`, `og.png` ? Deploy (live `*.pages.dev` instantly)
4	Cloudflare	Pages ? **Custom domains** ? add `khelsutra.guru` + `www` ? auto DNS + free HTTPS
5	Cloudflare	**Email ? Email Routing** ? `hello@khelsutra.guru` ? forward to your Gmail
6	Cloudflare	(optional) **Analytics ? Web Analytics** ? copy token ? paste into the snippet (or use automatic)

So: GoDaddy login = Step 2 only ? a 2-minute nameserver paste. Everything else is on Cloudflare.

Honest bit on "do your thing": I can't log into your GoDaddy/Cloudflare for you ? those are your credentials + 2FA. But I'll be right here narrating every field as you click, and if you connect the **Chrome extension**, I can even drive the non-login navigation alongside you. The files are prepped, validated, and ready to drag in.

The trigger: when you're back at your computer tonight, just say **"let's go live"** (or "I'm in GoDaddy") and we'll walk it together, click by click. The only thing that's ever slow is DNS propagation ? and that runs in the background while we finish the rest. ?

200. user (2026-06-19T19:02:20.896Z)

and before that can you tell me whether you used the product vision, roadmap, current state etc. as context to put the information in it?
feedback on image Remove the shuttle because we will do more sports later and maybe do something witty with the characters in the name

201. assistant (2026-06-19T19:06:05.124Z)

Two separate things ? let me take the honesty question head-on first, because it matters more than the logo.

Did I ground the page in your real docs? Yes ? here's exactly what fed it

I deliberately built the copy from your own material, not from imagination:

- **`docs/KHELSUTRA_PRODUCT_OVERVIEW.md`** (your product vision / early-user guide) ? this is the backbone. The page's "what it does," "who it's for," "gets sharper from real footage," the privacy/minors guardrail, "where it's going," "what it's not," and the FAQ are all adapted near-line-for-line from it, including its **honest hedges** (invite-only, *"not a switch that's already on,"* minors never used, pricing not set).
- **`docs/NEXT_STEPS.md`** (roadmap) ? drove the "where it's going" items (shot detection ? assistant coach, more sports, languages) and the multi-court "hard problem" framing.
- **`current-state.md` + this session** ? the multi-court / rally-detection focus you're actively on.

Two places I extrapolated ? flagging them honestly (per our attribution rule):
1. **The hero mockup numbers** ("44 rallies," "18 min," "6 over 10s," "1:32 reel") are **illustrative placeholders**, not measured output. Easy to swap for a real example from one of your processed videos.

2. **The conversational reel-builder** came from *your message this turn*, not the product overview ? which lists today's shipped features as detection + reel + summary. So the page currently presents *"Ask in plain English"* as a **today** capability when it's really the near-term build we just scoped. **Your call:** mark it as "coming," or confirm it's live enough to keep as-is. I'd lean toward a small "in alpha / rolling out" tag on that one to stay honest.

Now the logo. Good instinct ? let me de-badminton it.

202. user (2026-06-19T19:06:09.866Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\make_og.py has been updated successfully. (file state is current in your context ? no need to Read it back)

203. user (2026-06-19T19:06:13.548Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

204. user (2026-06-19T19:06:15.423Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

205. user (2026-06-19T19:06:18.711Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

206. user (2026-06-19T19:06:22.484Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\capture.html has been updated successfully. (file state is current in your context ? no need to Read it back)

207. user (2026-06-19T19:06:23.896Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\capture.html has been updated successfully. (file state is current in your context ? no need to Read it back)

208. user (2026-06-19T19:06:26.311Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\capture.html has been updated successfully. (file state is current in your context ? no need to Read it back)

209. assistant (2026-06-19T19:06:30.862Z)

Shuttle removed everywhere. Let me regenerate the OG image and verify the new sport-agnostic mark.

210. user (2026-06-19T19:06:32.006Z)

wrote artifacts/landing/og.png (1200, 630)

211. assistant (2026-06-19T19:06:43.793Z)

Shuttle's gone ? clean lime **KS** monogram now, fully sport-agnostic (the abstract rally-index bars stay, since they read as "indexed segments," not badminton). Let me confirm the live header updated too.

212. user (2026-06-19T19:06:45.524Z)

"reloading"

213. user (2026-06-19T19:06:50.761Z)

{"tagName":"span","text":"KS","className":"mk","id":"","styles":{"background-color":"rgb(25, 179, 107)","color":"rgb(4, 20, 13)"},"boundingBox":{"x":22,"y":17.984375,"width":30,"height":30}}

214. assistant (2026-06-19T19:07:27.161Z)

Shuttle removed across the board ? **OG image, site header, favicon, and the capture page** ? all now use a clean lime **KS** monogram (verified live: the header mark renders green with "KS"). The abstract rally-index bars stay since they read as "indexed segments," not a sport.

The `KS` monogram is the *safe* default ? but you asked for **witty**, and I'd rather land that with your taste than guess. Both halves of the name are great raw material: **Khel = play**, **Sutra = thread/formula**. So I have a couple of genuinely clever, sport-agnostic directions. Two quick calls and I'll finalize:

215. user (2026-06-19T19:08:22.886Z)

Your questions have been answered: "Which logo direction should I lock in? (All are sport-agnostic and work at favicon size.)"="Play-triangle: Khel ? Sutra", "The page currently shows the conversational reel-builder as a 'today' feature, but it's really the near-term build. How should it read?"="Move it to 'where it's going'". You can now continue with these answers in mind.

216. assistant (2026-06-19T19:12:48.830Z)

The user picked **play-triangle "Khel ? Sutra"** for the logo and **move the conversational builder to "where it's going"** (lead with the index + dead-time, which is real today). Both are clear, honest calls. Let me do the logo first.

217. user (2026-06-19T19:12:53.085Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

218. user (2026-06-19T19:12:55.398Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

219. user (2026-06-19T19:12:57.758Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

220. user (2026-06-19T19:12:59.147Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has been updated successfully. (file state is current in your context ? no need to Read it back)

221. user (2026-06-19T19:13:02.985Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\capture.html has been updated successfully. (file state is current in your context ? no need to Read it back)

222. user (2026-06-19T19:13:05.293Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\capture.html has been updated successfully. (file state is current in your context ? no need to Read it back)

223. user (2026-06-19T19:13:07.170Z)

The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\capture.html has been updated successfully. (file state is current in your context ? no need to Read it back)

224. user (2026-06-19T19:13:11.713Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\make_og.py has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

225. assistant (2026-06-19T19:13:17.911Z)

```
Logo's done. Now the content reframe ? lead with the index + dead-time (real today), move the
conversational builder into a "coming next" spotlight.
```

226. user (2026-06-19T19:13:23.028Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

227. user (2026-06-19T19:13:25.386Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

228. user (2026-06-19T19:13:28.671Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

229. user (2026-06-19T19:13:32.002Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

230. user (2026-06-19T19:13:39.504Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

231. user (2026-06-19T19:13:42.798Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

232. user (2026-06-19T19:13:48.441Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

233. user (2026-06-19T19:13:54.085Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

234. user (2026-06-19T19:13:57.416Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\index.html has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

235. user (2026-06-19T19:14:00.408Z)

```
The file C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\landing\make_og.py has
been updated successfully. (file state is current in your context ? no need to Read it back)
```

236. assistant (2026-06-19T19:14:05.403Z)

```
All edits applied. Let me regenerate the OG with the new play-triangle mark + today headline,
```

and verify it.

237. user (2026-06-19T19:14:06.435Z)

wrote artifacts/landing/og.png (1200, 630)

238. assistant (2026-06-19T19:14:21.071Z)

The OG nails it ? **"Khel ? Sutra"** with the play-triangle (witty: *khel* = play), and the honest today headline *"Every rally, indexed. The dead time, gone."* Let me reload and verify the live page.

239. user (2026-06-19T19:14:22.523Z)

"reloading"

240. user (2026-06-19T19:14:29.218Z)


```
{
  "askAfterSharper": true,
  "askKicker": "Coming next to the alpha",
  "brand": "KhelSutra",
  "hasPlayTri": true,
  "heroH1": "Every rally, indexed. The dead time, gone.",
  "heroHasAskInput": false,
  "sectionOrder": [
    "how",
    "diff",
    "sharper",
    "ask",
    "roadmap",
    "not",
    "faq",
    "join"
  ]
}
```


241. assistant (2026-06-19T19:14:38.250Z)

Everything verified on the live page:
- **Logo:** play-triangle present in the header (`Khel ? Sutra`) ?
- **Hero now leads with today:** *"Every rally, indexed. The dead time, gone."* ?, and the hero visual no longer shows a live "ask" input ?
- **Conversational builder moved to "where it's going":** section order is now `how ? diff ? sharper ? ask ? roadmap`, and its kicker reads **"Coming next to the alpha"** ?

Let me grab one visual of the new hero to confirm it looks right.

242. user (2026-06-19T19:14:38.933Z)

"top"

243. user (2026-06-19T19:15:12.036Z)

preview_screenshot timed out after 30s. The preview window may be stuck (modal dialog, navigation hang, or unresponsive renderer). Check preview_console_logs for errors.

244. assistant (2026-06-19T19:15:18.870Z)

Screenshot tool's being flaky again, but the DOM verification already confirms every change landed correctly. Quick console-error check to be safe, then I'll wrap up.

245. user (2026-06-19T19:15:19.318Z)

No console logs.

246. assistant (2026-06-19T19:15:40.246Z)

All done and verified ? here's where the page landed:

Logo ? "Khel ? Sutra" (sport-agnostic + witty)

The shuttle is gone everywhere and replaced with the **play-triangle** between the words ? *Khel* means "play," so the mark literally reads as a play button. Applied consistently to the **OG image** (you saw it: "Khel ? Sutra"), the **site header**, the **favicon** (lime triangle on an ink square), and the **capture page**. Nothing badminton-locked now, so it carries cleanly into table tennis and beyond.

Conversational builder ? moved to "where it's going" (honest framing)

Per your call, the page now leads with what's **real today** and frames the ask-feature as **coming**:
- **Hero** now reads *"Every rally, indexed. The dead time, gone."* ? and the hero visual shows the index ? highlight-reel output (no live "ask" input).
- The conversational demo is now a **"Coming next to the alpha"** section, repositioned **after the today sections, just before the roadmap** (verified order: `how ? diff ? sharper ? ask ? roadmap`), with a *"Rolling out now"* note.
- The "what it does today" cards, the FAQ, and the title/OG copy are all reworded to today-true (index + dead-time-removal + reel), so nothing overpromises.

Verified live: play-triangle present ?, hero H1 updated ?, ask-section relabeled "Coming next" and relocated ?, **zero console errors** ?. (The screenshot tool got flaky, so I confirmed via direct DOM checks ? but it's all rendering in your **preview panel** at http://localhost:8099 if you want to scroll through.)

And to close the loop on your first question: yes ? the copy was grounded in your real docs (`KHELSUTRA_PRODUCT_OVERVIEW.md` + `NEXT_STEPS.md` + current state), and the two things I'd extrapolated (the mockup numbers, and the now-fixed "ask as today") are exactly what we just made honest.

Nothing about tonight changes ? the page, capture checklist, OG image, and analytics are all prepped and validated. Whenever you're back at your computer, say **"let's go live"** and we'll do the GoDaddy ? Cloudflare walk-through together. ?

247. user (2026-06-19T19:15:43.044Z)

Also I think we will move towards a desktop application or a UI based "point me to your online URL and let me stream it and support essentially the smae desktop features so that folks who use GCS buckets or Drive subscriptions, MS 365 etc can set that up easily and it "utosyncs their recordings and prepares highlights"

248. assistant (2026-06-19T19:17:06.660Z)

Now you're sketching the *delivery model* ? and it's a meaty one worth thinking through properly. Let me make sure I reflect it back right:

> **"Bring your own storage" + autosync.** Instead of manually uploading, a user points KhelSutra at a cloud location they already pay for ? a **GCS bucket, Google Drive, OneDrive/MS 365, Dropbox** ? and KhelSutra **watches that folder, pulls/streams new recordings as they land, and auto-prepares highlights**, with the same processing the desktop app does. Set-it-and-forget-it: footage shows up ? reels appear.

This lands squarely on top of your existing platform plans (the `StorageBackend` registry, the locality model, the desktop app). Rather than riff from memory, let me ground my take in what those docs already say so I'm not inventing your roadmap. Reading them in parallel.

249. user (2026-06-19T19:17:28.996Z)

Workflow launched in background. Task ID: w33cc3fg4

Summary: Assess how the "connect your cloud storage (GCS/Drive/OneDrive) + autosync ? highlights" idea fits the existing locality/platform/desktop/storage/commercial docs
Transcript dir: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\subagents\workflows\wf_684a0618-ccb
Script file: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\byo-storage-autosync-fit-wf_684a0618-ccb.js
(Edit this file with Write/Edit and re-invoke Workflow with {scriptPath: "C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\byo-storage-autosync-fit-wf_684a0618-ccb.js"} to iterate without resending the script.)
Run ID: wf_684a0618-ccb
To resume after editing the script: Workflow({scriptPath: "C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\byo-storage-autosync-fit-wf_684a0618-ccb.js", resumeFromRunId: "wf_684a0618-ccb"}) ? completed agents return cached results.

You will be notified when it completes. Use /workflows to watch live progress.

250. assistant (2026-06-19T19:17:40.034Z)

I'll synthesize a grounded cofounder assessment once the doc-mapping returns ? covering where this slots architecturally (it's essentially the `StorageBackend` registry + a sync daemon), the honest risks (compute placement, egress, OAuth scopes/verification, stream-vs-download), and the PMF/cost angle ? plus where to capture it. Standing by for the workflow.

251. user (2026-06-19T19:19:34.045Z)

<task-notification>
<task-id>w33cc3fg4</task-id>
<tool-use-id>toolu_01BLGyQA6HporBtpyVQyn8sZ</tool-use-id>
<output-file>C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\w33cc3fg4.output</output-file>
<status>completed</status>
<summary>Dynamic workflow "Assess how the "connect your cloud storage (GCS/Drive/OneDrive) + autosync ? highlights" idea fits the existing locality/platform/desktop/storage/commercial docs" completed</summary>
<result>{"findings":[{"scope":"docs/VIDEO_LOCALITY_MODEL.md (read in full, 148 lines); cross-checked its named companion docs/PLATFORM_ARCHITECTURE.md §3 to locate where BYO-storage actually lives.","anticipates":"no","relevant_excerpts":["§2 ladder framing: 'From most to least bytes leaving the academy's Windows desktop' ? every tier is about what crosses the network FROM the academy box TO our box, not video already living in the user's cloud.","L4 ? sneakernet (VIDEO_LOCALITY_MODEL.md §2, line 60): 'a hard drive | 0 over network | a person (the field associate / owner) | how the pilot actually works today' ? the closest thing to an external-storage source, but it is a person carrying a disk, not a watched cloud folder.","§0 line 25-26: 'expecting academies to manage S3/GCS buckets is a non-starter. The 10 GB problem is real and the current design doesn't solve it.' ? the ONLY mention of S3/GCS, and it's dismissive, not a tier.","OQ9 (line 135): 'Recurring volume ? an academy recording daily generates 10 GB/day; even L2 accumulates | retention (OQ1) + maybe future delta thinking; not an alpha problem' ? the doc names the recurring-footage cadence that autosync would serve, but explicitly defers it.","Header line 7: companions include 'PLATFORM_ARCHITECTURE.md §3 (future StorageBackend/ComputeBackend registries)' ? the doc EXPLICITLY delegates storage backends to that doc rather than modeling them here.","PLATFORM_ARCHITECTURE.md §3.1 lines 129-130: 'Backends: local (upload dir), gdrive, s3, gcs, azure. BYO-storage: the tenant registers their own bucket/Drive; raw video never leaves their account.' ? this is where BYO-cloud-storage actually lives, NOT in the locality doc.","PLATFORM_ARCHITECTURE.md lines 131-134: 'Auth ? never persist users' long-lived keys... have the user create an IAM role you AssumeRole into... mint presigned URLs... For Drive, OAuth with the narrow drive.file scope.' ? the auth/scope design the autosync idea would inherit."],"how_it_slots":"It does NOT slot cleanly into this doc's tiers as written. The locality ladder (L0..L4) classifies designs by 'smallest artifact that must cross OUR network from the academy desktop', whereas BYO-storage+autosync is a different axis entirely ? WHERE the canonical bytes live and HOW ingestion is triggered. The natural placement is a NEW dimension orthogonal to the ladder: a 'source backend / ingestion trigger' axis. Conceptually the user's-cloud-watched-folder is the AUTOMATED, networked cousin of L4

(sneakernet): same 'the original already lives somewhere we pull from, ~0 of OUR-network upload by the user' property, but the disk-in-a-bag is replaced by a watched Drive/S3/GCS folder and the 'a person' actor is replaced by a poll/webhook watcher. Its true architectural home is the StorageBackend registry in PLATFORM_ARCHITECTURE.md §3.1 (a gdrive/s3/gcs backend the tenant registers) PLUS a new 'watch/sync trigger' component feeding the (committed-next) async job model. The locality doc's central insight ? 'timestamps are the product; the video is dead weight' (§1) and 'cut the reel wherever the original lives' (§1 bullet 3) ? is exactly what makes BYO-storage viable: with their cloud as a compute-adjacent or signed-URL source, you stream/pull the proxy and never move the 10 GB, which is the same logic L2 already exploits. So the idea is highly COMPATIBLE with the doc's model, but the doc does not yet contain it as a tier.","gaps":["A 'source/ingestion' axis distinct from the existing 'what-crosses-OUR-network' ladder ? the doc has no tier for 'the original canonically lives in the user's own cloud'; the watched-cloud-folder is best framed as automated-L4, but that mapping isn't drawn.","Any notion of a WATCH/POLL/WEBHOOK trigger ('reels appear automatically'). The doc is entirely pull-on-demand / batch-upload; the words watch/poll/sync/automatic/webhook do not appear. The 'auto' in §5 ('the same motion L2 will automate') refers to scripting the manual proxy step, not unattended folder watching.","An identity rule for files DISCOVERED (not uploaded) in a bucket ? OQ2 fixes video_id as the processed-proxy MD5, but doesn't cover 'detect a new object in a watched prefix, dedupe, decide proxy-vs-original' on arrival.","Where the proxy-prep step runs when bytes start in the cloud, not on the academy desktop. The whole §2 ladder assumes the client-side ffmpeg prep happens on the Windows desktop that holds the file; with cloud-origin footage that compute placement (their VM? our GPU-less box pulling via signed URL? the 'zero-infra BYO' middle tier?) is unaddressed here.","Auth/scopes/credential custody for the watched account (OAuth drive.file, AssumeRole, presigned URLs) ? flagged in the companion PLATFORM doc but absent from this locality model, even though autosync makes long-lived watch credentials a first-class concern.","Egress/cost-on-pull accounting ? the §4 frugality frame counts GPU + API, not the recurring cloud-egress / repeated-poll cost of pulling new footage out of the user's bucket on every sync (and OQ9's 'even L2 accumulates' x daily-cadence gets worse under autosync).","Streaming-vs-download and format/codec variety from arbitrary user clouds (phone uploads, NAS auto-backup, mixed containers) ? the doc assumes a known academy recording the associate handles, not heterogeneous auto-arriving files."],"risks_considerations":["Compute placement: §1's 'cut the reel wherever the original lives' is the enabling principle, but with cloud-origin footage 'wherever' is now the user's cloud ? forcing the question of pulling-to-our-box (egress + GPU-less orchestration, §3) vs the 'zero-infra BYO' middle tier (signed URLs, our GPU) named only in the companion doc.","Egress + recurring-poll cost: pulling fresh footage out of GCS/Drive/S3 on every sync incurs egress the §4 frugality frame ('don't burn GPU') doesn't model; OQ9 (daily 10 GB) compounds under unattended autosync.","Gemini consent surface (OQ3): even 'their cloud' footage still ships ?1280px chunks to Google on the Gemini path ? autosync makes this happen automatically/unattended, raising the bar on plain-language consent.","Credential/OAuth scope + key custody for a persistent watcher ? long-lived watch tokens are exactly what PLATFORM §3.1 warns against ('never persist users' long-lived keys'); narrow drive.file / AssumeRole + presigned URLs needed.","Streaming-vs-download: large MP4s from a bucket want byte-range streaming (the companion doc reaches for smart_open / fsspec) rather than full download before proxy-prep ? a primitive this locality doc doesn't mention.","Format/codec variety + 'is this even a match?' filtering: auto-arriving files from arbitrary user storage are unvetted (wrong codec, non-badminton, partial uploads/in-flight writes), unlike the associate-curated L4 drive the pilot assumes.","Privacy/liability is actually IMPROVED if raw video never leaves the tenant's account (the companion doc's 'central privacy/liability lever') ? but only if processing uses signed-URL/their-compute rather than copying the 10 GB onto our box, which would negate the benefit.","Retention (OQ1) gets murkier: if the canonical copy lives in the user's cloud, our retention policy may only need to cover proxies/intervals ? a simplification the doc hasn't reasoned through."],"fit_verdict":"Strong conceptual fit but NOT yet anticipated here: VIDEO_LOCALITY_MODEL.md has no user-cloud tier and no watch/autosync notion ? BYO-storage+autosync belongs in PLATFORM_ARCHITECTURE.md §3.1's StorageBackend registry plus a new ingestion-trigger component, and would enter this doc as an 'automated-L4' source axis layered onto the async-job model."},{"scope":"docs/PLATFORM_ARCHITECTURE.md (§2 layered architecture, §3 StorageBackend + ComputeBackend registries, §5 orchestration/async jobs, §6 phased roadmap, §7 privacy); cross-checked against docs/DEFERRED_HARDENING_PLAN.md (#16 async job model, now marked COMPLETED, PR #87 + #149).","anticipates":"partial","releva
... (truncated 33142 chars, full result in C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\w33cc3fg4.output)</result>
<usage><agent_count>5</agent_count><subagent_tokens>286878</subagent_tokens><tool_uses>31</tool_uses><duration_ms>125001</duration_ms></usage>

</task-notification>

252. assistant (2026-06-19T19:19:39.583Z)

The mapping is back ? and the key finding is already clear: `PLATFORM_ARCHITECTURE.md §3.1` **already names exactly this** ("Backends: local, gdrive, s3, gcs, azure. BYO-storage: the tenant registers their own bucket/Drive; raw video never leaves their account"), but the **autosync/watch trigger** is the genuinely new piece. Let me read the full output before synthesizing.

253. user (2026-06-19T19:19:40.757Z)


```
1       {
2         "summary": "Assess how the \"connect your cloud storage (GCS/Drive/OneDrive) + autosync ? highlights\" idea fits the existing locality/platform/desktop/storage/commercial docs",
3         "agentCount": 5,
4         "logs": [],
5         "result": {
6           "findings": [
7             {
8               "scope": "docs/VIDEO_LOCALITY_MODEL.md (read in full, 148 lines); cross-checked its named companion docs/PLATFORM_ARCHITECTURE.md §3 to locate where BYO-storage actually lives.",
9               "anticipates": "no",
10              "relevant_excerpts": [
11                "§2 ladder framing: 'From most to least bytes leaving the academy's Windows desktop' ? every tier is about what crosses the network FROM the academy box TO our box, not video already living in the user's cloud.",
12                "L4 ? sneakernet (VIDEO_LOCALITY_MODEL.md §2, line 60): 'a hard drive | 0 over network | a person (the field associate / owner) | how the pilot actually works today' ? the closest thing to an external-storage source, but it is a person carrying a disk, not a watched cloud folder.",
13                "§0 line 25-26: 'expecting academies to manage S3/GCS buckets is a non-starter. The 10 GB problem is real and the current design doesn't solve it.' ? the ONLY mention of S3/GCS, and it's dismissive, not a tier.",
14                "OQ9 (line 135): 'Recurring volume ? an academy recording daily generates 10 GB/day; even L2 accumulates | retention (OQ1) + maybe future delta thinking; not an alpha problem' ? the doc names the recurring-footage cadence that autosync would serve, but explicitly defers it.",
15                "Header line 7: companions include 'PLATFORM_ARCHITECTURE.md §3 (future StorageBackend/ComputeBackend registries)' ? the doc EXPLICITLY delegates storage backends to that doc rather than modeling them here.",
16                "PLATFORM_ARCHITECTURE.md §3.1 lines 129-130: 'Backends: local (upload dir), gdrive, s3, gcs, azure. BYO-storage: the tenant registers their own bucket/Drive; raw video never leaves their account.' ? this is where BYO-cloud-storage actually lives, NOT in the locality doc.",
17                "PLATFORM_ARCHITECTURE.md lines 131-134: 'Auth ? never persist users' long-lived keys... have the user create an IAM role you AssumeRole into... mint presigned URLs... For Drive, OAuth with the narrow drive.file scope.' ? the auth/scope design the autosync idea would inherit."
18              ],
19              "how_it_slots": "It does NOT slot cleanly into this doc's tiers as written. The locality ladder (L0..L4) classifies designs by 'smallest artifact that must cross OUR network from the academy desktop', whereas BYO-storage+autosync is a different axis entirely ? WHERE the canonical bytes live and HOW ingestion is triggered. The natural placement is a NEW dimension orthogonal to the ladder: a 'source backend / ingestion trigger' axis. Conceptually the user's-cloud-watched-folder is the AUTOMATED, networked cousin of L4 (sneakernet): same 'the original already lives somewhere we pull from, ~0 of OUR-network upload by the user' property, but the disk-in-a-bag is replaced by a watched Drive/S3/GCS folder and the 'a person' actor is replaced by a poll/webhook watcher. Its true architectural home is the StorageBackend registry in PLATFORM_ARCHITECTURE.md §3.1 (a gdrive/s3/gcs backend the tenant registers) PLUS a new 'watch/sync trigger' component feeding the (committed-next) async job model. The locality doc's central insight ? 'timestamps are the product; the video is dead weight' (§1) and 'cut the reel wherever the original lives' (§1 bullet 3) ? is exactly what makes BYO-storage viable: with
```

their cloud as a compute-adjacent or signed-URL source, you stream/pull the proxy and never move the 10 GB, which is the same logic L2 already exploits. So the idea is highly COMPATIBLE with the doc's model, but the doc does not yet contain it as a tier.",
20 "gaps": [
21 "A 'source/ingestion' axis distinct from the existing 'what-crosses-OUR-network' ladder ? the doc has no tier for 'the original canonically lives in the user's own cloud'; the watched-cloud-folder is best framed as automated-L4, but that mapping isn't drawn.",
22 "Any notion of a WATCH/POLL/WEBHOOK trigger ('reels appear automatically'). The doc is entirely pull-on-demand / batch-upload; the words watch/poll/sync/automatic/webhook do not appear. The 'auto' in §5 ('the same motion L2 will automate') refers to scripting the manual proxy step, not unattended folder watching.",
23 "An identity rule for files DISCOVERED (not uploaded) in a bucket ? OQ2 fixes video_id as the processed-proxy MD5, but doesn't cover 'detect a new object in a watched prefix, dedupe, decide proxy-vs-original' on arrival.",
24 "Where the proxy-prep step runs when bytes start in the cloud, not on the academy desktop. The whole §2 ladder assumes the client-side ffmpeg prep happens on the Windows desktop that holds the file; with cloud-origin footage that compute placement (their VM? our GPU-less box pulling via signed URL? the 'zero-infra BYO' middle tier?) is unaddressed here.",
25 "Auth/scopes/credential custody for the watched account (OAuth drive.file, AssumeRole, presigned URLs) ? flagged in the companion PLATFORM doc but absent from this locality model, even though autosync makes long-lived watch credentials a first-class concern.",
26 "Egress/cost-on-pull accounting ? the §4 frugality frame counts GPU + API, not the recurring cloud-egress / repeated-poll cost of pulling new footage out of the user's bucket on every sync (and OQ9's 'even L2 accumulates' x daily-cadence gets worse under autosync).",
27 "Streaming-vs-download and format/codec variety from arbitrary user clouds (phone uploads, NAS auto-backup, mixed containers) ? the doc assumes a known academy recording the associate handles, not heterogeneous auto-arriving files."
28],
29 "risks_considerations": [
30 "Compute placement: §1's 'cut the reel wherever the original lives' is the enabling principle, but with cloud-origin footage 'wherever' is now the user's cloud ? forcing the question of pulling-to-our-box (egress + GPU-less orchestration, §3) vs the 'zero-infra BYO' middle tier (signed URLs, our GPU) named only in the companion doc.",
31 "Egress + recurring-poll cost: pulling fresh footage out of GCS/Drive/S3 on every sync incurs egress the §4 frugality frame ('don't burn GPU') doesn't model; OQ9 (daily 10 GB) compounds under unattended autosync.",
32 "Gemini consent surface (OQ3): even 'their cloud' footage still ships ?1280px chunks to Google on the Gemini path ? autosync makes this happen automatically/unattended, raising the bar on plain-language consent.",
33 "Credential/OAuth scope + key custody for a persistent watcher ? long-lived watch tokens are exactly what PLATFORM §3.1 warns against ('never persist users' long-lived keys'); narrow drive.file / AssumeRole + presigned URLs needed.",
34 "Streaming-vs-download: large MP4s from a bucket want byte-range streaming (the companion doc reaches for smart_open / fsspec) rather than full download before proxy-prep ? a primitive this locality doc doesn't mention.",
35 "Format/codec variety + 'is this even a match?' filtering: auto-arriving files from arbitrary user storage are unvetted (wrong codec, non-badminton, partial uploads/in-flight writes), unlike the associate-curated L4 drive the pilot assumes.",
36 "Privacy/liability is actually IMPROVED if raw video never leaves the tenant's account (the companion doc's 'central privacy/liability lever') ? but only if processing uses signed-URL/their-compute rather than copying the 10 GB onto our box, which would negate the benefit.",
37 "Retention (OQ1) gets murkier: if the canonical copy lives in the user's cloud, our retention policy may only need to cover proxies/intervals ? a simplification the doc hasn't reasoned through."
38],
39 "fit_verdict": "Strong conceptual fit but NOT yet anticipated here: VIDEO_LOCALITY_MODEL.md has no user-cloud tier and no watch/autosync notion ? BYO-storage+autosync belongs in PLATFORM_ARCHITECTURE.md §3.1's StorageBackend registry plus a new ingestion-trigger component, and would enter this doc as an 'automated-L4' source axis layered onto the async-job model."
40 },
41 {
42 "scope": "docs/PLATFORM_ARCHITECTURE.md (§2 layered architecture, §3 StorageBackend + ComputeBackend registries, §5 orchestration/async jobs, §6 phased roadmap, §7

privacy); cross-checked against docs/DEFERRED_HARDENING_PLAN.md (#16 async job model, now marked COMPLETED, PR #87 + #149).",
43 "anticipates": "partial",
44 "relevant_excerpts": [
45 "§Purpose: 'users upload videos, get highlights stored and browsable by history, and choose where their data lives (Google Drive / cloud bucket / local)'",
46 "§3.1 StorageBackend registry: 'Backends: local (upload dir), gdrive, s3, gcs, azure. BYO-storage: the tenant registers their own bucket/Drive; raw video never leaves their account.'",
47 "§3.1: 'build the backend over fsspec (one matrix ? local / S3 (s3fs) / GCS (gcsfs) / Azure (adlfs) + Google Drive via gdrivefs, Dropbox via dropboxdrivefs), reaching for smart_open as the robust large-MP4 byte-streaming primitive.'",
48 "§3.1 interface includes 'def list(self, prefix: StorageRef) -> list[StorageRef]' and 'def stat(self, ref: StorageRef) -> StorageStat' ? the primitives a poller would need.",
49 "§3.1 Auth: 'For Drive, OAuth with the narrow drive.file scope; for academies, a service account with domain-wide delegation.' / 'For buckets, have the user create an IAM role you AssumeRole into... mint presigned URLs so raw video moves browser <-> the user's bucket directly, never through our servers.'",
50 "§5 Orchestration: jobs table 'id, tenant_id, video_id, type (index|compile|analyze), status...'; '/api/process returns 202 + job_id'; 'Idempotency keyed on (tenant_id, video_id, pipeline_version)'.",
51 "§5: 'the orchestrator enqueues; a worker (hosted or byo_worker) claims jobs for its tenant, resolves input via the tenant's StorageBackend, runs the pipeline'.",
52 "§6 roadmap P2: 'StorageBackend registry + BYO-storage ... local -> gdrive/s3/gcs; replace video_path with StorageRef; encrypted per-tenant creds'.",
53 "§1 gaps table: 'Storage | local filepath string; output to output_dir | no abstraction; can't BYO Drive/bucket; raw video must be on the box'.",
54 "Line 17 (Purpose): 'decouple the layers ? identity, ingestion, storage, orchestration, ...' ? 'ingestion' is named as a layer but never given its own section or a watch/trigger component."
55],
56 "how_it_slots": "The idea splits cleanly into a part the doc already owns and a part it doesn't. (1) The 'bring your own storage' half IS the §3.1 StorageBackend registry: GCS/Drive/OneDrive/S3/Dropbox are explicitly listed as backends (local, gdrive, s3, gcs, azure; Dropbox via dropboxdrivefs), built over fsspec, with BYO-storage as a first-class property ('the tenant registers their own bucket/Drive; raw video never leaves their account'). A 'watched folder' is just a StorageRef prefix; the ABC even exposes list(prefix) and stat() ? the exact primitives a poller needs. This lands on roadmap P2. (2) The 'autosync / watch the folder and auto-enqueue' half is a NEW ingestion/trigger component the doc names but never builds: line 17 lists 'ingestion' as one of the decoupled layers, but no section defines a watcher, change-feed listener, or webhook receiver. The natural slot is a small Ingestion/Sync service in the CONTROL plane (§2) that periodically lists each tenant's StorageBackend prefix (or subscribes to a provider change-feed / GCS Pub/Sub notification / Drive push channel), diffs against already-seen video_ids, and calls the existing enqueue path ? i.e. it becomes a new automated PRODUCER of the §5 jobs table (type=index), reusing the (tenant_id, video_id, pipeline_version) idempotency key as the natural dedup so re-listing the same file never double-processes. So: StorageBackend = the connector (exists in design); Sync/Watcher = the trigger (must be added); jobs table = the unchanged consumer.",
57 "gaps": [
58 "No watch/poll/webhook/change-feed component anywhere in the doc ? every grep for watch|sync|poll|webhook|trigger|enqueue-on-arrival comes back empty. The async job model (#16, now DONE per DEFERRED_HARDENING_PLAN) is PULL-triggered by a user POST to /api/process; nothing auto-creates a job when a file lands.",
59 "'ingestion' is listed as a decoupled layer (line 17) but, unlike storage/compute/orchestration, gets no §, no interface, and no roadmap phase ? it's a named-but-undesigned seam.",
60 "Per-provider change-notification mechanics are unspecified: GCS Pub/Sub object notifications, S3 event notifications, Drive 'changes' API + push channels, OneDrive/Graph delta + webhooks, Dropbox /list_folder/continue + webhooks all differ ? the doc's fsspec-based StorageBackend abstracts reads/writes but NOT change events, so 'watch' can't just be another fsspec method.",
61 "No 'seen-files' / sync-cursor state defined (a per-(tenant, backend, prefix) watermark or delta-token table) so the watcher knows what's new vs. already-enqueued.",
62 "No design for the desktop-app vs web-UI surface the idea proposes (a local

agent watching a cloud-synced folder vs. a server-side poller holding OAuth tokens) ? these have very different credential/egress profiles.",
63 "Scheduling/cadence + cost of continuous polling not addressed (how often to list a large bucket; back-off; quota on Drive/Graph APIs).",
64 "File-readiness / partial-upload handling absent: a recording still being written or synced must not be enqueued mid-flight (need a stat-size-stable / .complete-marker / mtime-quiet-period rule)."
65],
66 "risks_considerations": [
67 "Compute placement / egress: §3.2's middle tiers matter here. If watch+process runs on hosted GPUs while video sits in the user's bucket, every new recording is a full download (egress + the 'raw video never leaves their account' promise in §3.1 is weakened unless signed-URL streaming or BYO-compute is used). §3.2 'zero-infra BYO' (storage stays in user bucket via signed URLs, compute on ephemeral per-tenant sandbox) is the cleanest fit for autosync.",
68 "OAuth scope/credential lifetime: §3.1 says 'never persist users' long-lived keys' and use narrow drive.file scope + presigned URLs. But continuous server-side WATCHING needs a long-lived, refreshable credential or a push-channel registration ? that tension (ephemeral creds vs. always-on poller) is real and unaddressed. drive.file scope only sees files the app created/opened, which may be too narrow to discover arbitrary new recordings a user drops in (may need drive.readonly ? a broader, higher-trust ask).",
69 "Streaming vs download: ffmpeg/cv2/WSL need a local file (§3.1 download_to_local) while the provider can take a signed_url ? an autosync feeding the CV pipeline forces the download path, so per-recording bandwidth is non-trivial for 4K match video (the doc's own GPU/WSL 4K notes apply).",
70 "Format variety: a watched folder will contain arbitrary files (non-video, in-progress uploads, already-processed reels written back). Needs filtering + the partial-upload readiness rule above, plus a guard against re-ingesting KhelSutra's own highlight outputs written to the same bucket (§3.1 upload() target).",
71 "Privacy/minors (§7): auto-pulling 'whatever lands in the bucket' brushes against the §7/§5A guardrail that we must NEVER train on what merely passes through a tenant's bucket and must exclude minors ? fine for PROCESSING consent, but the autosync UX must make clear this is processing-only, not training contribution.",
72 "Idempotency / double-processing: safe by construction IF the watcher routes through the existing (tenant_id, video_id=content-MD5, pipeline_version) idempotency key (§5) ? but MD5 requires reading the whole file, so a cheap pre-check (size+mtime+path cursor) is needed before committing to a full hash/download.",
73 "Multi-tenant fan-out: a per-tenant poller or push-channel registration is per-tenant state to provision, renew (Drive/Graph webhook channels expire), and monitor ? operational surface the single-worker ThreadPoolExecutor drain loop (#16) was not built for."
74],
75 "fit_verdict": "Strong fit for the STORAGE half (it literally is the §3.1 StorageBackend registry / roadmap P2 ? GCS/Drive/OneDrive/S3/Dropbox already enumerated), but the AUTOSYNC/WATCH trigger is a genuine net-new ingestion-layer component the doc names ('ingestion' layer, line 17) yet never designs ? it belongs as a Sync/Watcher service in the control plane that produces jobs into the existing §5 async model."
76 },
77 {
78 "scope": "Primary: docs/DESKTOP_APP_PLAN.md and docs/UI_ALPHA_EXECUTION_PLAN.md (the two assigned surface docs). Cross-checked the docs they both anchor to for where a connector would actually live: docs/VIDEO_LOCALITY_MODEL.md (L0?L4 tier model) and docs/PLATFORM_ARCHITECTURE.md (StorageBackend registry §3.1).",
79 "anticipates": "no",
80 "relevant_excerpts": [
81 "DESKTOP_APP_PLAN.md §1: the ENTIRE source model is removable physical media ? 'User plugs the GoPro/phone SD card or camera USB into their laptop. The app detects the card, lists the match videos on it.' No cloud/online source anywhere.",
82 "DESKTOP_APP_PLAN.md D2 (Decisions locked): 'The app reads source video in place (off the card / a chosen folder)' ? 'a chosen folder' is the closest thing to a generic source, but it means a LOCAL folder, not a watched cloud bucket.",
83 "DESKTOP_APP_PLAN.md §4.1 the one-click flow is triggered by '[card inserted] ? (NEW) per-OS device-insert watcher ? mount path'. The only 'watcher' in the doc is a USB/SD hardware-insert watcher (Win WMI/WM_DEVICECHANGE, macOS DiskArbitration, Linux udev) ? NOT a cloud-folder change watcher.",
84 "DESKTOP_APP_PLAN.md §4.2 reuse map, NEW rows: 'USB/SD insert detection +

video listing' ? there is no row for a cloud connector or a sync daemon.",
85 "DESKTOP_APP_PLAN.md §1: 'no upload, no copy of the GBs of 4K originals ? fully on-device' ? the doc's thesis is explicitly the L0 'nothing leaves the machine' tier (VIDEO_LOCALITY_MODEL §0), which is in tension with pulling from a cloud account.",
86 "UI_ALPHA_EXECUTION_PLAN.md Phase L: the ONLY ingestion path is 'PR-2 · B2+B3 upload & download: chunked upload (init ? sequential ?95 MB parts ? complete ? MD5; ? mp4/mov allowlist)'. Source = manual browser upload; no external-storage connect.",
87 "UI_ALPHA_EXECUTION_PLAN.md L2 SPA 'M2: A2 Upload (chunked, guidelines checklist)' ? the UI surface is a manual upload form, no 'connect Drive/bucket' screen.",
88 "VIDEO_LOCALITY_MODEL.md line 25 (the only mention of self-managed cloud storage, and it is DISMISSIVE): 'expecting academies to manage S3/GCS buckets is a [miserable ?]' ? the locality doc actively argues AGAINST asking users to wrangle their own buckets.",
89 "PLATFORM_ARCHITECTURE.md §3.1 (OUT of assigned scope but is where the idea truly belongs): a 'StorageBackend' registry with backends 'local, gdrive, s3, gcs, azure' and explicit 'BYO-storage: the tenant registers their own bucket/Drive; raw video never leaves their account' ? scheduled as roadmap item P2.",
90 "PLATFORM_ARCHITECTURE.md §3.1 auth precedent: 'For Drive, OAuth with the narrow drive.file scope ? For buckets, have the user create an IAM role you AssumeRole into ? mint presigned URLs so raw video moves browser ? the user's bucket directly, never through our servers.'"
91],
92 "how_it_slots": "Neither assigned doc has a home for it as written; the connector belongs to the StorageBackend registry in PLATFORM_ARCHITECTURE.md §3.1 (roadmap P2, NOT yet built), and the autosync loop is a brand-new component with no precedent. Concretely: (1) The 'source' half ? GCS/Drive/OneDrive/Dropbox/S3 ? IS the planned StorageBackend ABC (local?gdrive/s3/gcs/azure, smart_open + s3fs/gcsfs/gdrivefs). (2) The 'autosync/watch daemon' half ? 'folder lands ? reels appear' ? has NO component in any doc: PLATFORM_ARCHITECTURE's async job model (P0, shipped) is request-driven (POST /api/process ? 202 + job_id), not event/poll-driven; it needs a new watcher that polls/subscribes the StorageBackend and enqueues a job per new object. (3) In the DESKTOP surface it slots beside the existing per-OS device-insert watcher (§4.1) as a second trigger type ? a 'cloud-folder watcher' generalizing the USB-insert watcher ? feeding the same proxy?detect?stitch flow; but it breaks D1/D2's 'no upload, no original-copy, fully on-device' invariants because the bytes now come down from a cloud account (egress + a local cache). (4) In the WEB-UI surface it slots as a new 'Connect storage' screen in L2 (alongside/replacing A2 manual Upload) plus a backend enabler in L1 (a new PR like 'B-x: StorageBackend connect + sync worker') that the async job model dispatches from. Tier-wise it is a new variant of L2/L1 in VIDEO_LOCALITY_MODEL: the proxy/chunks are pulled from the user's own cloud rather than browser-uploaded.",
93 "gaps": [
94 "A watch/autosync DAEMON concept ? 'new file arrives ? auto-process' ? exists in NO doc. PLATFORM_ARCHITECTURE's job model is pull-on-request (POST /api/process); there is no poller/webhook/inotify/change-feed component, no de-dup-by-md5-on-arrival, no 'already processed this object' ledger.",
95 "DESKTOP_APP_PLAN treats source as removable media only (SD/USB) or 'a chosen [local] folder'; it has no cloud-source connector, no OAuth, no streaming-from-cloud path, and its D1/D2 'no upload / no original-copy / fully on-device' invariants directly CONFLICT with pulling GBs down from a cloud account (which is download + local cache, i.e. an upload in reverse).",
96 "UI_ALPHA_EXECUTION_PLAN has no 'connect external storage' screen and no backend enabler for it ? its only ingest is manual chunked upload (PR-2 B2/B3); adding BYO-storage+sync would be a new PR in Phase L and a new SPA milestone, neither budgeted in the S1?S11 session map.",
97 "The DESKTOP per-OS device watcher (USB/SD) and a cloud-folder watcher are different beasts (DiskArbitration/udev vs. Drive change-feed / S3 event notifications / polling); the doc's §4.2 watcher row does not cover the cloud case.",
98 "VIDEO_LOCALITY_MODEL line 25 actively frames 'academies managing S3/GCS buckets' as a non-starter for the target (non-technical) user ? so the idea needs a much gentler 'connect your Google Drive/photos folder' UX than raw-bucket BYO, or it contradicts the stated persona.",
99 "OQ-level coverage exists for credential custody/consent in PLATFORM_ARCHITECTURE §3.1 (AssumeRole, drive.file OAuth, presigned URLs) but those guardrails are NOT referenced or inherited by either surface doc."
100],
101 "risks_considerations": [
102 "Compute placement: DESKTOP_APP_PLAN's whole gate is P0 'can a no-GPU laptop

run WASB at all?' (~3.4× RT even on a 2070S). A cloud-source autosync that processes locally inherits that unsolved make-or-break; processing in the cloud instead would flip the surface from L0/desktop to the hosted/BYO-compute path (PLATFORM_ARCHITECTURE §3.2), a different product.",
103 "Egress + duplication cost: pulling GBs of 4K from GCS/Drive/S3 to process them is download bandwidth + a local cache ? it breaks DESKTOP D1/D2's 'no upload, never copy the originals' selling point and re-introduces the exact bandwidth pain VIDEO_LOCALITY_MODEL §0 was built to avoid (10 GB ? ~1 h at 20 Mbps). Streaming-vs-full-download matters: detection needs only a proxy, but stitch reads the original ? so the original (or a ranged stream) must come down at reel time.",
104 "OAuth/scopes & key custody: PLATFORM_ARCHITECTURE §3.1 already sets the right posture (narrow drive.file scope, IAM AssumeRole into one bucket/prefix, presigned URLs, 'never persist users' long-lived keys') ? any connector MUST adopt this; a naive token-store would be a security regression.",
105 "Privacy/consent: a background daemon that auto-uploads/auto-processes whatever lands in a watched folder can sweep in unintended footage (and per the guardrails, minors must be excluded from any training pool, and the Gemini path still leaks ?1280px chunks to Google per VIDEO_LOCALITY_MODEL OQ3). 'Auto' makes the consent surface bigger, not smaller.",
106 "Format variety: the web ingest today is mp4/mov-allowlisted (UI_ALPHA PR-2); a watched cloud folder will contain arbitrary files (other codecs, photos, partial uploads still syncing) ? the watcher needs a stable-file/ready check and codec filtering the docs don't yet specify.",
107 "De-dup / idempotency: 'reels appear automatically' needs an arrival ledger keyed by video_id (MD5) so re-syncs or cloud-client re-writes don't reprocess; the indexer keys by MD5 already (VIDEO_LOCALITY_MODEL OQ2), but no doc defines the 'seen this object' bookkeeping for a sync loop.",
108 "Persona mismatch: VIDEO_LOCALITY_MODEL line 25 says expecting non-technical academies to manage buckets is 'miserable' ? raw BYO-bucket risks the wrong audience; a consumer-cloud (Drive/OneDrive/Dropbox) 'watch this folder' framing fits the persona better than S3/GCS."
109],
110 "fit_verdict": "Real fit, wrong doc: neither DESKTOP_APP_PLAN nor UI_ALPHA anticipates a cloud-storage source or an autosync daemon ? the connector belongs to PLATFORM_ARCHITECTURE.md §3.1's StorageBackend registry (roadmap P2, not built), and the 'watch ? auto-process' loop is a genuinely NEW component (no poller/webhook/change-feed exists anywhere) that should be specced as an autosync layer on top of the already-shipped async job model."
111 },
112 {
113 "scope": "docs/STORAGE_SHARING_MODEL.md (DRAFT v2, 2026-06-12, unratified) + docs/GOLDEN_DATA_SHARING.md (convention/plan 2026-06-15, tracked in #175).",
114 "anticipates": "partial",
115 "relevant_excerpts": [
116 "STORAGE_SHARING_MODEL §2: 'Source (raw match upload) | Processing INPUT | 1?4 GB each | kept through processing + correction window, then auto-deleted ... | optional \"keep my raw match\" toggle' vs 'Library (compiled reels...) | The PRODUCT | Durable | Yes ? this is what 2 GB means'",
117 "STORAGE_SHARING_MODEL §2: 'Do we store the uploaded video? ? Yes, transiently ? it's a processing input with a retention deadline, not a possession. The reel is the possession.'",
118 "STORAGE_SHARING_MODEL §5 (Beta/hosted): 'the bucket IS the upload path ? browser ? object storage directly via signed resumable URLs (PLATFORM P2 StorageBackend); the server never proxies bytes'",
119 "STORAGE_SHARING_MODEL §7d (USB-drive vision): 'Serious users plug a drive of match videos into their Windows/GPU machine; the app processes locally ... and syncs only reels + metadata to the cloud library/profile.' + 'Maps to PLATFORM P3 local/byo_worker compute backends ? productized.'",
120 "STORAGE_SHARING_MODEL §1a axes table: 'Compute (the cost) | Our GPU, capped runs | Our GPU, priority | Their GPU (USB-drive workflow)'",
121 "STORAGE_SHARING_MODEL §9: 'storage_ref follows P2's StorageRef (backend id + key)'",
122 "GOLDEN_DATA_SHARING §'Canonical shared folder': '%USERPROFILE%\\OneDrive\\rally-annotator\\golden-data\\ ... OneDrive is present + signed-in on both the GPU box and the dev machine under the same account, so it auto-syncs.'",
123 "GOLDEN_DATA_SHARING workflow: 'OneDrive syncs the JSONs to the dev machine

automatically (small files, seconds)' + the env-var/`scripts/sync_golden.py` copy ? shared-folder ergonomics."
124],
125 "how_it_slots": "The idea slots in as a new entry in the StorageBackend registry (PLATFORM P2, referenced in §5 Beta + §9 storage_ref = StorageRef{backend id + key}) PAIRED with a new ingest/watcher mode. Concretely: (1) the user's GCS/Drive/OneDrive/S3/Dropbox folder becomes a *read-only source StorageBackend* that the system polls/watches; (2) a watcher converts each newly-appeared file into the existing 'Source (raw match upload)' tier object in §2 and feeds the §4 CUJ ('upload source ? async job ? review ? compile reel ? reel lands in LIBRARY'), just with the manual upload step replaced by autosync; (3) the produced reels still land in KhelSutra's own Library tier (R2/GCS per §6). It is most naturally the *autosync front-door to the §7d 'BYO storage/compute' end-state* and the §1a 'Their GPU' compute axis ? i.e. a productized P3 local/byo_worker mode where the customer's cloud folder, not a USB drive, is the source of truth.",
126 "gaps": [
127 "No 'watch/poll an external user-owned folder' concept exists. The doc's only 'point at a bucket' (§5 Beta) is KhelSutra's OWN object storage as the upload TARGET via signed resumable URLs ? a push-from-browser path, not a pull/watch of a folder the customer already owns. The watcher (change feed / poll / webhook, new-file detection, dedup by video_id) is unspecified.",
128 "No OAuth / least-privilege scope model for connecting a third-party account (Drive/OneDrive/Dropbox/Microsoft 365 consent screens, S3/GCS service-account keys or HMAC). §8 guardrails cover consent-at-capture and share-link signing but not 'connect your cloud account' credential storage/rotation/revocation.",
129 "No 'read-only access to user-owned external storage' tier in the §2 object model. Sources are assumed to live in KhelSutra-controlled storage (transient, retention_deadline). A BYO-folder source the user keeps permanently changes the 'we store transiently, then auto-delete' premise ? the file already lives in the user's bucket; KhelSutra may only need a transient working copy.",
130 "Compute placement for BYO storage is unaddressed: §7d/§1a put 'their GPU' with a local USB drive; a cloud folder + our-GPU processing means egress to PULL multi-GB sources INTO our compute (the inverse of the §6 egress concern, which is about SERVING reels out). Who pays ingress/cross-cloud egress to fetch 1?4 GB per match is not modeled.",
131 "No streaming-vs-download decision for large sources (§5 caps assume an upload; pulling 1?4 GB from Drive/S3 to process is a different cost/latency profile).",
132 "Quota semantics for BYO storage are undefined: if the raw match lives in the user's own bucket, does it count against the 2 GB meter (§2 says quota is library/reels only)? The 'keep my raw match' toggle (§2) assumes the raw lives in OUR storage.",
133 "Schema (§9) has storage_ref for library/source objects but no 'connected external source' entity (provider, folder path, credential ref, last-synced cursor, watch state)."
134],
135 "risks_considerations": [
136 "Egress/ingress cost inversion: §6 frames the bill as egress on SERVING reels (R2 zero-egress chosen for this). BYO-cloud-storage adds the opposite cost ? pulling 1?4 GB raw matches FROM the user's GCS/S3/Drive INTO our GPU; cross-cloud egress on the USER's bucket (which they pay) plus our ingress/compute. Format/streaming-vs-download matters (§5 sizing).",
137 "Compute placement tension: the doc's BYO end-state (§7d, §1a 'Their GPU') assumes LOCAL processing on the customer's machine so cloud stays 'storage + social only.' A watch-my-cloud-folder + OUR-GPU variant re-centralizes compute and reintroduces the per-run Gemini/$ cost the §7e abuse economics tries to avoid ? needs an explicit 'free never touches Gemini' (§7a) ruling for autosynced runs and an N-runs/mo cap (§7b).",
138 "OAuth scopes / privacy: connecting Drive/OneDrive/Dropbox/365 means broad third-party read tokens; §8's consent-at-capture and training-opt-in guardrails must extend to 'we read your cloud folder' ? and GOLDEN_DATA_SHARING §Guardrails ('keep the OneDrive folder private ? Proprietary corpus, no public share links') is the proof that even internal cloud-folder sync is treated as privacy-sensitive.",
139 "Format variety: a watched folder receives arbitrary files; the §2/§7e relevance validator and library 'pipeline-outputs-only' abuse guard must run on auto-ingested files too, or autosync becomes a 'free CDN / process-anything' abuse vector the doc explicitly designed against.",
140 "Retention-model collision: §2's 'source = transient, auto-deleted' assumes WE hold the raw. With BYO storage the raw is durably the user's; deciding whether we keep a transient working copy (and its retention_deadline) vs re-pull on demand affects both cost and the 'not a possession' framing.",

141 "Abuse/quota: an always-watching pull from a user-controlled, externally-mutable folder needs rate-limiting + dedup (video_id = MD5, per CLAUDE.md) so re-appearing/renamed files don't re-trigger paid processing."
142],
143 "fit_verdict": "Strong strategic fit and clearly the autosync front-door to the §7d BYO-storage/§1a 'their compute' end-state ? but NOT yet anticipated as a feature; it belongs as a new read-only 'connected external source' StorageBackend + watcher mode layered on the existing §2 source-tier and §4 CUJ, with OAuth/egress/compute-placement decisions still to be designed (candidate co-design item #15)."
144 },
145 {
146 "scope": "Primary: docs/PLATFORM_ARCHITECTURE.md (§3 pluggable Storage/Compute registries), docs/COMMERCIALIZATION.md (§6 PMF + dashboard), docs/PMF_MARKET_RESEARCH.md (segments/pricing), docs/archives/MONETIZATION_AUDIT.md (licensing/codec/egress). Supporting (read because they carry the storage/egress/autosync threads the named docs point to): docs/STORAGE_SHARING_MODEL.md (point-at-a-bucket upload + egress economics), docs/VIDEO_LOCALITY_MODEL.md (the locality ladder / cost story), docs/DESKTOP_APP_PLAN.md (the device-insert \"watcher ? auto-prepare\" precedent).",
147 "anticipates": "partial",
148 "relevant_excerpts": [
149 "PLATFORM_ARCHITECTURE.md §3.1: 'Backends: local (upload dir), gdrive, s3, gcs, azure. BYO-storage: the tenant registers their own bucket/Drive; raw video never leaves their account.'",
150 "PLATFORM_ARCHITECTURE.md §3.1 (auth): 'never persist users' long-lived keys... mint presigned URLs so raw video moves browser ? the user's bucket directly, never through our servers. For Drive, OAuth with the narrow drive.file scope; for academies, a service account with domain-wide delegation.'",
151 "PLATFORM_ARCHITECTURE.md, competitive note: 'none of Veo / Hudl / SwingVision / Spiideo offer true BYO-storage or BYO-compute... So \"your video, your storage, your compute\" is a genuine differentiator for privacy-sensitive academies.'",
152 "PLATFORM_ARCHITECTURE.md §3.1 (recommended impl): 'build the backend over fsspec (one matrix ? local / S3 (s3fs) / GCS (gcsfs) / Azure (adlfs) + Google Drive via gdrivefs, Dropbox via dropboxdrivefs)... Treat Drive/Dropbox as add-ons ? less battle-tested... cover with integration tests.'",
153 "PLATFORM_ARCHITECTURE.md §3.2 'Middle tier ? zero-infra BYO': 'storage stays in the user's bucket (signed URLs) while compute runs on our GPUs in a strictly isolated, per-tenant, ephemeral sandbox (nothing persists our side).'",
154 "STORAGE_SHARING_MODEL.md §5: 'Beta (hosted): the bucket IS the upload path ? browser ? object storage directly via signed resumable URLs (PLATFORM P2 StorageBackend); the server never proxies bytes.'",
155 "STORAGE_SHARING_MODEL.md §6: 'The real bill is egress, not storage... Storing a free user's 2 GB costs < ?5/month anywhere. Serving is the risk: on GCS, egress is ~$0.12/GB... on Cloudflare R2, egress is $0.'",
156 "VIDEO_LOCALITY_MODEL.md §1: 'any design that moves the 10 GB is moving the wrong thing... what is the smallest artifact that must cross the network for this tier of service.' L1-direct: 'chunks go straight from the client to Gemini ? zero video bytes through our server.'",
157 "DESKTOP_APP_PLAN.md §4.1: '[card inserted] ? (NEW) per-OS device-insert watcher ? mount path ? [list videos] ? [make on-device proxy] ? [detect] ? [stitch] ? [export reel].' (the local twin of cloud-folder autosync)",
158 "COMMERCIALIZATION.md §6 beachhead: 'Sell to the academy, not the individual'; pricing anchors B2C USD 60-180/yr, B2B academy/club USD 300-1,500/yr; PMF_MARKET_RESEARCH.md: 'static-camera constraint is a B2B filter... coaches/academies/clubs already record from a fixed angle.'"
159],
160 "how_it_slots": "It is a new ingestion-TRIGGER layer sitting on top of the already-designed StorageBackend registry (PLATFORM_ARCHITECTURE §3.1). The 'point at storage you own' half is exactly the existing BYO-storage backend (gdrive/s3/gcs/azure/dropbox over fsspec, with signed-URL + drive.file-OAuth + never-persist-keys auth). What does NOT yet exist is the 'WATCH the folder and auto-process on arrival' half ? a Watcher/Ingest-trigger component that calls StorageBackend.list()/stat() on a poll (or a provider push: S3 EventBridge, GCS Pub/Sub, Drive changes-watch, Dropbox webhooks), dedupes by the existing video_id=MD5, and submits an async job (the shipped /api/process ? 202 + job_id model, EXECUTION_POLICY.md). On the desktop path it is the cloud-folder analog of DESKTOP_APP_PLAN §4.1's per-OS device-insert watcher ? same 'watcher ? auto-prepare' shape, different source. Compute placement then chooses among the

§3.2 ComputeBackend tiers: 'zero-infra BYO' (their bucket via signed URL, our ephemeral GPU sandbox) is the natural default; byo_worker pulls from their own storage for the full-privacy tier.",
161 "gaps": [
162 "No watch/autosync/ingest-trigger concept anywhere in the docs ? grep for watch/auto-sync/poll/webhook/'new recording' finds only the device-insert watcher (DESKTOP_APP_PLAN) and unrelated job-polling. The trigger component (poll vs provider-push, dedupe, debounce on still-uploading files) is entirely unspecified.",
163 "Who-pays-egress is unanalyzed for the PULL direction. STORAGE_SHARING_MODEL §6 only costs egress for OUR serving (library?R2). If WE pull each new recording from the user's GCS/Drive to our GPU, the user eats provider download/egress on every match (GCS ~$0.12/GB × 1-4 GB/match, daily) ? a recurring surprise bill the docs never model. The desktop-pulls variant avoids this but isn't framed as a BYO-cloud flow.",
164 "No OAuth/scopes/token-lifecycle design for a LONG-LIVED watch. §3.1's drive.file scope only sees files the app created ? a watcher of the user's existing footage folder needs broader read scope (drive.readonly) + refresh-token storage + Google's app-verification/security-review, none of which is addressed. Same for Dropbox/OneDrive consumer OAuth and MS365 admin consent.",
165 "Streaming-vs-download decision for the watcher's input is unmade. PLATFORM exposes both open_read (stream) and download_to_local (ffmpeg needs a local file); locality ladder says proxy/chunks, not the 10 GB original, should cross ? but a naive autosync that download_to_local's every full original is the anti-pattern §1 warns against.",
166 "No retention/privacy posture for the 'we briefly hold their video' window. OQ1 (retention per artifact class) and OQ3 (Gemini-path consent: 1280px chunks still leak to Google even in 'local' modes) are open in VIDEO_LOCALITY_MODEL; an autopilot that silently ingests every new file makes consent-at-capture and the auto-delete deadline more load-bearing, not less.",
167 "Format/source variety unhandled: GoPro Quik / phone-Photos auto-upload produces HEVC + chaptered/segmented clips + non-match footage. The relevance validator + HEVC decode exposure (C6 ??) + clip-stitching are mentioned elsewhere but not tied to an auto-ingest firehose that has no human in the loop to reject junk.",
168 "No PMF/segment framing for BYO-cloud specifically. The docs' privacy-BYO story is bucket/Drive REGISTRATION; 'footage auto-uploads to my cloud, reels just appear' is an unstated value-prop and isn't priced or placed in a tier (STORAGE_SHARING_MODEL §7b tiers are storage-size, not autosync)."
169],
170 "risks_considerations": [
171 "Compute placement decides who pays: if WE pull ? user pays cloud egress + we pay GPU + we briefly hold raw video (privacy surface). If the DESKTOP app pulls ? zero egress through us, no raw video on our infra, but needs a standing local agent (DESKTOP_APP_PLAN P0 compute-feasibility gate: WASB on a no-GPU laptop is unverified, possibly hours/clip).",
172 "Egress is the real bill (STORAGE_SHARING §6) ? but that analysis only covers serving; pulling from BYO-cloud inverts it and makes the USER's provider the meter. The whole cost story ('we don't store giant videos') holds for storage but NOT automatically for egress on a WE-pull design.",
173 "Locality discipline (VIDEO_LOCALITY §1: don't move the 10 GB) argues the watcher should pull a proxy/chunks, not the original ? pushing the proxy step to the client (L1/L2), which a pure cloud-bucket-watch can't do without a client. This tensions 'zero manual upload, fully cloud' against 'don't move the wrong bytes'.",
174 "OAuth scope creep + app-verification: watching the user's existing footage folder needs read scopes broader than the privacy-minimal drive.file the doc chose, plus Google/Dropbox/MS security review and refresh-token custody ? a real ongoing security + compliance load.",
175 "Privacy/consent: BYO-storage's headline ('raw video never leaves your account') is WEAKENED the moment we pull-and-process ? we transiently hold it and (on the Gemini path) ship 1280px chunks to Google (OQ3). Auto-ingest removes the human checkpoint, so consent-at-capture (minors-excluded guardrail) and a hard retention deadline become mandatory, not optional.",
176 "Streaming vs download: ffmpeg/cv2/WSL need a local file (download), the LLM can take a signed URL (stream) ? an autosync must choose per stage or it defaults to downloading full originals (worst case).",
177 "Format variety: HEVC decode royalty exposure (C6 ??), GoPro chaptering, and non-sport footage in a watched folder all hit a no-human-in-loop pipeline; the relevance validator and input-quality reports become the only gate.",
178 "Support burden (PLATFORM §3.2 ??): a standing watcher/agent on academy infra

is the 'it stopped after our IT change' ticket the doc already flags for byo_worker ? applies equally to a desktop autosync agent."
179   ],
180   "fit_verdict": "Strong strategic fit and on-thesis, but only HALF-anticipated: the BYO-storage backend is already designed (PLATFORM_ARCHITECTURE §3.1) ? the missing piece is a new Watcher/ingest-trigger layer on top of it; it belongs in PLATFORM_ARCHITECTURE (a §3.x 'IngestTrigger/Watcher' seam) with the who-pays-egress + OAuth-scope + retention questions added to VIDEO_LOCALITY_MODEL's open-questions ledger before any greenlight."
181   }
182   ]
183   }
184   }

254. assistant (2026-06-19T19:21:12.361Z)

Grounded the assessment in the actual platform docs (quoting them, so this isn't me riffing) ? and the verdict is clean:

This is half-already-designed, on-thesis, and a real differentiator

The **"bring your own storage"** half is *already in your roadmap*. `PLATFORM_ARCHITECTURE.md §3.1` literally specs a `StorageBackend` registry ? **"Backends: local, gdrive, s3, gcs, azure ? BYO-storage: the tenant registers their own bucket/Drive; raw video never leaves their account"** ? built over `fsspec`/`smart_open`, with the auth posture already chosen (narrow `drive.file` OAuth, IAM AssumeRole + presigned URLs, *never persist long-lived keys*). It's roadmap item **P2**. And your own doc already makes the strategic case: *"none of Veo / Hudl / SwingVision / Spiideo offer true BYO-storage or BYO-compute ? 'your video, your storage, your compute' is a genuine differentiator for privacy-sensitive academies."*

You've even **proven the ergonomics on yourselves**: `GOLDEN_DATA_SHARING.md` already runs a OneDrive folder that auto-syncs both ways between your GPU box and dev machine. That's exactly this pattern, in production, internally.

What's genuinely new is the autosync trigger. Across all the docs, there is *no* watcher/poller/webhook/"new file ? process" component. The shipped async job model (#16) is **pull-on-request** (`POST /api/process ? 202`). "ingestion" is even named as a layer but never built. So the work splits cleanly:

Piece	Status
Connect GCS / Drive / OneDrive / S3 / Dropbox	**Designed** (`§3.1` StorageBackend, P2) ? not built
`list()`/`stat()` primitives a poller needs	**Already in the ABC**
Dedup so re-syncs don't reprocess	**Already there** ? `video_id` = content MD5 idempotency key
Watch folder ? auto-enqueue a job	**Net-new** ? the one real new component
Reels land back in the library	Existing flow

The honest risks ? and the one decision that unblocks everything

Compute placement is the fork that determines who pays and how private it really is:
- **Desktop app pulls** (their machine watches their synced folder, processes locally): zero egress through us, no raw video ever touches our infra ? the strongest privacy story. **But it's gated on your unsolved desktop P0 spike** (WASB is ~3.4× real-time even on a 2070S; a no-GPU laptop could be hours/clip or infeasible).
- **We pull (hosted):** simplest UX, but we briefly hold their video (which *weakens* the "never leaves your account" headline) and someone eats recurring egress. The clean middle is your own `§3.2` **"zero-infra BYO"** ? storage stays in their bucket via signed URLs, compute on an ephemeral per-tenant GPU sandbox.

Pick that, and the rest falls out. The other live ones:
- **Egress / "surprise bill":** pulling 1?4 GB per match *out* of their cloud meters egress on *their* provider. Your cost story ("we don't store giant videos") holds for storage but **not automatically for egress** on a we-pull design. Locality discipline says pull a *proxy*, not the 10 GB original ? which really wants a client (desktop).
- **OAuth reality:** the privacy-minimal `drive.file` scope only sees files *the app created*.

Watching a user's *existing* footage folder needs `drive.readonly` (broader, higher-trust) + **Google/Microsoft/Dropbox app-verification/security review** + refresh-token custody. MS365 needs admin consent. That's a real weeks-long compliance gate, not a weekend.
- **"Stream it, never download" isn't fully true:** detection can stream a proxy, but **stitch reads the original**, and ffmpeg/cv2 want a local file ? so originals (or ranged streams) come down at reel time. Be honest in the copy.
- **Autosync hygiene** (the new component's actual work): a watched folder is a firehose ? in-flight partial uploads, non-video, HEVC/GoPro-chaptered clips, non-badminton footage, **and your own reels written back to the same bucket** (re-ingest loop). Needs a file-ready check, relevance/codec filtering, and the MD5 dedup ledger. Plus per-provider change-feeds all differ (GCS Pub/Sub, S3 events, Drive push channels that *expire*, Graph delta) ? `fsspec` abstracts reads, not change events.
- **Consent gets *bigger*:** a background daemon that auto-processes "whatever lands" removes the human checkpoint ? consent-at-capture, minors-exclusion, retention deadline, and "free tier never hits Gemini" become mandatory, not optional.

Persona nuance worth locking

Your locality doc explicitly calls "expecting academies to manage S3/GCS buckets" a **non-starter**. So split it: **consumer cloud (Drive/OneDrive/Dropbox) ? "watch this folder, reels appear"** for the non-technical academy/prosumer; **raw S3/GCS bucket BYO** for the technical/B2B tenant. Same engine, two front doors.

My recommendation

Strong yes as a direction ? but it's a platform layer, not the next thing. It sits *after* the immediate launch (tonight's page, the alpha web app, corpus growth) and rides on the **desktop P0 compute spike**, because the make-or-break question (where compute runs / who pays egress) is the same unknown that spike resolves. Beautifully, it composes with the new positioning: **autosync ? index rallies ? (coming) ask for your reel** = a zero-touch funnel.

When you're ready, I'd capture it as a tracked design doc (it spans three existing ones): connector half stays in `PLATFORM_ARCHITECTURE §3.1` (P2), and I'd add a new **"IngestTrigger / Watcher"** seam there + push the egress/OAuth/retention open-questions into `VIDEO_LOCALITY_MODEL`'s ledger ? it's effectively the autosync front-door to `STORAGE_SHARING_MODEL §7d` (your co-design item #15). I won't start it unprompted (owner-gated strategy), but say the word and I'll draft the doc with the compute-placement decision teed up as the lead question.

For tonight, though ? this changes nothing about go-live. Page's ready when you are. ?

255. user (2026-06-20T02:01:28.198Z)

Right after my tea I will come back and we will do the godaddy thing, can you please spin up the server again?

256. assistant (2026-06-20T02:01:40.758Z)

Of course ? enjoy your tea! Spinning the server back up now.

257. user (2026-06-20T02:01:43.837Z)


```
{
  "serverId": "5426fa31-5192-4c48-b88c-77fe9f64c03d",
  "port": 8099,
  "name": "landing",
  "reused": false
}
```

Server started successfully on port 8099.

258. assistant (2026-06-20T02:01:53.049Z)

Server's back up at **http://localhost:8099**. Quick sanity check that it's serving the latest version (play-triangle logo + today framing).

259. user (2026-06-20T02:01:57.623Z)

```
HTTP status:
200
Title + key markers:
<title>KhelSutra ? every rally indexed, the dead time gone</title>
class="play"
Every rally, indexed
class="play"
Every rally, indexed
Coming next to the alpha
class="play"
Assets:
  /og.png -> 200
  /capture.html -> 200
```

260. assistant (2026-06-20T02:02:03.403Z)

```
All set and serving the latest build:

- **http://localhost:8099** ? 200 ?
- Title: *"KhelSutra ? every rally indexed, the dead time gone"* ?
- Play-triangle logo (`class="play"`) ? · today-framed hero ("Every rally, indexed") ? · "Coming
next to the alpha" section ?
- `og.png` ? 200 ? · `capture.html` ? 200 ?

It's loaded in the **preview panel** too. Enjoy the tea ? ? when you're back, just say **"let's
go live"** and we'll do the GoDaddy ? Cloudflare walkthrough together (the only GoDaddy step is
the nameserver paste; everything else is Cloudflare). I'll be right here.
```

261. user (2026-06-20T02:19:16.037Z)

```
yes, LGTM! And after its ive you will checkin the PR?
```

262. assistant (2026-06-20T02:20:09.494Z)

```
Yes ? happy to. There's just **one decision to nail before I commit anything**, and it's a real
one: *which repo* the marketing site lives in. Putting a static landing page inside the
badminton-engine repo mixes concerns; the cleaner home is the brand repo (and it lets Cloudflare
auto-redeploy on push). Let me check what's actually on disk so I recommend accurately rather
than guess.
```

263. user (2026-06-20T02:20:12.579Z)

```
=== sibling repos under Projects/ ===
Annotation Setup
CODE WITH ERIC
Email-Mining
KhelSutraGuru.jpeg
Vairagi
Verifiers
badminton-highlight-indexer
gmail-attachments-to-gdrive
gtest
khelsutra
muneem
rally-annotator
sports-data-collector
sports-obsreport

=== is artifacts/ gitignored in the indexer repo? ===
NOT ignored (trackable)

=== any khelsutra repo locally? ===
```

```
/c/Users/avidu/Projects/khelsutra

=== current branch / HEAD (indexer) ===
docs/court-floor-rally-end
430de76 docs(rally): capture the shuttle-floor-contact rally-END signal (owner note)
```

264. assistant (2026-06-20T02:20:31.085Z)

Good ? that settles it: the private **`khelsutra` brand repo is already cloned** at `C:\Users\avidu\Projects\khelsutra`, and the indexer's `artifacts/` is just a scratch area (not where a marketing site should live). Let me peek at the khelsutra repo's shape so my plan is concrete.

265. user (2026-06-20T02:20:33.852Z)

```
=== git? ===
true
=== remote ===
origin  https://github.com/avidullu/khelsutra.git (fetch)
origin  https://github.com/avidullu/khelsutra.git (push)
=== default/current branch ===
master
=== top-level contents ===
.git
.gitignore
README.md
docs
web
=== any existing site/landing/index.html? ===
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

266. assistant (2026-06-20T02:21:06.897Z)

That's the answer, then. The **`khelsutra` brand repo** (`github.com/avidullu/khelsutra`, already cloned, already has a `web/` dir) is the right home ? not the indexer's scratch `artifacts/`. It keeps the marketing site out of the ML-engine repo, and Cloudflare Pages can **auto-redeploy on every push** to it, so future tweaks go live by `git push`.

So yes ? here's the plan, **after** we go live:
1. **Tonight:** go live via Cloudflare **direct upload** (fast, no git needed) so khelsutra.guru is up the moment we finish the GoDaddy nameserver step.
2. **Then I check it in** ? pull `khelsutra`, branch off `origin/master`, add the four files (`index.html`, `capture.html`, `og.png`, `make_og.py`) under its `web/` convention, **one clean PR** (explicit paths, branch from origin/master, re-verify HEAD before commit ? the usual shared-checkout safety). I'll also tidy the temporary `launch.json` I dropped in the indexer.
3. **Optional follow-up:** connect that repo to Cloudflare Pages so it auto-deploys ? then you never touch the dashboard again.

One quick call so it's ready when you're back:

267. user (2026-06-20T02:21:37.312Z)

Your questions have been answered: "Which repo should the landing page be checked into?"="khelsutra brand repo (Recommended)". You can now continue with these answers in mind.

268. assistant (2026-06-20T02:21:49.546Z)

Locked in: the **`khelsutra` brand repo**, under its `web/` convention.

So the full sequence when you're back:
1. ? ? **"let's go live"** ? GoDaddy nameservers ? Cloudflare Pages (direct upload) ? custom domain ? email routing. khelsutra.guru is live.
2. **Then I check it in:** `git pull` khelsutra ? look at how `web/` is laid out ? branch off `origin/master` ? add `index.html`, `capture.html`, `og.png`, `make_og.py` ? **one clean PR**

(explicit paths, HEAD re-verified before commit, no stacking). I'll open the PR, not push to master.
3. Offer to wire that repo to Cloudflare Pages so future edits deploy on `git push`.

I'll hold the check-in until it's actually live, exactly as you said. Whenever you're back ? I'm ready. ?

269. user (2026-06-20T02:22:19.921Z)

i logged into my godaddy account

270. assistant (2026-06-20T02:23:14.273Z)

Perfect timing. One thing to flag so we don't stall: the **two nameserver values you'll paste into GoDaddy come *from* Cloudflare** ? so we grab those first, then drop them into the GoDaddy screen you're on. Keep that GoDaddy tab open.

Step 1 ? Cloudflare: add the domain, get the nameservers (~3 min)

1. New tab ? **dash.cloudflare.com** ? log in, or **Sign up** (free, ~1 min) if you don't have an account.
2. Click **Add a domain** (a.k.a. "Add site") ? type `khelsutra.guru` ? **Continue**.
3. Choose the **Free** plan ? **Continue**.
4. Cloudflare scans existing DNS (there's basically nothing ? it's a parked domain) ? **Continue**.
5. You'll land on a screen titled something like **"Update your nameservers."** It shows **2 values** ending in `.ns.cloudflare.com` (e.g. `aria.ns.cloudflare.com` and `rob.ns.cloudflare.com`).

? Tell me those two values (or just keep them on screen). Those are what go into GoDaddy.

Step 2 ? GoDaddy: swap the nameservers (the one GoDaddy step)

On your GoDaddy tab:
1. **Domain Portfolio** (or **My Products**) ? click **khelsutra.guru**.
2. Find **Nameservers** ? **Change** (might be under a "DNS" or "Manage DNS" section).
3. Choose **"I'll use my own nameservers"** / **"Enter my own nameservers"**.
4. Replace the existing entries with Cloudflare's **two** values ? **Save**.
5. If GoDaddy warns that this moves DNS management away from GoDaddy ? that's expected, **confirm it**.

That's the whole GoDaddy part. Nothing goes down ? the domain's just parked ? and propagation runs in the background (minutes to a couple hours) while we set up the page.

Two quick checks so I guide you right:
- **Do you already have a Cloudflare account, or are you creating one now?**
- Once you've pasted the nameservers in GoDaddy and hit Save, **say "nameservers done"** and we'll go straight to deploying the page (Cloudflare Pages ? upload ? custom domain ? email).

(If you'd rather I co-navigate the Cloudflare screens with you, you can connect the Chrome extension ? but the logins and the Save click stay yours.)

271. user (2026-06-20T02:28:28.617Z)

we're ready for deploy

272. assistant (2026-06-20T02:29:01.340Z)

You've reached the fun part. Let me prep a clean deploy folder first so the upload is foolproof ? just the 3 live files, no build script.

273. user (2026-06-20T02:29:03.800Z)

```
=== ready to upload (drag this folder): artifacts/site_deploy/ ===
  total 80
```

```
drwxr-xr-x 1 avidu 197609     0 Jun 20 07:59 .
drwxr-xr-x 1 avidu 197609     0 Jun 20 07:59 ..
-rw-r--r-- 1 avidu 197609  6851 Jun 20 07:59 capture.html
-rw-r--r-- 1 avidu 197609 28646 Jun 20 07:59 index.html
-rw-r--r-- 1 avidu 197609 40635 Jun 20 07:59 og.png
```

274. assistant (2026-06-20T02:29:17.748Z)

Deploy folder ready ? **`C:\Users\avidu\Projects\badminton-highlight-indexer\artifacts\site_deploy\`** (just `index.html`, `capture.html`, `og.png`). That's the folder you'll drag in.

Step 3 ? Deploy to Cloudflare Pages

1. Cloudflare dash ? left sidebar **Workers & Pages** ? **Create** ? **Pages** tab ? **Upload assets**.
2. Project name: **`khelsutra-site`** ? **Create project**. (This only sets the temporary `*.pages.dev` subdomain; the real domain comes next.)
3. **Drag the `site_deploy` folder** into the upload box (or select all 3 files so they sit at the top level).
4. Click **Deploy site**. In ~10 seconds you'll get a live URL like **`khelsutra-site.pages.dev`**.

? Open that `*.pages.dev` URL, confirm it looks right, and paste it here ? I'll fetch it and double-check it's serving correctly (logo, OG, capture page).

Step 4 ? Point khelsutra.guru at it

In the Pages project ? **Custom domains** tab ? **Set up a domain** ? enter `khelsutra.guru` ? **Continue/Activate**. Then repeat for `www.khelsutra.guru`. Cloudflare auto-creates the DNS records + free SSL.
- If it says **"pending nameserver update,"** that just means GoDaddy's nameserver change hasn't fully propagated yet ? it'll finish on its own. The `*.pages.dev` URL works immediately regardless.

Step 5 ? Make the email CTA work

Cloudflare dash ? click the **`khelsutra.guru`** zone ? **Email** ? **Email Routing** ? **Enable/Get started**:
1. Add a **destination address** = your Gmail ? click the **verify** link Cloudflare emails you.
2. Create address **`hello@khelsutra.guru`** ? forward to that Gmail.
3. Cloudflare adds the MX/TXT records automatically. Now every "Request access" click lands in your inbox.

Do **Step 3** first and drop me the `pages.dev` link ? I'll verify it live before we wire the domain. ?

275. user (2026-06-20T02:29:53.003Z)

are you going to make this a new repo after deploying?

276. assistant (2026-06-20T02:30:27.906Z)

No ? not a new repo. Per what you picked earlier, it goes into your **existing `khelsutra` brand repo** (`github.com/avidullu/khelsutra`), under its `web/` folder, as a **single PR** ? after it's live. No fresh repo gets created.

One honest nuance worth knowing, though, since it affects the *next* step (not today's deploy):
- Today's deploy is a **direct upload** to Cloudflare Pages ? no git involved at all.
- If you later want **auto-deploy on `git push`**, Cloudflare connects to a repo. In the `khelsutra` repo I'd just set the Pages project's **"root directory"** to the site subfolder (e.g. `web/landing`) so only that ships. That works fine.
- The *only* reason you'd want a dedicated `khelsutra-site` repo is that it's marginally simpler for git auto-deploy (Pages points at the repo root, nothing else in the repo to trigger rebuilds). Both are valid ? you chose brand-cohesion, which I think is the right call.

So: **existing `khelsutra` repo, one PR, after go-live.** If you'd rather flip to a dedicated site repo for the cleaner auto-deploy, just say the word and I'll do that instead.

Nothing to decide right now though ? keep going with the upload and drop me the `*.pages.dev` link so I can verify it live. ?

277. user (2026-06-20T02:31:12.480Z)

my bad I meant the existing khelsutra repo! So we deploy first then checkin?

278. assistant (2026-06-20T02:31:30.806Z)

Exactly right: **deploy first, then check in.**

- **Now:** direct-upload to Cloudflare Pages ? live in minutes (no git needed). That's the fastest path to khelsutra.guru being up.
- **Right after it's live:** I pull the `khelsutra` repo, branch off `origin/master`, add `index.html` + `capture.html` + `og.png` (+ `make_og.py` as the build helper) under its `web/` folder, and open **one clean PR**. Version control with zero pressure on tonight's launch.

The deploy and the check-in are independent ? getting it live doesn't depend on the repo at all, so no reason to wait.

Go ahead and finish the upload (Step 3) and drop me the **`*.pages.dev` link** ? I'll fetch it and verify it's serving correctly before we wire up the custom domain and email. ?

279. user (2026-06-20T02:33:26.431Z)

there is no workers and pages

280. assistant (2026-06-20T02:34:04.484Z)

That's almost certainly a **"you're inside the domain, not at the account level"** thing ? Workers & Pages lives at the **account home**, not in the sidebar of a specific domain. When you added `khelsutra.guru`, Cloudflare dropped you *into* that domain's view, where the sidebar only shows domain stuff (DNS, SSL, Email, etc.).

Get back to account level:
1. Click the **Cloudflare logo** (top-left), or the **account name / "Account Home"** in the breadcrumb at the top. That exits the domain view.
2. Now the left sidebar should show account-wide items. Look for any of these labels (Cloudflare keeps renaming it): **"Workers & Pages"**, **"Compute (Workers)"**, **"Workers"**, or **"Pages"**.
3. Click it ? **Create** ? choose the **Pages** tab ? **Upload assets** ? drag the `site_deploy` folder.

Fastest shortcut: there's a **search bar at the very top** of the Cloudflare dashboard ? type **`Pages`** and it'll jump you straight there.

If it's *still* not showing, tell me **exactly what items you see in the left sidebar** right now (e.g. "Overview, DNS, SSL/TLS, Email, ?" vs "Websites, Workers & Pages, ?") and I'll pinpoint where to click. Newer Cloudflare accounts sometimes file it under a **"Compute"** or **"Developer Platform"** heading.

281. user (2026-06-20T02:46:42.475Z)

"Hostname 'khelsutra.guru' already has externally managed DNS records (A, CNAME, etc). Delete them first or try a different hostname." and "DNS Provider: Cloudflare
This domain's DNS is managed outside GoDaddy
Your domain is registered at GoDaddy, but its DNS is currently managed elsewhere, so DNS records can?t be updated here." -- waiting

282. assistant (2026-06-20T02:47:31.524Z)

Two things, and the first is **good news**:

? **The GoDaddy message means you're done at GoDaddy** ? it's confirming your nameservers propagated and Cloudflare now manages DNS. Don't touch GoDaddy again.

The Pages error is just **leftover GoDaddy "parking" records** that Cloudflare auto-imported when you added the domain. There's an `A`/`CNAME` record sitting on `khelsutra.guru` (pointing at GoDaddy's parking page), and Pages won't overwrite it. Delete those and retry ? 2 minutes.

Fix: clear the parking records, then re-add the domain

1. Cloudflare dash ? click the **`khelsutra.guru`** domain ? left sidebar **DNS ? Records**.
2. You'll see a few auto-imported records. **Delete** these:
   - The **`A`** record with **Name `khelsutra.guru`** (or `@`) ? that's the parking IP.
   - Any **`A`/`CNAME`** with **Name `www`** ? parking too.
   - (Safe to also delete any stray parking `CNAME`/`TXT`. **Leave** the `NS` records and a `_domainconnect` CNAME if present ? those are harmless.)
   - Since this is a freshly-parked domain, **all existing records are GoDaddy junk** ? nothing of yours to lose.
3. Go back to **Workers & Pages ? `khelsutra-site` ? Custom domains ? Set up a domain** ? enter `khelsutra.guru` again. It'll now succeed and **auto-create the correct record** (proxied) + start provisioning SSL.
4. Repeat for **`www.khelsutra.guru`** (or set a redirect ? your call).

After you delete the A record, the apex briefly has no record ? that's expected; Pages recreates the right one. SSL can take a few minutes to go green.

While you're at it ? **what's the `*.pages.dev` URL** from your deploy? Paste it and I'll fetch it right now to confirm the page itself is serving correctly (logo, OG image, capture page) so we know the upload was clean before the domain finishes wiring up.

283. user (2026-06-20T02:52:18.613Z)

can you verify

284. assistant (2026-06-20T02:52:45.919Z)

Let me check both the custom domain and the likely Pages URL in parallel.

285. user (2026-06-20T02:52:48.242Z)

The server returned HTTP 403 Forbidden.

The response body was not retrieved. If this URL requires authentication, use an authenticated tool (e.g. `gh` for GitHub, or an MCP-provided fetch tool) instead of WebFetch.

286. user (2026-06-20T02:52:49.208Z)

ECONNREFUSED

287. assistant (2026-06-20T02:53:23.738Z)

The custom domain returned **403 to my fetcher** (Cloudflare often blocks bots ? that doesn't necessarily mean it's broken for browsers), and my *guessed* `khelsutra-site.pages.dev` didn't resolve (so the real project URL is different). Let me check more reliably with a real browser user-agent.

288. user (2026-06-20T02:53:29.977Z)

```
=== https://khelsutra.guru ===
  HTTP/1.1 200 OK
  Content-Type: text/html
  Server: cloudflare
  CF-RAY: a0e77bc1bf973e53-SIN
```

```
  --- title/markers ---
  <title>KhelSutra ? every rally indexed, the dead time gone</title>
  Every rally, indexed
  Every rally, indexed
  Coming next to the alpha
=== https://www.khelsutra.guru ===
  --- title/markers ---
  <title>KhelSutra ? every rally indexed, the dead time gone</title>
  Every rally, indexed
  Every rally, indexed
  Coming next to the alpha
=== asset checks (khelsutra.guru) ===
curl: (43) A libcurl function was given a bad argument
  /og.png -> 000
curl: (43) A libcurl function was given a bad argument
  /capture.html -> 000
```

289. assistant (2026-06-20T02:53:47.895Z)

```
**It's LIVE** ? ? `https://khelsutra.guru` returns **200 OK**, served by **Cloudflare**, with
valid HTTPS, and it's *our* page:
- Title: *"KhelSutra ? every rally indexed, the dead time gone"* ?
- *"Every rally, indexed"* ? · *"Coming next to the alpha"* ?
- `www.khelsutra.guru` serves the same content ?

(The earlier 403 was just Cloudflare bot-blocking my fetcher ? with a browser user-agent it's a
clean 200.) The asset check hit a curl quoting bug, not a site issue ? let me confirm the OG
image + capture page cleanly.
```

290. user (2026-06-20T02:53:51.806Z)

```
og.png       -> 000    0 bytes
curl: (43) A libcurl function was given a bad argument
capture.html -> 000
curl: (43) A libcurl function was given a bad argument
apex http->https redirect: 200 (final: http://khelsutra.guru/)
```

291. assistant (2026-06-20T02:54:08.871Z)

```
The `-w` format tripped this curl build (not a site issue ? the apex returned 200). Let me
confirm the assets with a plain HEAD request.
```

292. user (2026-06-20T02:54:11.902Z)

```
=== /og.png ===
  HTTP/1.1 200 OK
  Content-Type: image/png
=== /capture.html ===
  HTTP/1.1 307 Temporary Redirect
```

293. assistant (2026-06-20T02:54:23.695Z)

```
`/og.png` is 200 ?. `/capture.html` returns **307** ? that's just Cloudflare Pages' default
"clean URL" behavior (it redirects `.html` ? extensionless `/capture`). Let me confirm the
redirect lands on a live page.
```

294. user (2026-06-20T02:54:27.680Z)

```
=== /capture.html redirect target ===
  Location: /capture
=== /capture (followed) ===
  final: HTTP/1.1 200 OK
  title:
    <title>KhelSutra ? 1-page capture checklist</title>
```

```
    1-page capture checklist
```

295. assistant (2026-06-20T02:55:07.776Z)

```
? **Verified ? khelsutra.guru is LIVE and healthy.** Full proof:

| Check | Result |
|---|---|
| `https://khelsutra.guru` | **200**, HTTPS valid, served by Cloudflare ? |
| Page content | Title + *"Every rally, indexed"* + *"Coming next to the alpha"* ? |
| `www.khelsutra.guru` | Same page ? |
| `/og.png` (social previews) | **200**, `image/png` ? |
| `/capture` (checklist) | **200**, *"KhelSutra ? 1-page capture checklist"* ? |

(`/capture.html` cleanly 307-redirects to `/capture` ? that's just Cloudflare Pages' default
clean-URLs behavior. The footer link works; I'll tidy it to `/capture` during the check-in so it
skips the hop.)

**One thing still needed before the page fully "works"** ?
- **Email routing (Step 5).** Right now if someone clicks **"Request access," the email goes
nowhere** until you set up `hello@khelsutra.guru` ? your Gmail in Cloudflare ? **Email ? Email
Routing**. That's the only must-do left ? the CTA depends on it.

**Two optional polish toggles (30 seconds each):**
- Cloudflare ? **SSL/TLS ? Edge Certificates ? "Always Use HTTPS"** ON (right now plain
`http://` serves without forcing https).
- Cloudflare ? **Analytics ? Web Analytics ? enable for khelsutra.guru** (automatic setup, zero
code ? your page already has the beacon placeholder; this turns it on).

---

You're **live, cofounder.** ? The domain you bought this morning is now a real product page.

Now that it's deployed ? **want me to do the repo check-in?** I'll pull the `khelsutra` repo,
branch off `origin/master`, drop the site under `web/` (matching its layout), fix the `/capture`
link, and open one clean PR. Just say go and I'll start.
```

296. user (2026-06-20T02:55:10.807Z)

```
hidden-king-khelsutra.avi-dullu.workers.dev
```

297. assistant (2026-06-20T02:55:30.882Z)

```
That's the underlying deployment URL ? so you deployed via Cloudflare's newer **Workers**
static-assets flow (`<name>.<account>.workers.dev`) rather than classic Pages. Totally fine, it
serves identically, and your custom domain is already attached to it (we just verified
khelsutra.guru ? 200). Let me confirm this URL serves the same page.
```

298. user (2026-06-20T02:55:35.465Z)

```
=== https://hidden-king-khelsutra.avi-dullu.workers.dev ===
  HTTP/1.1 200 OK
  Content-Type: text/html
  Server: cloudflare
  --- markers ---
  <title>KhelSutra ? every rally indexed, the dead time gone</title>
  class="play"
  Every rally, indexed
  class="play"
=== /og.png + /capture ===
  /og.png -> HTTP/1.1 200 OK
  /capture -> HTTP/1.1 200 OK
```

299. assistant (2026-06-20T02:55:51.648Z)

Confirmed ? the `workers.dev` URL serves the **identical page** (200, our title, play logo, `/og.png` 200, `/capture` 200). So your deployment is healthy on both the default Workers URL *and* the custom domain. Everything's consistent.

Quick orientation so it's clear:
- **`khelsutra.guru`** = the public face (what you share, what's on the page/OG).
- **`hidden-king-khelsutra.avi-dullu.workers.dev`** = Cloudflare's auto-named internal/default URL for the same deployment. Cosmetic only ? you can rename the Worker later for a tidier default URL, but no one needs to see it.

So the launch is **done and verified end-to-end.** ? The only outstanding item that affects functionality is **email routing** (so the "Request access" CTA actually delivers to your Gmail) ? everything else is optional polish.

Ready for the **repo check-in** whenever you want it: I'll pull `khelsutra`, branch off `origin/master`, add the site under `web/`, fix the `/capture` link, and open one clean PR. Say go and I'll start. ?

300. user (2026-06-20T02:55:53.853Z)

do we have the right SSLs so that some security config doesnt flag this as unprotected?

301. assistant (2026-06-20T02:56:25.057Z)

Good instinct to check ? let me actually inspect the live cert + redirect behavior rather than assume.

302. user (2026-06-20T02:56:40.337Z)

```
=== HTTP -> HTTPS redirect? (apex) ===
  HTTP/1.1 200 OK
=== HTTPS headers (apex) ===
  HTTP/1.1 200 OK
  Server: cloudflare
=== TLS cert: khelsutra.guru ===
  issuer=C=US, O=Google Trust Services, CN=WE1
  subject=CN=khelsutra.guru
  notBefore=Jun 20 01:51:30 2026 GMT
  notAfter=Sep 18 02:51:27 2026 GMT
=== SAN (does cert cover apex + www?) ===
      DNS:khelsutra.guru, DNS:*.khelsutra.guru
=== TLS version / cipher ===
=== www cert valid? ===
```

303. assistant (2026-06-20T02:57:26.051Z)

Checked the live cert and redirect behavior. **The encryption itself is solid ? but there's one real gap a security scanner *will* flag.** Here's the honest picture:

? Certificate is correct and trusted:
- Issued by **Google Trust Services** (Cloudflare's Universal SSL CA) ? publicly trusted, no warnings.
- Covers **`khelsutra.guru` *and* `*.khelsutra.guru`** (SAN), so `www` is covered by the wildcard too (and we already saw `www` serve over HTTPS fine).
- Valid **Jun 20 ? Sep 18 2026**, auto-renewing. HTTPS to the site validates cleanly (no chain errors).

?? The thing that gets flagged ? no forced HTTPS:
- `http://khelsutra.guru` currently returns **200 over plain HTTP** ? it does **not** redirect to HTTPS. So the page is reachable insecurely, and browsers/scanners will flag "available over HTTP / not secure."
- There's also **no HSTS header**, which security audits commonly flag.

Fix ? 2 toggles in Cloudflare (1 minute), no redeploy:

Cloudflare dash ? **`khelsutra.guru`** ? **SSL/TLS ? Edge Certificates**:
1. **"Always Use HTTPS" ? ON.** This makes `http://` 301-redirect to `https://`. *This is the one that closes the main flag.*
2. **"HTTP Strict Transport Security (HSTS)" ? Enable** ? max-age ~6 months. (One caveat: HSTS is sticky ? once set, browsers force HTTPS for the domain ? but for a static site that's always on HTTPS, it's exactly what you want.)
3. *(Optional hygiene)* same page ? **"Minimum TLS Version" ? 1.2**.

Also worth a glance: **SSL/TLS ? Overview ? encryption mode** should read **Full (strict)** ? though for a Workers/Pages deployment the origin *is* Cloudflare, so this is largely automatic; just make sure it's **not** set to "Flexible."

Net: your data-in-transit is genuinely protected (valid TLS cert), but flip **"Always Use HTTPS" ON** so nothing's reachable over plain HTTP ? that's the toggle that keeps a security scanner from marking it "unprotected." Want me to re-verify the http?https redirect once you've flipped it?

304. user (2026-06-20T03:02:08.790Z)

please merge all the PRs and I won't connect github to cloudflare as yet

305. assistant (2026-06-20T03:03:13.139Z)

Before I merge anything, two things matter: I shouldn't blindly merge PRs I haven't created/reviewed, and I need to set up the landing-page check-in correctly. Let me see what's actually open and check the `khelsutra` repo state first.

306. user (2026-06-20T03:03:20.085Z)

```
=== OPEN PRs ? khelsutra ===

=== OPEN PRs ? badminton-highlight-indexer ===

=== khelsutra repo state ===
current branch: master
origin/master tip: a6f616c init: KhelSutra Guru product repo
working tree (dirty?):
--- web/ layout (tracked files) ---
web/README.md
```

307. assistant (2026-06-20T03:04:07.190Z)

No open PRs anywhere ? so "all the PRs" means **the landing-page check-in** (one clean PR), which I'll create and merge. The `khelsutra` repo is a fresh init on `master`, clean, with only `web/README.md`. Let me read that + the gitignore so I place files correctly and `og.png` isn't excluded.

308. user (2026-06-20T03:04:07.710Z)

```
1       SPA scaffold lands here at execution-plan milestone M1 (Vite + React + TS + Tailwind).
2       Dev: talks to the engine at http://localhost:8000 via Vite proxy; no auth locally
(DEV_USER_EMAIL).
3
```

309. user (2026-06-20T03:04:10.106Z)

```
=== khelsutra/.gitignore ===
node_modules/
dist/
.env
.env.*
*.local
.DS_Store
Thumbs.db
```

```
=== would og.png be ignored? ===
  not ignored (OK)
=== root README (first 30 lines) ===
```

KhelSutra Guru ?

```
**Private product repo** ? the heart and brain of [khelsutra.guru](https://khelsutra.guru):
the web app (alpha SPA) now, the hosted control plane at beta.

- **Engine** (pipeline, API, eval): sibling repo `badminton-highlight-indexer` (rename pending).
- **Canonical planning docs** (live in the engine repo's `docs/` for now, migrate here at beta):
  `UI_PROPOSAL.md` · `HOSTING_PLAN.md` · `UI_ALPHA_EXECUTION_PLAN.md`.
- **Status:** pre-alpha scaffold. SPA lands in `web/` per execution-plan milestone M1.
```

Structure

```
- `web/` ? React SPA (alpha; Vite + React + TS + Tailwind; Cloudflare Pages build root)
- `docs/` ? product docs as they migrate from the engine repo
- `control-plane/` ? (beta) Cloud Run FastAPI control plane
```

License

```
All rights reserved (private; no OSS license).
```

310. assistant (2026-06-20T03:05:04.407Z)

```
`web/` is earmarked for the React SPA ("Cloudflare Pages build root"), so the **marketing
landing page belongs in its own top-level `site/`**, not `web/` ? cleaner separation. (Course-
correcting from what I said earlier now that I've read the repo's README.) `og.png` won't be
gitignored. Let me set up the branch and stage the files.
```

311. user (2026-06-20T03:05:08.126Z)

```
branch 'feat/marketing-landing-page' set up to track 'origin/master'.
Switched to a new branch 'feat/marketing-landing-page'
=== branch + files staged in working tree ===
branch: feat/marketing-landing-page
  total 84
  drwxr-xr-x 1 avidu 197609     0 Jun 20 08:35 .
  drwxr-xr-x 1 avidu 197609     0 Jun 20 08:35 ..
  -rw-r--r-- 1 avidu 197609  6851 Jun 20 08:35 capture.html
  -rw-r--r-- 1 avidu 197609 28646 Jun 20 08:35 index.html
  -rw-r--r-- 1 avidu 197609  2262 Jun 20 08:35 make_og.py
  -rw-r--r-- 1 avidu 197609 40635 Jun 20 08:35 og.png
```

312. assistant (2026-06-20T03:05:22.978Z)

```
Branch is set from `origin/master`, files staged. Now let me fix `make_og.py`'s output path for
its new location, add a `site/README.md`, and document `site/` in the root README. Reading the
two files I need to edit first.
```

313. user (2026-06-20T03:05:23.936Z)

```
42      d.text((M, 436), "A full match becomes a clean highlight reel ? only the real play.",
43             font=reg(30), fill=MUT)
44
45      # --- badge pill ---
46      pill = "Invite-only alpha · Bengaluru"
47      pf = bold(24)
48      pw = d.textlength(pill, font=pf)
49      d.rounded_rectangle([M, 506, M + pw + 44, 552], radius=23, fill=(199, 242, 176, 255),
50                          outline=None)
51      # darker lime text on lime pill
```

314. user (2026-06-20T03:05:24.098Z)

```
1       # KhelSutra Guru ?
2
3       **Private product repo** ? the heart and brain of
[khelsutra.guru](https://khelsutra.guru):
4       the web app (alpha SPA) now, the hosted control plane at beta.
5
6       - **Engine** (pipeline, API, eval): sibling repo `badminton-highlight-indexer` (rename
pending).
7       - **Canonical planning docs** (live in the engine repo's `docs/` for now, migrate here
at beta):
8         `UI_PROPOSAL.md` · `HOSTING_PLAN.md` · `UI_ALPHA_EXECUTION_PLAN.md`.
9       - **Status:** pre-alpha scaffold. SPA lands in `web/` per execution-plan milestone M1.
10
11      ## Structure
12
13      - `web/` ? React SPA (alpha; Vite + React + TS + Tailwind; Cloudflare Pages build root)
14      - `docs/` ? product docs as they migrate from the engine repo
15      - `control-plane/` ? (beta) Cloud Run FastAPI control plane
16
17      ## License
18
19      All rights reserved (private; no OSS license).
20
```

315. user (2026-06-20T03:05:44.529Z)

```
1       """Generate og.png (1200x630) ? the social-share card for khelsutra.guru.
2       Re-run after copy/brand tweaks:  python artifacts/landing/make_og.py
3       """
4       from PIL import Image, ImageDraw, ImageFont
5
6       W, H = 1200, 630
7       INK, INK2 = (12, 38, 31), (20, 62, 47)
8       GREEN, LIME = (25, 179, 107), (199, 242, 176)
```

316. user (2026-06-20T03:06:06.352Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\make_og.py has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

317. user (2026-06-20T03:06:06.560Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\make_og.py has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

318. assistant (2026-06-20T03:06:10.591Z)

site/ ? public marketing landing page

```
The static one-page site served at **[khelsutra.guru](https://khelsutra.guru)** (plus the
`/capture` one-page recording checklist). Plain HTML/CSS, one web font, no build step.
```

Files

```
- `index.html` ? the landing page.
- `capture.html` ? the printable "1-page capture checklist" (linked from the footer; served at
`/capture`).
- `og.png` ? 1200×630 social-share image (referenced by the OG / Twitter meta tags).
- `make_og.py` ? regenerates `og.png` (Pillow + Windows fonts). Re-run after brand/copy tweaks:
`python site/make_og.py`.
```

Deploy (current)

```
Deployed to **Cloudflare** via **direct upload** of this folder (Workers static assets);
custom domain `khelsutra.guru` + `www`, free SSL, and email routing for `hello@khelsutra.guru`.

> Git-based auto-deploy is **not connected yet** ? pushing to this repo does **not** redeploy.
> To publish changes, re-upload this folder in the Cloudflare dashboard.
```

Notes

```
- The `hello@khelsutra.guru` "Request access" CTA relies on Cloudflare Email Routing.
- The Cloudflare Web Analytics beacon in `index.html` has a placeholder token
  (`REPLACE_WITH_CF_TOKEN`) ? paste the real token, or enable Automatic Setup in Cloudflare
  and delete the snippet.
- Minor cosmetic follow-up: the footer link points to `/capture.html` (Cloudflare 307-redirects
  it to `/capture`); can be changed to `/capture` directly on the next content edit + re-upload.
```

319. user (2026-06-20T03:06:10.626Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\README.md (file state is
current in your context ? no need to Read it back)
```

320. user (2026-06-20T03:06:12.138Z)

```
The file C:\Users\avidu\Projects\khelsutra\README.md has been updated successfully. (file state
is current in your context ? no need to Read it back)
```

321. assistant (2026-06-20T03:06:33.112Z)

```
Files ready. Now staging explicit paths, re-verifying HEAD (git-safety), and committing.
```

322. user (2026-06-20T03:06:40.330Z)

```
warning: in the working copy of 'README.md', LF will be replaced by CRLF the next time Git
touches it
warning: in the working copy of 'site/README.md', LF will be replaced by CRLF the next time Git
touches it
warning: in the working copy of 'site/capture.html', LF will be replaced by CRLF the next time
Git touches it
warning: in the working copy of 'site/index.html', LF will be replaced by CRLF the next time Git
touches it
warning: in the working copy of 'site/make_og.py', LF will be replaced by CRLF the next time Git
touches it
branch: feat/marketing-landing-page
HEAD before commit: a6f616c init: KhelSutra Guru product repo
=== staged (explicit paths only) ===
  M     README.md
  A     site/README.md
  A     site/capture.html
  A     site/index.html
  A     site/make_og.py
  A     site/og.png
=== unstaged/untracked left behind (should be empty) ===
[feat/marketing-landing-page f81f908] feat(site): add the khelsutra.guru marketing landing page
 6 files changed, 670 insertions(+)
 create mode 100644 site/README.md
 create mode 100644 site/capture.html
 create mode 100644 site/index.html
 create mode 100644 site/make_og.py
 create mode 100644 site/og.png
=== after commit ===
  f81f908 feat(site): add the khelsutra.guru marketing landing page
  a6f616c init: KhelSutra Guru product repo
=== origin/master..HEAD (MUST be only my commit) ===
  f81f908 feat(site): add the khelsutra.guru marketing landing page
```

323. assistant (2026-06-20T03:06:57.511Z)

Clean commit ? `origin/master..HEAD` shows **only my commit** (no siblings rode in), explicit paths only. Pushing and opening the PR.

324. user (2026-06-20T03:07:07.281Z)

```
remote:
branch 'feat/marketing-landing-page' set up to track 'origin/feat/marketing-landing-page'.
To https://github.com/avidullu/khelsutra.git
 * [new branch]      feat/marketing-landing-page -> feat/marketing-landing-page
=== create PR ===
https://github.com/avidullu/khelsutra/pull/1
```

325. assistant (2026-06-20T03:07:20.446Z)

PR **#1** is up: https://github.com/avidullu/khelsutra/pull/1. Now merging it (owner-authorized) and tidying the local repo back to a clean `master`.

326. user (2026-06-20T03:07:30.548Z)

```
=== verify merged ===
state=MERGED merged=2026-06-20T03:07:26Z commit=5d6eb83a1
=== sync local master ===
Switched to branch 'master'
 create mode 100644 site/make_og.py
 create mode 100644 site/og.png
Deleted branch feat/marketing-landing-page (was f81f908).
=== master now ===
  5d6eb83 feat(site): add the khelsutra.guru marketing landing page (#1)
  a6f616c init: KhelSutra Guru product repo
=== site/ tracked on master ===
  site/README.md
  site/capture.html
  site/index.html
  site/make_og.py
  site/og.png
```

327. assistant (2026-06-20T03:08:10.654Z)

Done ? merged. Here's the honest accounting:

PR #1 ? merged ?

[avidullu/khelsutra#1](https://github.com/avidullu/khelsutra/pull/1) ? squash-merged into `master` (commit `5d6eb83`), branch deleted, local `master` synced. The landing site now lives in version control at **`khelsutra/site/`** (`index.html`, `capture.html`, `og.png`, `make_og.py`, `README.md`), with `site/` documented in the root README.

"All the PRs" was exactly one ? I checked both repos first and there were **no other open PRs** (khelsutra or the indexer), so nothing's left hanging.

A few things worth being explicit about:
- **The live site is unaffected** ? this was purely version control. khelsutra.guru still serves from the Cloudflare direct-upload deployment.
- **GitHub?Cloudflare is NOT connected** (per your call). So pushing to this repo does *not* redeploy. To publish future changes: re-upload the `site/` folder in the Cloudflare dashboard. (I noted this clearly in `site/README.md` so it's not a surprise later.)
- **I did *not* touch the indexer repo** ? the `artifacts/` scratch + the temp `launch.json` stay as local-only scratch; marketing HTML doesn't belong in the engine repo.

Let me jot the live-state into memory so a future session knows where the site lives.

328. user (2026-06-20T03:08:16.350Z)

The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\khelsutra-domain-launch.md has been updated successfully. (file state is current

in your context ? no need to Read it back)

329. assistant (2026-06-20T03:08:29.435Z)

All wrapped, cofounder. ?

Today you went from a parked domain to a live, version-controlled product page:
- ? **khelsutra.guru is live** ? landing page (rally-index + dead-time positioning, "Khel ? Sutra" play-triangle mark, "coming next" framing for the conversational builder), printable `/capture` checklist, OG share image ? all on Cloudflare with valid SSL.
- ? **Checked in & merged** ? [khelsutra#1](https://github.com/avidullu/khelsutra/pull/1) into `master`; source lives at `khelsutra/site/`. (No git?Cloudflare auto-deploy, exactly as you wanted.)

Two things still on your side, whenever you're ready (both Cloudflare, ~1 min each, no code):
1. **Email Routing ? `hello@khelsutra.guru` ? your Gmail** ? until this is set, the "Request access" button delivers nowhere. This is the one that actually matters.
2. **SSL/TLS ? Edge Certificates ? "Always Use HTTPS" ON** ? closes the "reachable over plain HTTP" security flag we found.

Ping me after you flip either and I'll re-verify live. Otherwise ? that's a hell of a morning's work. Congrats on the launch. ?

330. user (2026-06-20T03:08:32.621Z)

can you tell ABOUT EMAIL REDIRECT

331. assistant (2026-06-20T03:09:06.108Z)

Sure ? here's the full picture on the **email redirect** (Cloudflare calls it **Email Routing**), which is what makes your `hello@khelsutra.guru` "Request access" button actually work.

What it is

It's **free email forwarding**. You don't get a new inbox to log into ? instead, anything sent to `hello@khelsutra.guru` is **automatically forwarded to an inbox you already use** (your Gmail). So when someone clicks "Request access" on the site, the email lands in your normal Gmail, and you reply from there.

- It's **receive-only forwarding** ? think of it as a mail-forwarding sticker on a PO box, not a new mailbox.
- It's **free**, runs entirely on Cloudflare (works because your domain's DNS is now on Cloudflare).
- You can have many addresses: `hello@`, `avi@`, `support@` ? all forwarding wherever you want.

How to set it up (~2 min, no code)

Cloudflare dashboard ? click the **`khelsutra.guru`** domain ? left sidebar **Email ? Email Routing** ? **Get started / Enable**, then:
1. **Add a destination address** = your Gmail (e.g. `avi.dullu@gmail.com`). Cloudflare sends a **verification email** to that Gmail ? open it and click the confirm link. (One-time; proves you own the inbox.)
2. **Create a custom address**: `hello` @ `khelsutra.guru` ? **Action: Send to** ? pick your verified Gmail ? **Save**.
3. *(Optional)* Toggle **Catch-all** so *anything*@khelsutra.guru (typos, future names) also forwards to you.

Cloudflare then **auto-adds the DNS records** it needs ? three **MX** records (`*.mx.cloudflare.net`) and an **SPF TXT** record. You don't touch those; it does it for you. Since khelsutra.guru is a fresh domain with no existing email, there's nothing to conflict with.

That's it ? test it by emailing `hello@khelsutra.guru` from any account; it should appear in your Gmail within seconds.

The one honest caveat: receiving ? sending

Email Routing **forwards inbound mail only.** When you reply, by default it goes **from your Gmail address**, not from `hello@khelsutra.guru`. For an invite-only alpha that's totally fine ? you'll just reply personally.

If later you want replies to **appear as `hello@khelsutra.guru`**, that's a separate step (Gmail "Send mail as" needs an outbound SMTP relay, which Cloudflare Routing doesn't provide ? you'd add a small sender service like a transactional email provider). Not worth it now; flag it for beta.

Why it matters right now

Your landing page's button is `mailto:hello@khelsutra.guru?subject=KhelSutra alpha access?`. Until Email Routing is set up, **that address doesn't exist** and those emails bounce. Once you do the 3 steps above, every "Request access" click quietly lands in your inbox. ?

Want me to verify it's working once you've set it up? I can send a quick check or confirm the MX/SPF records resolved correctly.

332. user (2026-06-20T03:15:09.100Z)

I have set it up, can you verify?

333. assistant (2026-06-20T03:15:50.825Z)

Let me check the DNS records Email Routing should have added (MX + SPF) using a public resolver, so we're seeing the real propagated state.

334. user (2026-06-20T03:15:55.141Z)

```
=== MX records (khelsutra.guru) ===
  25 route1.mx.cloudflare.net.
  41 route3.mx.cloudflare.net.
  37 route2.mx.cloudflare.net.
=== TXT / SPF (khelsutra.guru) ===
  v=spf1 include:_spf.mx.cloudflare.net ~all
=== DMARC (_dmarc.khelsutra.guru) ===
  (none - optional)
```

335. assistant (2026-06-20T03:16:22.002Z)

Verified ? the **email infrastructure is correctly configured and live** ?

Record	Found	Verdict
MX	`route1` / `route2` / `route3.mx.cloudflare.net`	? Cloudflare Email Routing's mail servers ? inbound mail to the domain now routes to Cloudflare
SPF (TXT)	`v=spf1 include:_spf.mx.cloudflare.net ~all`	? Correct SPF for Cloudflare forwarding
DMARC	none	? Optional ? not needed for forwarding to work

Those MX/SPF records **only get added when Email Routing is enabled**, so this confirms it's switched on and propagated publicly. Mail sent to your domain will hit Cloudflare and forward per your rule.

Two honest caveats ? DNS can't see these, so I'm trusting your setup + recommending one quick test:
1. DNS doesn't show the *specific* `hello@ ? your Gmail` rule or whether the destination Gmail finished its **verification click** ? those live inside Cloudflare's UI. You said you set them up, so it should be good.
2. The only true end-to-end proof is a real message. **10-second test:** from your phone or any other email account, send a blank email to **`hello@khelsutra.guru`** ? it should land in your Gmail within a few seconds. (Don't send it *from* the same Gmail you're forwarding to ? some providers dedupe self-sends.)

Do that and your **"Request access" button is fully wired** ? clicks ? your inbox. ?

Optional, later (not needed now): add a **DMARC** record (`_dmarc.khelsutra.guru` TXT ? `v=DMARC1; p=none; rua=mailto:hello@khelsutra.guru`) to improve deliverability and anti-spoofing once you start sending. Want me to re-verify the **"Always Use HTTPS"** toggle too while we're at it, or send a follow-up once you've done the test email?

336. user (2026-06-20T03:16:24.677Z)

you can read my gmail right?

337. assistant (2026-06-20T03:17:04.064Z)

Yes ? there's a **Gmail integration connected to this session**, so I can search and read your mail. I'll keep it **tightly scoped to the email-routing check only** (searching for messages to `hello@khelsutra.guru` and the Cloudflare setup mail) ? I'm not going to browse your inbox. Let me load the tool and look.

338. assistant (2026-06-20T03:17:28.792Z)

Searching narrowly ? just messages tied to the domain / `hello@khelsutra.guru` in the last week (subjects + snippets only, no full bodies).

339. user (2026-06-20T03:17:30.652Z)

{"resultCountEstimate":"2","threads":[{"id":"19ee2ff66f24c05b","messages":[{"id":"19ee2ff66f24c05b","labelIds":["UNREAD","Label_8940358401084726341"],"snippet":"Adds the public marketing landing page served at https://khelsutra.guru. site/index.html ? landing page (rally-index + dead-time positioning; "Coming next to the alpha" for the conversational"},{"id":"19ee2ffabd87f3fa","labelIds":["UNREAD","Label_8940358401084726341"],"snippet":"Merged #1 into master. ? Reply to this email directly, view it on GitHub, or unsubscribe. You are receiving this because you are subscribed to this thread. Message ID: <avidullu/khelsutra/pull/1/"}]},{"id":"19ee2e60bac0f079","messages":[{"id":"19ee2e60bac0f079","labelIds":["UNREAD","Label_8940358401084726341","IMPORTANT"],"snippet":"khelsutra.guru is now faster and more secure, thanks to Cloudflare's network. Here are tips on how to improve performance even more. Cloudflare logo Cloudflare's network is now boosting"}]}]}

340. assistant (2026-06-20T03:17:57.133Z)

Yes ? I can. Here's exactly what my scoped search turned up (last 7 days, domain-related only):

- ? **Cloudflare welcome email** ? *"khelsutra.guru is now faster and more secure, thanks to Cloudflare's network"* ? confirms the domain is active on Cloudflare.
- ? **GitHub PR #1 notifications** ? the landing-page PR + its merge (so GitHub?Gmail works, unrelated to routing).

But here's the honest catch: **I don't see any email *to* `hello@khelsutra.guru` yet** ? because none has been sent. So what's proven is the *infrastructure* (DNS MX/SPF + domain active), not the actual forwarding loop. I can't confirm a forward that hasn't happened.

The definitive test (and I'll verify it for you):
1. From your **phone or any non-Gmail account** (not the Gmail it forwards to ? self-sends can get deduped), send a quick email to **`hello@khelsutra.guru`** ? subject like "routing test".
2. Tell me **"sent"** and I'll re-search your inbox and confirm it forwarded through ? that's the real end-to-end proof that the "Request access" button works.

One note while I'm in here: I also didn't spot the Cloudflare **Email Routing destination-verification** email in the results ? if you already clicked its verify link and it got archived, that's fine; the test above will settle it definitively either way. Want to fire off that test email?

341. user (2026-06-20T05:34:50.903Z)

I saw the webpage and can we iterate on it with the same diligence as of the badminton

highlihgts indexer repository and add the relevant docs and next steps etc. As a starter, based on your understandin of the repo requirements, can we update the Request Access to serve both Academies and Individual enthusiasts and ask questions around "do you GPU machines for do the processing or how would you store your videos so the tool can process them" in a form and collect these reponses but also leave a free text message as right nwo

342. assistant (2026-06-20T05:37:38.924Z)

Yes ? let's bring the indexer's discipline to the `khelsutra` repo: I'll treat this as a proper tracked work item. Concretely, mirroring the engine repo's conventions, this iteration gets:
- A **`NEXT_STEPS.md`** in `khelsutra/` (the live "what's next" for the brand/site).
- A **tracked project doc** (e.g. `docs/REQUEST_ACCESS_FORM.md`) using the same `PROJECT_DOC_TEMPLATE` shape ? status header · decisions-locked · progress tracker · definition-of-done.
- Shipped as a **PR** to `khelsutra/master` (no auto-deploy, so it goes live when you re-upload).

And your instinct on the two questions is sharp ? **"do you have a GPU?" and "where do your videos live?" are exactly the compute-placement and storage-backend signals** that decide how we'd serve each user (the BYO-compute / BYO-storage + autosync direction we mapped earlier). So this form doubles as PMF + architecture research, not just lead capture.

Proposed form

- **You are:** ? Coaching academy / club ? Individual player *(drives the framing)*
- **Name** · **Email** *(so we can reach you)* · **City** *(+ academy/club name if academy)*
- **Courts you record** *(academy only ? 1 / 2?4 / 5+; optional)*
- **Do you have a GPU machine that could process video locally?** ? Yes (gaming/GPU PC or laptop) / Not sure / No / I'd prefer cloud
- **Where do your recordings live / how would you share them?** *(pick any)* ? Google Drive · OneDrive (MS 365) · Dropbox · cloud bucket (S3/GCS) · SD card / local drive · Not sure
- **Anything else?** ? free-text box *(same open message as today)*
- ? Okay to email me about the KhelSutra alpha

The one decision that unblocks "collecting responses"

The page is **static** right now, so a form needs *somewhere to send data*. This is a genuine fork (privacy + how you deploy from here on), so I'd lock it in the tracked doc:

343. user (2026-06-20T05:37:53.797Z)

Your questions have been answered: "How should the Request Access form collect & store submissions? (The site is static today, so the form needs a backend.)"="Cloudflare Worker + D1 (owned) ? Recommended". You can now continue with these answers in mind.

344. assistant (2026-06-20T05:41:30.471Z)

Excellent choice ? **Cloudflare Worker + D1** keeps the lead data in your own account, queryable, and on-brand with the BYO/privacy ethos. And those two questions become real architecture signals you can later `SELECT` over.

I'll bring the indexer's discipline: a **tracked design doc** (decisions-locked + progress tracker + definition-of-done) and a **`NEXT_STEPS.md`**, plus the form itself ? as a PR to `khelsutra`. Let me branch and lay down the docs first.

345. user (2026-06-20T05:41:35.313Z)

Switched to a new branch 'feat/request-access-form'
branch: feat/request-access-form base: 5d6eb83 feat(site): add the khelsutra.guru marketing landing page (#1)
  docs: README.md

346. assistant (2026-06-20T05:41:40.241Z)

Branch is set from `origin/master`. Now the tracked design doc.

347. assistant (2026-06-20T05:42:16.577Z)

Request-Access form ? segmented lead capture + setup discovery

```
> **Status:** `IN PROGRESS` · **Owner:** Avi · **Created:** 2026-06-20 · **Last updated:**
2026-06-20
> **Lifecycle:** `DRAFT ? IN PROGRESS ? DONE ? archived`
> **Relation:** first tracked work item in the `khelsutra` repo; mirrors the engine repo's
> `PROJECT_DOC_TEMPLATE` discipline (status header · decisions-locked · §6 tracker · DoD).
> Feeds the platform direction in the engine repo's `PLATFORM_ARCHITECTURE.md §3`
(StorageBackend /
> ComputeBackend) and `VIDEO_LOCALITY_MODEL.md` ? see §7.

---
```

0. TL;DR

```
Replace the landing page's `mailto:` "Request access" CTA with a real form that (a) segments
**academies vs individual enthusiasts**, (b) asks the two questions that actually decide how
we'd
serve a user ? **"do you have a GPU machine?"** (compute placement) and **"where do your videos
live?"** (storage backend) ? and (c) keeps a free-text message. Submissions are stored in our
own
**Cloudflare D1** via a small **Worker** (`POST /api/request-access`). The form doubles as PMF +
architecture research.
```

1. Problem & goal

```
Today the CTA is `mailto:hello@khelsutra.guru` ? zero structure, nothing collected, no
segmentation.
We want **structured, queryable** interest signals from two distinct audiences, captured the
moment
someone raises a hand, **without shipping lead data to a third party** (privacy-first ethos).
**Goal:** a clean on-brand form whose responses land in our own Cloudflare account, queryable
later.
```

2. Decisions locked

```
- **Backend = Cloudflare Worker + D1** (owner decision, 2026-06-20). Data stays in our
Cloudflare
  account; queryable; no third party. (Alternatives considered: hosted form (Tally/Formspree ?
  third-party data), Worker?email-only (no DB), structured `mailto` (no collection). Rejected
for
  the reasons in parens.)
- **Two audience segments:** `academy` and `individual` ? the form lightly adapts (academy also
  asks club name + #courts).
- **The two discovery questions are first-class** (not optional polish): GPU/compute + storage ?
  they map 1:1 onto the BYO-compute / BYO-storage architecture forks (§7).
- **Free-text message retained** ("Anything else?").
- **Privacy:** store only what's submitted + a coarse country (from `request.cf.country`); **no
raw
  IP**; consent checkbox for "email me about the alpha"; this data is **operations/contact
only**,
  never a training source (same guardrail as the engine repo).
- **Deploy model changes:** the site moves from dashboard direct-upload ? **Wrangler** (Worker
with
  static assets + D1 binding). Documented in §5; owner runs the `wrangler` deploy (our
account/auth).
```

3. Form spec

```
| Field | Type | Notes |
|---|---|---|
| Segment | radio | `academy` \| `individual` (required) |
| Name | text | required |
| Email | email | required (so we can reply) |
```

City	text	optional
Club / academy name	text	shown for `academy`
Courts recorded	select	`1` | `2-4` | `5+` (academy; optional)
GPU machine for local processing?	radio	`yes` | `unsure` | `no` | `prefer-cloud`
Where do your recordings live / how would you share them?	checkboxes (multi)	`gdrive` | `onedrive` | `dropbox` | `bucket` (S3/GCS) | `local` (SD/drive) | `unsure`
Message	textarea	free text ("Anything else?")
Consent	checkbox	"Okay to email me about the KhelSutra alpha"
`hp`	hidden honeypot	anti-spam; if filled ? silently drop

4. Design / architecture

- **Frontend** (`site/index.html`): the header + hero "Request access" buttons scroll to the `#join` form; `#join` becomes the form. On submit, JS `POST`s JSON to `/api/request-access`. **Progressive enhancement:** if the endpoint isn't live yet (pre-Worker), it falls back to a prefilled `mailto:hello@` so the form is useful from the first re-upload.
- **Worker** (`/api/request-access`): validates (email present, honeypot empty), inserts into D1, returns `200 {ok:true}`; serves the static assets for all other routes.
- **D1 schema** (`access_requests`): `id, created_at, segment, name, email, city, org_name, courts, has_gpu, storage (csv), message, consent, country, user_agent`.
- **Anti-spam:** honeypot first; add **Cloudflare Turnstile** in P2 if spam appears.
- **Notify (P2):** optional email ping to `hello@` per submission (MailChannels / Cloudflare Email).

5. Deploy (owner-run; our Cloudflare account)

```
wrangler d1 create khelsutra-access            # once; copy the database_id into wrangler.toml
wrangler d1 execute khelsutra-access --file=site/schema.sql   # create the table
wrangler deploy                                # ships the Worker + static assets + D1 binding
```

> Until the Worker ships, the form works via the `mailto` fallback (re-upload `site/` as today).
> Git?Cloudflare auto-deploy still intentionally not connected.

6. Deliverables & progress tracker ? source of truth

Legend: ? Todo · ? In progress · ? Done. One PR per row off `origin/master`.

ID	Deliverable	Gated?	Status	PR
P0	This design doc + `NEXT_STEPS.md`	No	?	this PR
P1	Form UI in `site/index.html` (segments + GPU + storage + free text + consent) w/ `mailto` fallback	No	?	this PR
P2	Worker + D1 + `wrangler.toml` + `schema.sql` (real collection)	owner deploy	?	?
P3	Anti-spam (Turnstile) + per-submission email notify	No	?	?
P4	Tiny admin/export view (auth'd) to read responses	owner	?	?

7. Why this matters beyond lead capture

The two questions are **the architecture forks** from the engine repo's BYO analysis:

- **GPU?** ? compute placement: desktop/local (their GPU) vs hosted vs zero-infra-BYO (`PLATFORM_ARCHITECTURE.md §3.2`, `DESKTOP_APP_PLAN.md` P0 spike).
- **Storage?** ? which `StorageBackend` + whether autosync is worth building (`PLATFORM_ARCHITECTURE.md §3.1`, the cloud-storage/autosync thread).

So early responses directly inform what to build next ? not just who to email.

8. Open questions

1. Turnstile now or wait for spam? (Default: honeypot now, Turnstile if needed.)
2. Email-notify per submission, or just query D1 weekly? (Default: notify in P2.)
3. Do we localize the form (Kannada/Hindi) for academies now or at beta? (Default: English now.)

9. Definition of done

- [ ] Form replaces the `mailto` CTA, segments academy/individual, asks GPU + storage + free text.
- [ ] Submissions persist to Cloudflare D1 via the Worker; honeypot drops bots.

- [ ] Privacy posture honored (no raw IP; consent; contact-only).
- [ ] Deployed live on khelsutra.guru; a test submission appears in D1.
- [ ] Doc tracker + DoD updated; archived when P1?P3 land.

348. user (2026-06-20T05:42:16.629Z)

File created successfully at: C:\Users\avidu\Projects\khelsutra\docs\REQUEST_ACCESS_FORM.md
(file state is current in your context ? no need to Read it back)

349. assistant (2026-06-20T05:42:21.935Z)

Now the repo's `NEXT_STEPS.md` (the live what's-next for the brand/site).

350. assistant (2026-06-20T05:42:35.505Z)

Next steps (live) ? KhelSutra brand / site

The prioritized "what to do next" for the **`khelsutra` product repo** (brand site now, SPA + control
plane later). Mirrors the engine repo's `docs/NEXT_STEPS.md` discipline. Refresh at each checkpoint.

Where we are (2026-06-20): `khelsutra.guru` is **LIVE** ? static marketing landing page + `/capture` checklist + OG image, on Cloudflare (Workers static-assets, direct upload; custom domain
+ free SSL + email routing for `hello@`). Source in `site/` (PR #1). Git?Cloudflare auto-deploy intentionally **not** connected ? publish by re-uploading `site/` in the dashboard.

In progress

- **Request-Access form** ? segmented (academy / individual) + GPU + storage discovery questions +
  free text, collected in **Cloudflare D1** via a Worker. Tracked: [`docs/REQUEST_ACCESS_FORM.md`](docs/REQUEST_ACCESS_FORM.md).
  - P1 form UI (this PR) ? P2 Worker + D1 (owner-run `wrangler` deploy) ? P3 anti-spam + email notify.

Queued (site)

- **Always-Use-HTTPS + HSTS** toggles in Cloudflare (close the plain-HTTP security flag).
- **Cloudflare Web Analytics** ? paste the token into `site/index.html` (or enable automatic setup).
- **`/capture` link tidy** ? footer points to `/capture.html` (307?`/capture`); change to `/capture`.
- **OG image refresh** when copy/brand changes (`python site/make_og.py`).
- **Social/DMARC** ? add `_dmarc` TXT once we start sending mail (deliverability).
- **Connect git auto-deploy** (deferred by owner) ? Cloudflare Pages/Workers Build pointed at `site/`.
- **`send mail as hello@`** for replies (needs an outbound SMTP relay) ? beta.

Bigger arcs (not started; design-first per the engine repo's convention)

- **Alpha SPA** lands in `web/` (Vite + React + TS + Tailwind) per `UI_ALPHA_EXECUTION_PLAN.md` (engine repo).
- **BYO-storage + autosync** ? connect Drive/OneDrive/GCS/S3, watch the folder, auto-prepare
  highlights (engine repo `PLATFORM_ARCHITECTURE.md §3` + the autosync thread). The Request-Access
  GPU/storage answers are the demand signal for this.
- **Control plane** (`control-plane/`) at beta.

Convention

Substantial work gets a tracked doc in `docs/` (template as in the engine repo); one PR per item off
`origin/master`; don't connect prod to git auto-deploy without owner say-so.

351. user (2026-06-20T05:42:35.532Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\NEXT_STEPS.md (file state is
current in your context ? no need to Read it back)
```

352. assistant (2026-06-20T05:42:51.884Z)

Now the form itself. Let me read the current `site/index.html` to anchor the edits precisely.

353. user (2026-06-20T05:42:52.444Z)

```
88        .reel-video .court{position:absolute;inset:16px;border:2px solid
rgba(199,242,176,.30);border-radius:6px}
89        .reel-video .court:before{content:"";position:absolute;left:50%;top:0;bottom:0;width:2
px;background:rgba(199,242,176,.30)}
90        .reel-video .court:after{content:"";position:absolute;top:50%;left:0;right:0;height:2p
x;background:rgba(199,242,176,.22)}
91        .reel-video .play{width:60px;height:60px;border-
radius:50%;background:var(--lime);display:flex;align-items:center;justify-
content:center;color:#0C261F;font-size:1.4rem;z-index:2;padding-left:4px;box-shadow:0 10px 26px
rgba(0,0,0,.35)}
92        .reel-video .badge{position:absolute;top:12px;left:12px;background:rgba(12,38,31,.72);
color:var(--lime);font-size:.7rem;font-weight:700;letter-spacing:.06em;padding:.26rem
.6rem;border-radius:999px;z-index:2}
93        .reel-video
.wm{position:absolute;bottom:10px;right:12px;color:rgba(234,243,238,.78);font-size:.66rem;font-
weight:500;z-index:2}
94        .reel-body{padding:.95rem 1.05rem 1.1rem}
95        .reel-body .lbl{font-size:.72rem;font-weight:600;letter-spacing:.04em;text-
transform:uppercase;color:var(--mut)}
96        .reel-track{position:relative;height:12px;border-
radius:6px;background:#E8ECE6;margin:.55rem 0 .95rem}
97        .reel-track span{position:absolute;top:0;bottom:0;background:var(--green);border-
radius:6px}
98        .reel-stats{display:flex;gap:10px}
99        .reel-stats .m{flex:1;background:var(--paper);border-radius:10px;padding:.6rem .85rem}
100       .reel-stats .m b{display:block;font-size:1.25rem;color:var(--tx);font-weight:700}
101       .reel-stats .m em{font-size:.76rem;color:var(--mut);font-style:normal}
102       .ask{display:flex;align-items:center;gap:.55rem;margin:.85rem 1.05rem
.2rem;padding:.6rem .7rem;border:1px solid var(--line);border-
radius:12px;background:var(--paper)}
103       .ask-ic{color:var(--green-d);font-size:.95rem}
104       .ask-tx{flex:1;font-size:.88rem;color:var(--tx);font-weight:500;line-height:1.3}
105       .ask-go{width:26px;height:26px;border-
radius:8px;background:var(--green);color:#04140D;display:flex;align-items:center;justify-
content:center;font-size:.85rem;flex:none}
106       .lbl{display:flex;justify-content:space-between;align-items:center}
107       .lbl .match{color:var(--green-d);font-weight:700}
108       .reel-track span.off{background:#D7DDD4}
109       .reel-track span.on{background:var(--green)}
110       .result{display:flex;align-items:center;gap:.7rem;margin-top:.2rem;padding:.65rem
.8rem;background:var(--ink);border-radius:12px}
111       .result .play{width:34px;height:34px;border-
radius:50%;background:var(--lime);color:#0C261F;display:flex;align-items:center;justify-
content:center;flex:none;font-size:.95rem;padding-left:2px}
112       .result-tx{flex:1}
113       .result-tx b{display:block;font-size:.84rem;font-weight:700;color:#fff}
```

354. user (2026-06-20T05:43:02.972Z)

```
48:  .btn-dark{background:var(--ink);color:#fff}
49:  .btn-dark:hover{background:var(--ink2)}
70:  .hero-cta{display:flex;flex-wrap:wrap;gap:.8rem;margin-top:2rem}
170:  .cta .hero-cta{justify-content:center}
```

```
189:</style>
195:    <a class="brand" href="#top">Khel<span class="play" aria-hidden="true"></span>Sutra</a>
202:    <a class="btn btn-dark" href="mailto:hello@khelsutra.guru?subject=KhelSutra%20alpha%20ac
cess&amp;body=Hi%20KhelSutra%20team%2C%20I'd%20like%20to%20try%20the%20alpha.%0A%0AAcademy%20%2F
%20name%3A%20%0ACity%3A%20%0ANumber%20of%20courts%3A%20%0AHow%20you%20record%20today%3A%20">Requ
est access</a>
213:      <div class="hero-cta">
214:        <a class="btn btn-primary" href="mailto:hello@khelsutra.guru?subject=KhelSutra%20alp
ha%20access&amp;body=Hi%20KhelSutra%20team%2C%20I'd%20like%20to%20try%20the%20alpha.%0A%0AAcadem
y%20%2F%20name%3A%20%0ACity%3A%20%0ANumber%20of%20courts%3A%20%0AHow%20you%20record%20today%3A%2
0">Request alpha access</a>
413:<section id="join" style="padding-bottom:80px">
418:      <div class="hero-cta">
419:        <a class="btn btn-primary" href="mailto:hello@khelsutra.guru?subject=KhelSutra%20alp
ha%20access&amp;body=Hi%20KhelSutra%20team%2C%20I'd%20like%20to%20try%20the%20alpha.%0A%0AAcadem
y%20%2F%20name%3A%20%0ACity%3A%20%0ANumber%20of%20courts%3A%20%0AHow%20you%20record%20today%3A%2
0">Request alpha access</a>
443:</body>
```

355. user (2026-06-20T05:43:12.154Z)

```
411     </section>
412
413     <section id="join" style="padding-bottom:80px">
414       <div class="wrap">
415         <div class="cta">
416           <h2>Want to see your academy's own footage as a reel?</h2>
417           <p>That's the best way to get a feel for KhelSutra. Tell us a little about your
setup and we'll get you in.</p>
418           <div class="hero-cta">
419             <a class="btn btn-primary" href="mailto:hello@khelsutra.guru?subject=KhelSutra%2
0alpha%20access&amp;body=Hi%20KhelSutra%20team%2C%20I'd%20like%20to%20try%20the%20alpha.%0A%0AAc
ademy%20%2F%20name%3A%20%0ACity%3A%20%0ANumber%20of%20courts%3A%20%0AHow%20you%20record%20today%
3A%20">Request alpha access</a>
420             <a class="btn btn-ghost"
href="mailto:hello@khelsutra.guru">hello@khelsutra.guru</a>
421           </div>
422         </div>
423       </div>
424     </section>
425
426     <footer>
427       <div class="wrap foot">
428         <div>
429           <div class="brand" style="font-size:1.05rem">Khel<span class="play" aria-
hidden="true"></span>Sutra</div>
430           <div class="stamp" style="margin-top:.4rem">Automatic badminton highlights ·
Bengaluru, India</div>
431         </div>
432         <div style="text-align:right">
433           <div><a href="/capture.html">Capture checklist</a> · <a
href="mailto:hello@khelsutra.guru">hello@khelsutra.guru</a> · <a
href="https://khelsutra.guru">khelsutra.guru</a></div>
434           <div class="stamp" style="margin-top:.4rem">© 2026 KhelSutra · Invite-only alpha ·
Reels carry a "Generated by Khelsutra.guru" mark</div>
435         </div>
436       </div>
437     </footer>
438
439     <!-- Cloudflare Web Analytics: paste your token from Cloudflare dashboard ? Analytics ?
Web Analytics ? Add a site.
440          (Or skip this snippet entirely and enable "Automatic setup" once the domain is
proxied through Cloudflare.) -->
441     <script defer src="https://static.cloudflareinsights.com/beacon.min.js" data-cf-
beacon='{"token": "REPLACE_WITH_CF_TOKEN"}'></script>
```

```
442
443     </body>
```

356. user (2026-06-20T05:43:46.934Z)

```
210           <span class="eyebrow">Invite-only alpha · Bengaluru</span>
211           <h1>Every rally, indexed. <span class="hl">The dead time, gone.</span></h1>
212           <p class="lead">KhelSutra finds every rally in your session and strips out the
knock-ups, walks and gaps ? so a full match becomes a clean highlight reel you can watch and
download. No editing, no scrubbing.</p>
213           <div class="hero-cta">
214             <a class="btn btn-primary" href="mailto:hello@khelsutra.guru?subject=KhelSutra%2
0alpha%20access&amp;body=Hi%20KhelSutra%20team%2C%20I'd%20like%20to%20try%20the%20alpha.%0A%0AAc
ademy%20%2F%20name%3A%20%0ACity%3A%20%0ANumber%20of%20courts%3A%20%0AHow%20you%20record%20today%
3A%20">Request alpha access</a>
215             <a class="btn btn-ghost" href="#how">See how it works</a>
216           </div>
217           <p class="hero-note">A phone or GoPro on a tripod is all you need. No special
hardware.</p>
218           <div class="loop">
```

357. assistant (2026-06-20T05:44:47.709Z)

```
All anchors confirmed. Now building the form ? replacing the `mailto` CTAs with a real segmented
form (academy/individual + GPU + storage + free text + consent), plus the submit JS that posts
to `/api/request-access` with a `mailto` fallback so it works even before the Worker ships.
```

358. user (2026-06-20T05:44:49.894Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\index.html has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

359. user (2026-06-20T05:44:53.171Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\index.html has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

360. user (2026-06-20T05:45:02.109Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\index.html has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

361. user (2026-06-20T05:45:17.192Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\index.html has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

362. user (2026-06-20T05:45:25.923Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\index.html has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

363. assistant (2026-06-20T05:45:36.241Z)

```
Form built. Let me validate it renders correctly ? copying the updated file to the running
preview and checking the form (and the academy-toggle behavior).
```

364. user (2026-06-20T05:45:39.620Z)

```
copied to preview dir
=== quick HTML sanity (form present, no obvious breakage) ===
  form#af count: 1
        1 name="consent"
```

```
        1 name="email"
        4 name="has_gpu"
        1 name="message"
        2 name="segment"
        6 name="storage"
```

365. assistant (2026-06-20T05:45:48.287Z)

```
Fields all present. Let me get the preview server and render the form.
```

366. user (2026-06-20T05:45:51.354Z)

```
{
  "serverId": "ff446b01-fc06-4b97-830d-352da5178798",
  "port": 8099,
  "name": "landing",
  "reused": false
}
Server started successfully on port 8099.
```

367. user (2026-06-20T05:45:55.358Z)

```
Viewport reset to native size (desktop).
```

368. user (2026-06-20T05:46:00.394Z)

```
[
  6211,
  false,
  0
]
```

369. assistant (2026-06-20T05:46:12.810Z)

```
Preview tab still showing the pre-form version (cached). Forcing a hard reload.
```

370. user (2026-06-20T05:46:13.412Z)

```
"reloading"
```

371. user (2026-06-20T05:46:26.421Z)

```
[
  "NO FORM",
  "KhelSutra ? every rally indexed, the dead time gone"
]
```

372. assistant (2026-06-20T05:46:38.975Z)

```
The renderer's still stale. Let me check what the server *actually* serves (disk + HTTP) to
separate a copy problem from a browser-cache problem.
```

373. user (2026-06-20T05:46:41.877Z)

```
=== disk: artifacts/site_deploy/index.html has the form? ===
  form#af on disk: 1
=== what the SERVER returns (curl) ===
```

374. assistant (2026-06-20T05:46:57.313Z)

```
Found it ? the preview server serves `artifacts/landing/`, but I copied to
`artifacts/site_deploy/`. Copying to the right dir and confirming the server now serves the
form.
```

375. user (2026-06-20T05:46:59.348Z)

```
copied to artifacts/landing/ (the served dir)
=== server now returns the form? ===
        1 Get into the KhelSutra alpha
        1 id="af"
        4 name="has_gpu"
```

376. assistant (2026-06-20T05:47:05.845Z)

```
Server now serves the form. Let me reload the preview and capture it.
```

377. user (2026-06-20T05:47:07.120Z)

```
"reloading"
```

378. user (2026-06-20T05:47:12.301Z)


```
{
  "gpu": 4,
  "scrollY": 6900,
  "storage": 6
}
```


379. assistant (2026-06-20T05:47:16.953Z)

```
Form's loaded (4 GPU radios, 6 storage checkboxes). Let me capture it.
```

380. assistant (2026-06-20T05:47:32.412Z)

```
The form renders cleanly ? segment toggle, the GPU question, the storage multi-select, free
text, consent, and the privacy note. Let me verify the academy-toggle (the club-name + courts
fields should appear when "Coaching academy" is picked).
```

381. user (2026-06-20T05:47:34.998Z)


```
{
  "acadFieldsDisplay": "block",
  "hasCourts": true,
  "hasOrgField": true,
  "isAcademyClass": true
}
```


382. assistant (2026-06-20T05:47:51.521Z)

```
Academy toggle works (selecting "Coaching academy" reveals the club-name + courts fields). Let
me confirm no JS errors, then commit and open the PR.
```

383. user (2026-06-20T05:47:52.088Z)

```
No console logs.
```

384. assistant (2026-06-20T05:48:16.291Z)

```
Validated: form renders cleanly, the academy-toggle works, **no JS errors**. Committing and
opening the PR.
```

385. user (2026-06-20T05:48:24.596Z)

```
warning: in the working copy of 'NEXT_STEPS.md', LF will be replaced by CRLF the next time Git
touches it
warning: in the working copy of 'docs/REQUEST_ACCESS_FORM.md', LF will be replaced by CRLF the
next time Git touches it
```

```
branch: feat/request-access-form  base HEAD: 5d6eb83 feat(site): add the khelsutra.guru
marketing landing page (#1)
=== staged ===
  A     NEXT_STEPS.md
  A     docs/REQUEST_ACCESS_FORM.md
  M     site/index.html
[feat/request-access-form e78e34b] feat(site): segmented Request-Access form (academy/individual
+ GPU + storage discovery)
 3 files changed, 289 insertions(+), 10 deletions(-)
 create mode 100644 NEXT_STEPS.md
 create mode 100644 docs/REQUEST_ACCESS_FORM.md
=== origin/master..HEAD (only my commit) ===
  e78e34b feat(site): segmented Request-Access form (academy/individual + GPU + storage
discovery)
branch 'feat/request-access-form' set up to track 'origin/feat/request-access-form'.
To https://github.com/avidullu/khelsutra.git
 * [new branch]      feat/request-access-form -> feat/request-access-form
```

386. assistant (2026-06-20T05:48:29.584Z)

Committed (only my commit on the branch) and pushed. Opening the PR.

387. user (2026-06-20T05:48:40.449Z)

https://github.com/avidullu/khelsutra/pull/2

388. assistant (2026-06-20T05:49:10.319Z)

Done ? PR **#2** is up: https://github.com/avidullu/khelsutra/pull/2

What shipped (validated live in the preview)

The `mailto` CTA is now a real **segmented form**:
- **I'm a?** Coaching academy / club · Individual player ? picking *academy* reveals club-name + #courts.
- **Do you have a GPU machine?** (Yes / Not sure / No / prefer cloud) ? the **compute-placement** signal.
- **Where do your recordings live?** (Drive · OneDrive · Dropbox · S3/GCS · SD/local · Not sure) ? the **storage-backend** signal.
- Free-text "Anything else?" + consent + the minors/privacy note. Honeypot keeps bots out.

I verified: renders cleanly, the academy-toggle works, all 14 fields present, **zero console errors**. Those two questions aren't just lead-gen ? they're literally the BYO-compute / BYO-storage forks from the platform analysis, so early answers tell you *what to build next*.

The diligence layer (as you asked)

- **`docs/REQUEST_ACCESS_FORM.md`** ? tracked design doc: decisions-locked (Worker + D1), form spec, D1 schema, API contract, privacy, the deploy-model change, a §6 progress tracker, and a definition-of-done.
- **`NEXT_STEPS.md`** ? the repo's live what's-next.

Two honest notes

1. **Not live until you publish.** No git auto-deploy, so the form goes live when you **re-upload `khelsutra/site/`** to Cloudflare (that folder is the source of truth now).
2. **Until the Worker ships, the form uses a `mailto` fallback** ? submissions open a prefilled email rather than hitting a database. That's by design (P1 ships value today; P2 wires real collection).

Your call on two things

- **Merge #2 now, or review the design doc + form first?** (I validated it; it's a low-risk static change ? but it's your repo and the doc locks decisions, so your call.)
- **Proceed to P2 ? the Worker + D1 backend?** I'd write `wrangler.toml` + the Worker (`/api/request-access` ? D1) + `schema.sql` as the next PR, and hand you the exact `wrangler d1 create` / `deploy` commands (those run under your Cloudflare account). Heads-up: P2 moves

deploys from dashboard-upload to the Wrangler CLI ? I'll document it fully.

389. user (2026-06-20T05:49:12.677Z)

and oh I verified the email working when my sister clicked on request access
should we link the github with cloudflare so that the deployments are quick and safe?

390. assistant (2026-06-20T05:50:43.746Z)

? **That's the real proof** ? your sister clicking "Request access" and it landing in your inbox confirms the whole **Email Routing loop works end-to-end** (the live `mailto` CTA ? `hello@khelsutra.guru` ? your Gmail). That's the one thing DNS couldn't prove for me earlier. Closed. ?

Should you link GitHub ? Cloudflare? **Yes ? it's both quicker *and* safer.**

The wins are exactly what you'd want:
- **Quick:** `git push` / merge ? Cloudflare builds + deploys automatically. No more manual drag-upload.
- **Safe ? three ways:**
  - **Versioned rollback** ? every deploy is tied to a commit; one click to roll back to any previous deployment.
  - **PR preview URLs** ? each PR gets its own live preview link, so you (and I) can *see* a change rendered before it touches production. That's a real upgrade to the review discipline you asked for.
  - **No drift** ? the repo becomes the single source of truth; "what's in git" == "what's live." (Literally the trap we just hit when I copied to the wrong folder ? git-deploy eliminates it.)

The timing is perfect ? pair it with the Worker (P2)

Here's the one nuance: Cloudflare deploys a Worker from a **`wrangler.toml`** in the repo (via "Workers Builds"). We're *about* to add that anyway for the form's D1 backend. So the clean sequence is:

1. **I build P2** ? `wrangler.toml` + the Worker (`/api/request-access` ? D1) + `schema.sql`, in a PR. This makes the repo "connect-ready."
2. **You connect the repo** in Cloudflare (Workers Builds ? pick `avidullu/khelsutra`, root dir `site/`, bind D1). One-time, ~5 min ? I'll give exact click-by-click steps.
3. From then on: **merge to `master` ? auto-deploys to khelsutra.guru; PRs ? preview URLs.**

One behavior change to know (it's the "safe" part): production only moves when you merge to `master`. Feature branches/PRs deploy to *preview*, not production ? so nothing goes live by accident.

A couple of setup details we'll handle during the connect: re-attaching the `khelsutra.guru` custom domain to the git-deployed Worker, and the D1 binding. All straightforward; I'll walk it.

So, where this leaves us

- **PR #2 (the form)** is still awaiting your merge call. Once git is connected, **merging it = it deploys automatically.** (Until then, you'd re-upload `site/` to publish ? but I'd just wait for the git connection so the form ships cleanly.)
- **Recommendation:** let me **build P2 now** (so the repo is connect-ready), then you do the one-time Cloudflare connect, and from that point everything is push-to-deploy with previews + rollback.

Want me to go ahead and build P2 (the Worker + D1 + `wrangler.toml`)?

391. user (2026-06-20T05:51:49.666Z)

yes please go ahead and merge these PRs as well ensuring all the tests.

392. assistant (2026-06-20T05:53:00.425Z)

On it. First: confirm Node is available (for the Worker tests) and merge the form PR, then I'll

build P2 with real tests.

393. user (2026-06-20T05:53:10.542Z)

```
=== node / npm ===
v26.1.0
11.13.0
=== merge PR #2 (form ? validated in-browser: renders, academy-toggle works, no console errors;
khelsutra has no automated suite yet, P2 adds one) ===
=== sync khelsutra master ===
Deleted branch feat/request-access-form (was e78e34b).
master now: 2274faa feat(site): segmented Request-Access form (academy/individual + GPU +
storage discovery) (#2)
```

394. assistant (2026-06-20T05:55:21.991Z)

Node 26 + npm 11 are available, and **PR #2 (form) merged** ? master is `2274faa`. Now P2: I'll convert `site/` into a proper Worker project (static assets + `/api/request-access` ? D1), with **real unit tests** I can run green before merging.

Let me branch and restructure.

395. user (2026-06-20T05:55:24.450Z)

```
Switched to a new branch 'feat/access-worker-d1'
branch: feat/access-worker-d1
=== restructure (staged moves) ===
  R  site/capture.html -> site/public/capture.html
  R  site/index.html -> site/public/index.html
  R  site/og.png -> site/public/og.png
```

396. assistant (2026-06-20T05:55:29.655Z)

Files moved. Now the Worker, config, schema, and tests.

397. assistant (2026-06-20T05:55:40.716Z)

```
// Cloudflare Worker for khelsutra.guru ? serves the static site and handles the
// Request-Access form: POST /api/request-access -> validate -> insert into D1 (binding DB).
// Pure helpers are exported so they can be unit-tested without a live Worker/D1.

const SEGMENTS = ["academy", "individual"];
const GPU = ["yes", "unsure", "no", "prefer-cloud"];
const COURTS = ["1", "2-4", "5+"];
const STORAGE = ["gdrive", "onedrive", "dropbox", "bucket", "local", "unsure"];

function str(v, max) {
  if (v == null) return null;
  const s = String(v).trim();
  return s ? s.slice(0, max) : null;
}

// Honeypot: humans never fill the hidden "company" field.
export function isBot(p) {
  return !!(p && typeof p.company === "string" && p.company.trim() !== "");
}

export function validate(p) {
  if (!p || typeof p !== "object") return { ok: false, error: "bad payload" };
  const email = String(p.email || "").trim();
  if (!/^[^@\s]+@[^@\s]+\.[^@\s]+$/.test(email)) return { ok: false, error: "valid email
required" };
  return { ok: true };
}
```

```
export function toRow(p, meta = {}) {
  const storage = Array.isArray(p.storage)
    ? p.storage.filter((s) => STORAGE.includes(s)).join(",")
    : "";
  return {
    segment: SEGMENTS.includes(p.segment) ? p.segment : null,
    name: str(p.name, 200),
    email: str(p.email, 320),
    city: str(p.city, 200),
    org_name: str(p.org_name, 200),
    courts: COURTS.includes(p.courts) ? p.courts : null,
    has_gpu: GPU.includes(p.has_gpu) ? p.has_gpu : null,
    storage,
    message: str(p.message, 4000),
    consent: p.consent ? 1 : 0,
    country: str(meta.country, 8),
    user_agent: str(meta.userAgent, 400),
  };
}

function json(obj, status = 200) {
  return new Response(JSON.stringify(obj), { status, headers: { "content-type":
"application/json" } });
}

export async function handleRequest(request, env) {
  let body;
  try {
    body = await request.json();
  } catch {
    return json({ ok: false, error: "invalid json" }, 400);
  }
  if (isBot(body)) return json({ ok: true }); // silently drop bots
  const v = validate(body);
  if (!v.ok) return json({ ok: false, error: v.error }, 400);

  const r = toRow(body, {
    country: request.headers.get("cf-ipcountry"),
    userAgent: request.headers.get("user-agent"),
  });

  await env.DB.prepare(
    `INSERT INTO access_requests
       (segment, name, email, city, org_name, courts, has_gpu, storage, message, consent,
country, user_agent)
     VALUES (?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?)`
  )
    .bind(r.segment, r.name, r.email, r.city, r.org_name, r.courts, r.has_gpu, r.storage,
r.message, r.consent, r.country, r.user_agent)
    .run();

  return json({ ok: true });
}

export default {
  async fetch(request, env) {
    const url = new URL(request.url);
    if (url.pathname === "/api/request-access") {
      if (request.method !== "POST") return json({ ok: false, error: "method not allowed" },
405);
      return handleRequest(request, env);
    }
    return env.ASSETS.fetch(request); // everything else: static assets
  },
};
```

398. user (2026-06-20T05:55:40.737Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\src\index.js (file state is
current in your context ? no need to Read it back)
```

399. assistant (2026-06-20T05:55:58.305Z)

```
Now the tests (the verification that matters here ? the preview can't exercise a Worker).
```

400. assistant (2026-06-20T05:56:12.601Z)

```
// Unit tests for the Request-Access Worker logic. Run: node --test src/index.test.js
import { test } from "node:test";
import assert from "node:assert/strict";
import worker, { validate, toRow, isBot, handleRequest } from "./index.js";

function fakeDB() {
  const inserts = [];
  return {
    inserts,
    prepare(sql) {
      return {
        bind(...args) {
          return { async run() { inserts.push({ sql, args }); return { success: true }; } };
        },
      };
    },
  };
}

function req(payload, headers = {}) {
  return new Request("https://khelsutra.guru/api/request-access", {
    method: "POST",
    headers: { "content-type": "application/json", ...headers },
    body: JSON.stringify(payload),
  });
}

test("validate requires a well-formed email", () => {
  assert.equal(validate({}).ok, false);
  assert.equal(validate({ email: "nope" }).ok, false);
  assert.equal(validate({ email: "a@b.co" }).ok, true);
});

test("isBot detects a filled honeypot only", () => {
  assert.equal(isBot({}), false);
  assert.equal(isBot({ company: "" }), false);
  assert.equal(isBot({ company: "spam" }), true);
});

test("toRow maps + sanitizes a valid payload", () => {
  const r = toRow(
    { segment: "academy", email: "a@b.co", storage: ["gdrive", "bucket", "evil"], has_gpu:
"yes", courts: "2-4", consent: true },
    { country: "IN", userAgent: "x" }
  );
  assert.equal(r.segment, "academy");
  assert.equal(r.storage, "gdrive,bucket"); // unknown "evil" dropped
  assert.equal(r.has_gpu, "yes");
  assert.equal(r.courts, "2-4");
  assert.equal(r.consent, 1);
  assert.equal(r.country, "IN");
});

test("toRow rejects invalid enum values", () => {
```

```
  const r = toRow({ segment: "hacker", has_gpu: "maybe", courts: "9", consent: 0 });
  assert.equal(r.segment, null);
  assert.equal(r.has_gpu, null);
  assert.equal(r.courts, null);
  assert.equal(r.consent, 0);
  assert.equal(r.storage, "");
});

test("handleRequest inserts a valid submission", async () => {
  const db = fakeDB();
  const res = await handleRequest(req({ email: "a@b.co", segment: "individual", storage:
["gdrive"] }), { DB: db });
  assert.equal(res.status, 200);
  assert.deepEqual(await res.json(), { ok: true });
  assert.equal(db.inserts.length, 1);
  assert.equal(db.inserts[0].args[2], "a@b.co"); // email is the 3rd bound column
});

test("handleRequest drops a bot without inserting", async () => {
  const db = fakeDB();
  const res = await handleRequest(req({ email: "a@b.co", company: "spam" }), { DB: db });
  assert.equal(res.status, 200);
  assert.equal(db.inserts.length, 0);
});

test("handleRequest rejects a bad email with 400", async () => {
  const db = fakeDB();
  const res = await handleRequest(req({ email: "nope" }), { DB: db });
  assert.equal(res.status, 400);
  assert.equal(db.inserts.length, 0);
});

test("handleRequest rejects invalid JSON with 400", async () => {
  const db = fakeDB();
  const bad = new Request("https://khelsutra.guru/api/request-access", {
    method: "POST", headers: { "content-type": "application/json" }, body: "{not json",
  });
  const res = await handleRequest(bad, { DB: db });
  assert.equal(res.status, 400);
  assert.equal(db.inserts.length, 0);
});

test("fetch routes: POST /api hits the handler, other paths delegate to ASSETS", async () => {
  const db = fakeDB();
  const apiRes = await worker.fetch(req({ email: "a@b.co" }), { DB: db });
  assert.equal(apiRes.status, 200);

  let assetCalled = false;
  const env = { DB: db, ASSETS: { fetch: async () => { assetCalled = true; return new
Response("ok"); } } };
  const assetRes = await worker.fetch(new Request("https://khelsutra.guru/"), env);
  assert.equal(assetCalled, true);
  assert.equal(await assetRes.text(), "ok");
});

test("fetch rejects non-POST to the API with 405", async () => {
  const res = await worker.fetch(new Request("https://khelsutra.guru/api/request-access"), { DB:
fakeDB() });
  assert.equal(res.status, 405);
});
```

401. user (2026-06-20T05:56:12.632Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\src\index.test.js (file
state is current in your context ? no need to Read it back)
```

402. assistant (2026-06-20T05:56:16.993Z)

```
Now the config, schema, and `package.json`.
```

403. assistant (2026-06-20T05:56:18.274Z)

```
{
  "name": "khelsutra-site",
  "private": true,
  "type": "module",
  "scripts": {
    "test": "node --test src/",
    "deploy": "wrangler deploy",
    "dev": "wrangler dev",
    "og": "python make_og.py"
  }
}
```

404. user (2026-06-20T05:56:18.292Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\package.json (file state is
current in your context ? no need to Read it back)
```

405. assistant (2026-06-20T05:56:25.730Z)

Cloudflare Worker for khelsutra.guru ? static site + Request-Access form API.

#

One-time setup (owner, our Cloudflare account):

npx wrangler d1 create khelsutra-access # copy the printed database_id below

npx wrangler d1 execute khelsutra-access --remote --file=schema.sql

npx wrangler deploy # or connect this repo via Workers Builds (git)

```
#
name = "khelsutra-site"
main = "src/index.js"
compatibility_date = "2025-06-01"

[assets]
directory = "./public"
binding = "ASSETS"

[[d1_databases]]
binding = "DB"
database_name = "khelsutra-access"
database_id = "REPLACE_WITH_D1_DATABASE_ID"   # from `wrangler d1 create khelsutra-access`
```

406. user (2026-06-20T05:56:25.758Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\wrangler.toml (file state
is current in your context ? no need to Read it back)
```

407. assistant (2026-06-20T05:56:32.957Z)

```
-- D1 schema for Request-Access submissions.
-- Apply: npx wrangler d1 execute khelsutra-access --remote --file=schema.sql
CREATE TABLE IF NOT EXISTS access_requests (
  id         INTEGER PRIMARY KEY AUTOINCREMENT,
  created_at TEXT NOT NULL DEFAULT (datetime('now')),
  segment    TEXT,            -- 'academy' | 'individual'
  name       TEXT,
  email      TEXT NOT NULL,
```

```
  city       TEXT,
  org_name   TEXT,            -- academy / club name
  courts     TEXT,            -- '1' | '2-4' | '5+'
  has_gpu    TEXT,            -- 'yes' | 'unsure' | 'no' | 'prefer-cloud'
  storage    TEXT,            -- csv of gdrive,onedrive,dropbox,bucket,local,unsure
  message    TEXT,            -- free text
  consent    INTEGER DEFAULT 0,
  country    TEXT,            -- coarse, from cf-ipcountry (no raw IP stored)
  user_agent TEXT
);
```

408. user (2026-06-20T05:56:32.989Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\schema.sql (file state is
current in your context ? no need to Read it back)
```

409. assistant (2026-06-20T05:56:45.751Z)

```
Now the key step ? running the Worker tests (this is the real verification, since the preview
can't exercise a Worker).
```

410. user (2026-06-20T05:56:46.512Z)

```
? validate requires a well-formed email (0.9817ms)
? isBot detects a filled honeypot only (1.1112ms)
? toRow maps + sanitizes a valid payload (0.2514ms)
? toRow rejects invalid enum values (0.1955ms)
? handleRequest inserts a valid submission (24.1246ms)
? handleRequest drops a bot without inserting (0.5218ms)
? handleRequest rejects a bad email with 400 (0.4594ms)
? handleRequest rejects invalid JSON with 400 (0.4835ms)
? fetch routes: POST /api hits the handler, other paths delegate to ASSETS (1.0593ms)
? fetch rejects non-POST to the API with 405 (0.329ms)
? tests 10
? suites 0
? pass 10
? fail 0
? cancelled 0
? skipped 0
? todo 0
? duration_ms 121.7249
```

411. assistant (2026-06-20T05:57:07.445Z)

```
**All 10 tests pass.** ? The Worker logic is verified (validation, honeypot, payload
sanitization, enum rejection, D1 insert, routing, 400/405 paths). Now updating the supporting
files (the moved structure changed the deploy story + OG output path) before I commit.
```

412. user (2026-06-20T05:57:08.392Z)

```
1       """Generate og.png (1200x630) ? the social-share card for khelsutra.guru.
2       Re-run after copy/brand tweaks:  python site/make_og.py
3       """
4       from PIL import Image, ImageDraw, ImageFont
5
```

413. user (2026-06-20T05:57:08.596Z)

```
1       # site/ ? public marketing landing page
2
3       The static one-page site served at **[khelsutra.guru](https://khelsutra.guru)** (plus
the
4       `/capture` one-page recording checklist). Plain HTML/CSS, one web font, no build step.
5
```

```
6       ## Files
7       - `index.html` ? the landing page.
8       - `capture.html` ? the printable "1-page capture checklist" (linked from the footer;
served at `/capture`).
9       - `og.png` ? 1200×630 social-share image (referenced by the OG / Twitter meta tags).
10      - `make_og.py` ? regenerates `og.png` (Pillow + Windows fonts). Re-run after brand/copy
tweaks: `python site/make_og.py`.
11
12      ## Deploy (current)
13      Deployed to **Cloudflare** via **direct upload** of this folder (Workers static assets);
14      custom domain `khelsutra.guru` + `www`, free SSL, and email routing for
`hello@khelsutra.guru`.
15
16      > Git-based auto-deploy is **not connected yet** ? pushing to this repo does **not**
redeploy.
17      > To publish changes, re-upload this folder in the Cloudflare dashboard.
18
19      ## Notes
20      - The `hello@khelsutra.guru` "Request access" CTA relies on Cloudflare Email Routing.
21      - The Cloudflare Web Analytics beacon in `index.html` has a placeholder token
22        (`REPLACE_WITH_CF_TOKEN`) ? paste the real token, or enable Automatic Setup in
Cloudflare
23        and delete the snippet.
24      - Minor cosmetic follow-up: the footer link points to `/capture.html` (Cloudflare
307-redirects
25        it to `/capture`); can be changed to `/capture` directly on the next content edit +
re-upload.
26
```

414. assistant (2026-06-20T05:58:01.589Z)

```
Tests green. Now syncing the docs to the new Worker structure (and fixing the OG output path).
```

415. user (2026-06-20T05:58:03.042Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\make_og.py has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

416. user (2026-06-20T05:58:04.912Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\make_og.py has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

417. assistant (2026-06-20T05:58:10.668Z)

site/ ? public marketing site + Request-Access API (Cloudflare Worker)

The public site at **khelsutra.guru** ? a static landing page + `/capture` checklist ? plus a small Worker that backs the Request-Access form with **Cloudflare D1**.

Layout

- `public/` ? static assets served as-is (`index.html`, `capture.html`, `og.png`).
- `src/index.js` ? the Worker: serves `public/` and handles `POST /api/request-access` ? D1.
- `src/index.test.js` ? unit tests for the form logic.
- `wrangler.toml` ? Worker + static-assets + D1 binding config.
- `schema.sql` ? the `access_requests` D1 table.
- `make_og.py` ? regenerates `public/og.png` (Pillow + Windows fonts): `python make_og.py`.

Develop / test

```
npm test       # node --test src/  (no install needed)
npm run dev    # npx wrangler dev ? local Worker + D1
```

```
```

Deploy (owner ? our Cloudflare account)

One-time:

```
npx wrangler d1 create khelsutra-access                          # paste database_id into
wrangler.toml
npx wrangler d1 execute khelsutra-access --remote --file=schema.sql
```

Then either:
- **Manual:** `npx wrangler deploy`, or
- **Git (recommended):** connect this repo in Cloudflare ? Workers Builds (root dir `site/`), bind the
  D1 database ? every merge to `master` deploys; PRs get preview URLs.

Notes

- The form posts to `/api/request-access`; if the Worker is unreachable it falls back to a prefilled
  `mailto:hello@`, so it degrades gracefully.
- `hello@khelsutra.guru` relies on Cloudflare Email Routing.
- The Web Analytics beacon in `public/index.html` has a placeholder token (`REPLACE_WITH_CF_TOKEN`)
  ? paste the real token or enable Cloudflare automatic setup.

418. user (2026-06-20T05:58:10.699Z)

The file C:\Users\avidu\Projects\khelsutra\site\README.md has been updated successfully. (file state is current in your context ? no need to Read it back)

419. assistant (2026-06-20T05:58:15.697Z)

Now the tracker + NEXT_STEPS status.

420. user (2026-06-20T05:58:19.846Z)

The file C:\Users\avidu\Projects\khelsutra\docs\REQUEST_ACCESS_FORM.md has been updated successfully. (file state is current in your context ? no need to Read it back)

421. user (2026-06-20T05:58:29.880Z)

The file C:\Users\avidu\Projects\khelsutra\NEXT_STEPS.md has been updated successfully. (file state is current in your context ? no need to Read it back)

422. user (2026-06-20T05:58:32.299Z)

The file C:\Users\avidu\Projects\khelsutra\NEXT_STEPS.md has been updated successfully. (file state is current in your context ? no need to Read it back)

423. assistant (2026-06-20T05:59:01.936Z)

Re-running the suite, then staging explicit paths for review before commit.

424. user (2026-06-20T05:59:06.736Z)

```
=== re-run tests (pre-merge gate) ===
  ? tests 10
  ? pass 10
  ? fail 0
  EXIT: GREEN
=== stage explicit paths ===
warning: in the working copy of 'site/README.md', LF will be replaced by CRLF the next time Git
touches it
warning: in the working copy of 'site/package.json', LF will be replaced by CRLF the next time
```

```
Git touches it
warning: in the working copy of 'site/schema.sql', LF will be replaced by CRLF the next time Git
touches it
warning: in the working copy of 'site/src/index.js', LF will be replaced by CRLF the next time
Git touches it
warning: in the working copy of 'site/src/index.test.js', LF will be replaced by CRLF the next
time Git touches it
warning: in the working copy of 'site/wrangler.toml', LF will be replaced by CRLF the next time
Git touches it
branch: feat/access-worker-d1
=== staged ===
  M     NEXT_STEPS.md
  M     docs/REQUEST_ACCESS_FORM.md
  M     site/README.md
  M     site/make_og.py
  A     site/package.json
  R100  site/capture.html       site/public/capture.html
  R100  site/index.html site/public/index.html
  R100  site/og.png     site/public/og.png
  A     site/schema.sql
  A     site/src/index.js
  A     site/src/index.test.js
  A     site/wrangler.toml
=== anything unstaged/untracked left? ===
```

425. assistant (2026-06-20T05:59:29.614Z)

```
**10/10 tests green**, everything staged cleanly (the moves registered as pure renames R100).
Committing, pushing, opening the PR.
```

426. user (2026-06-20T05:59:45.728Z)

```
HEAD before: 2274faa feat(site): segmented Request-Access form (academy/individual + GPU +
storage discovery) (#2)
[feat/access-worker-d1 590def8] feat(site): Request-Access Worker + D1 backend (+ 10 unit tests)
 12 files changed, 285 insertions(+), 26 deletions(-)
 create mode 100644 site/package.json
 rename site/{ => public}/capture.html (100%)
 rename site/{ => public}/index.html (100%)
 rename site/{ => public}/og.png (100%)
 create mode 100644 site/schema.sql
 create mode 100644 site/src/index.js
 create mode 100644 site/src/index.test.js
 create mode 100644 site/wrangler.toml
=== origin/master..HEAD (must be only my commit) ===
  590def8 feat(site): Request-Access Worker + D1 backend (+ 10 unit tests)
To https://github.com/avidullu/khelsutra.git
 * [new branch]      feat/access-worker-d1 -> feat/access-worker-d1
=== create PR ===
https://github.com/avidullu/khelsutra/pull/3
```

427. assistant (2026-06-20T05:59:54.790Z)

```
PR **#3** is up: https://github.com/avidullu/khelsutra/pull/3 (clean ? only my commit). Tests
green ? merging it.
```

428. user (2026-06-20T06:00:05.163Z)

```
=== verify merged ===
state=MERGED commit=f69984786
=== sync local master ===
Deleted branch feat/access-worker-d1 (was 590def8).
master now: f699847 feat(site): Request-Access Worker + D1 backend (+ 10 unit tests) (#3)
=== final tree (site/) ===
```

```
  site/README.md
  site/make_og.py
  site/package.json
  site/public/capture.html
  site/public/index.html
  site/public/og.png
  site/schema.sql
  site/src/index.js
  site/src/index.test.js
  site/wrangler.toml
```

429. assistant (2026-06-20T06:00:49.685Z)

Both PRs merged, tests green. ?

- **PR #2** (form) ? merged.
- **PR #3** (Worker + D1 backend + **10/10 unit tests**) ? merged (`f699847`).

`master` now has the whole thing: the segmented form, the Worker (`/api/request-access` ? D1), `wrangler.toml`, `schema.sql`, unit tests, and the synced docs/tracker. Let me capture this state in memory, then give you the honest "what's live vs not."

430. user (2026-06-20T06:00:57.774Z)

The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\khelsutra-domain-launch.md has been updated successfully. (file state is current in your context ? no need to Read it back)

431. assistant (2026-06-20T06:01:13.793Z)

The honest status ? what's live vs not

Everything is **in `master` and tested**, but **nothing changed on the live site yet.** khelsutra.guru still serves the original (form-less) page, because the Worker hasn't been *deployed* ? that's an owner step on your Cloudflare account.

What I verified: the form (rendering + academy-toggle, in-browser) and the Worker logic (10 unit tests: validation, honeypot, sanitization, D1 insert, routing, 400/405). **What only the deploy can prove:** the live D1 round-trip ? that's confirmed the moment you ship.

To actually go live (owner steps ? I'll co-pilot like the first launch)

This is the same session where we connect git, as you decided:
1. **Create the database:** `npx wrangler d1 create khelsutra-access` ? it prints a `database_id`.
2. **Wire it:** paste that id into `site/wrangler.toml` (it's not a secret ? safe to commit), then apply the schema: `npx wrangler d1 execute khelsutra-access --remote --file=site/schema.sql`.
3. **Connect git:** Cloudflare ? Workers & Pages ? **Create ? Workers ? Connect to Git** ? `avidullu/khelsutra`, **root directory `site/`**, bind the D1 database ? it deploys the `khelsutra-site` Worker. (From then on: merge to `master` ? deploys; PRs ? preview URLs.)
4. **Cutover the domain** (the one gotcha): khelsutra.guru is currently attached to the *old* auto-named Worker (`hidden-king-khelsutra`). Move the custom domain (`khelsutra.guru` + `www`) onto the new `khelsutra-site` Worker, then delete the old one.

After that: submitting the form writes a row to D1, and you can `wrangler d1 execute ? "SELECT * FROM access_requests"` to read responses ? including those GPU/storage answers that tell you what to build next.

Want to do this deploy + git-connect now, the same way we did the first launch (you click, I walk each step)? Or park it for later ? the repo's ready whenever you are. Either way: form + backend built, **10/10 tests green, both PRs merged.** ?

432. user (2026-06-20T06:02:10.886Z)

would this bring us into the non free tier?

433. assistant (2026-06-20T06:03:12.262Z)

Good question to ask *before* committing to the architecture. Let me confirm the current free-tier numbers rather than quote them from memory.

434. user (2026-06-20T06:03:17.885Z)

Based on the pricing documentation, the Cloudflare D1 FREE plan includes these limits:

- **Rows read**: "5 million / day"
- **Rows written**: "100,000 / day"
- **Storage**: "5 GB (total)"

These allowances are tied to the Workers Free plan and reset daily at 00:00 UTC. Once exceeded, the D1 API returns errors preventing further queries until the next day's reset.

435. user (2026-06-20T06:03:18.917Z)

Cloudflare Workers Free Plan Limits

Based on the pricing documentation, here are the specific free-tier limitations:

Requests: The free plan permits "100,000 per day" of inbound requests to your Worker.

CPU Time: You receive "10 milliseconds of CPU time per invocation" at no charge, with no duration limit otherwise.

Static Assets: Excellent news for static content?"Requests to static assets are free and unlimited," meaning they don't count against your request quota and incur no charges.

Additionally, the free plan includes other complimentary services like Workers Logs (200,000 log events daily with 3-day retention), Workers KV access, and Hyperdrive database queries (100,000 per day).

436. assistant (2026-06-20T06:03:48.603Z)

Confirmed from Cloudflare's pricing docs ? **no, this stays comfortably in the free tier**, by a wide margin. And critically: on the free plan, exceeding a limit returns errors (rate-limits you), it **does not auto-bill** ? so there's no surprise-charge risk, and no card needed.

Here's the math against your actual (invite-only alpha) usage:

Component	Free-tier limit	Your usage	Headroom
Workers ? static serving	**Free & unlimited** (doesn't count)	All landing-page traffic	?
Workers ? dynamic requests	100,000 / **day**	1 per form submit ? dozens?	~99.9% unused
D1 ? rows written	100,000 / **day**	1 per submit	microscopic
D1 ? rows read	5,000,000 / **day**	your occasional `SELECT`	microscopic
D1 ? storage	5 GB	~1 KB/row ? ~5M submissions to fill	never
Email Routing · SSL · DNS · Analytics	Free	?	?

The key insight from the docs: **serving the site is free and uncounted** ? only the *form POSTs* hit the Worker, and you'd need **100,000 submissions in a single day** to touch the limit. Git-connected builds also have a generous free monthly allotment, far above your commit cadence.

So you'd only ever cross into paid if you deliberately opt into the Workers Paid plan ($5/mo) ? which you'd do for *scale* (millions of requests) or extras (Durable Objects, higher CPU), none of which the landing page + alpha form needs.

One honest framing, so the cost picture is clear: the real money in this product lives in **video processing** ? GPU compute, the Gemini API, and video storage/egress (the stuff we cost-analyzed in the engine repo). That's a separate system. **The marketing site + lead form is, and will stay, free.** The BYO-storage/compute direction those form answers feed into is precisely the lever that keeps the *expensive* part cheap too.

437. user (2026-06-20T06:03:51.252Z)

BTW we are also planning on owning an always on VM to host the website eventually if cloudflare acts out, so design keeping in that plan

438. assistant (2026-06-20T06:07:09.178Z)

Smart to bake that in now, before the coupling spreads. Good news: the design is *already* ~80% portable (the handler uses Web-standard `Request`/`Response`, and the logic is pure functions). The one real Cloudflare coupling is the **storage call**. Let me make that an explicit swap-point and document the VM migration path ? so "if Cloudflare acts out" is a clean adapter swap, not a rewrite.

Let me branch and add the portability seam.

439. user (2026-06-20T06:07:13.061Z)

```
Switched to a new branch 'feat/portable-store'
branch: feat/portable-store  base: f699847 feat(site): Request-Access Worker + D1 backend (+ 10
unit tests) (#3)
```

440. assistant (2026-06-20T06:08:36.560Z)

Perfect ? you asked the right diligence questions, and the answers are good (stays free; portable to a VM). The portability seam is a small follow-up ? I'll do it *after* you're live (it'll actually be the first change to ride your new git auto-deploy, proving that pipeline works). Let me set that branch aside and get you deployed now.

441. user (2026-06-20T06:08:39.201Z)

```
Deleted branch feat/portable-store (was f699847).
on master @ f699847 feat(site): Request-Access Worker + D1 backend (+ 10 unit tests) (#3)
=== schema.sql (paste this into the D1 console in Step 2) ===
-- D1 schema for Request-Access submissions.
-- Apply: npx wrangler d1 execute khelsutra-access --remote --file=schema.sql
CREATE TABLE IF NOT EXISTS access_requests (
  id         INTEGER PRIMARY KEY AUTOINCREMENT,
  created_at TEXT NOT NULL DEFAULT (datetime('now')),
  segment    TEXT,            -- 'academy' | 'individual'
  name       TEXT,
  email      TEXT NOT NULL,
  city       TEXT,
  org_name   TEXT,            -- academy / club name
  courts     TEXT,            -- '1' | '2-4' | '5+'
  has_gpu    TEXT,            -- 'yes' | 'unsure' | 'no' | 'prefer-cloud'
  storage    TEXT,            -- csv of gdrive,onedrive,dropbox,bucket,local,unsure
  message    TEXT,            -- free text
  consent    INTEGER DEFAULT 0,
  country    TEXT,            -- coarse, from cf-ipcountry (no raw IP stored)
  user_agent TEXT
);
```

442. assistant (2026-06-20T06:09:00.008Z)

Reset to clean `master`. You're ready. ?

Quick context: the site is now a **Worker project** (static `public/` + the form API + D1), so this deploy goes through **git-connect**, not a folder re-upload ? which is exactly the auto-

deploy you wanted. It's all in the Cloudflare dashboard (no local CLI), ~15 min, and I'll verify at the end.

Step 1 ? Create the database *(you)*

Cloudflare dashboard (account level ? click the Cloudflare logo if you're inside the domain) ? left sidebar **Storage & Databases ? D1 SQL Database ? Create database** ? name it **exactly** `khelsutra-access` ? **Create**.

On its page, copy the **Database ID** (a UUID) and **paste it here.** I'll commit it into `wrangler.toml` so the deploy can bind it.

Step 2 ? Create the table *(you)*

Same database ? **Console** tab ? paste this and **Execute**:

```sql
CREATE TABLE IF NOT EXISTS access_requests (
  id         INTEGER PRIMARY KEY AUTOINCREMENT,
  created_at TEXT NOT NULL DEFAULT (datetime('now')),
  segment    TEXT, name TEXT, email TEXT NOT NULL, city TEXT,
  org_name   TEXT, courts TEXT, has_gpu TEXT, storage TEXT,
  message    TEXT, consent INTEGER DEFAULT 0, country TEXT, user_agent TEXT
);
```

Then I'll do **Step 3** (wire the Database ID into `wrangler.toml` + push), and walk you through:
- **Step 4 ? Connect git:** Workers & Pages ? **Create ? Workers ? Connect to Git** ? `avidullu/khelsutra`, **root directory `site/`** ? it builds + deploys the new `khelsutra-site` Worker (binding name `DB`).
- **Step 5 ? Move the domain:** attach `khelsutra.guru` (+ `www`) to the new Worker (it's currently on the old auto-named `hidden-king-khelsutra` one), then delete the old.
- **Step 6 ? I verify:** I'll fetch khelsutra.guru to confirm the form is live, then POST a test submission and confirm it returns `200` ? and you'll see the row appear in the D1 console.

Go ahead with Steps 1?2, and drop me that Database ID. I'll take it from there.

443. user (2026-06-20T06:14:28.238Z)

"This query successfully executed."

444. assistant (2026-06-20T06:14:44.228Z)

Table's created. ? Now I just need the **Database ID** to wire it in (Step 3).

You'll find it on the **`khelsutra-access` database's overview page** ? look for **"Database ID"** (a UUID like `a1b2c3d4-?`); there's usually a copy icon next to it.

Paste that here and I'll commit it into `wrangler.toml` and push, so the git deploy can bind it. Then we connect the repo.

445. user (2026-06-20T06:14:50.553Z)

cd02c6bf-4bd6-4a33-a333-ced86e36b511

446. assistant (2026-06-20T06:15:39.924Z)

Got it ? `cd02c6bf-4bd6-4a33-a333-ced86e36b511`. Wiring it into `wrangler.toml` and shipping it to `master` (so the git deploy binds the right database). Quick PR, tests still green.

447. user (2026-06-20T06:15:58.073Z)

Switched to a new branch 'chore/wire-d1-id'

```
=== wrangler.toml d1 block ===
  4:#   npx wrangler d1 create khelsutra-access          # copy the printed database_id below
  14:binding = "ASSETS"
  17:binding = "DB"
  18:database_name = "khelsutra-access"
  19:database_id = "cd02c6bf-4bd6-4a33-a333-ced86e36b511"   # from `wrangler d1 create
khelsutra-access`
TESTS: GREEN
warning: in the working copy of 'site/wrangler.toml', LF will be replaced by CRLF the next time
Git touches it
=== origin/master..HEAD ===
  d72a687 chore(site): wire D1 database_id (khelsutra-access) for the deploy binding
 * [new branch]      chore/wire-d1-id -> chore/wire-d1-id
PR: https://github.com/avidullu/khelsutra/pull/4
Deleted branch chore/wire-d1-id (was d72a687).
master now: a45e4af chore(site): wire D1 database_id (khelsutra-access) for the deploy binding
(#4)
```

448. assistant (2026-06-20T06:16:29.765Z)

Repo's now fully deploy-ready ? **PR #4 merged**, `master` has the form + Worker + schema + the bound `database_id`. Two steps left, both in the dashboard:

Step 4 ? Connect the repo (this is your git auto-deploy)

Cloudflare ? **Workers & Pages** (account level ? Cloudflare logo if you're inside the domain) ? **Create ? Workers** tab ? **Connect to Git** / **Import a repository**:
- Pick **`avidullu/khelsutra`**.
- **Root directory: `site/`** ? important (that's where `wrangler.toml` lives).
- Build command: leave default/empty (no build step). Deploy is `wrangler deploy` (default).
- **Save and Deploy.** It reads `wrangler.toml`, binds the `khelsutra-access` D1 (binding name `DB`), and deploys a Worker named **`khelsutra-site`**. (If it asks you to confirm the D1 binding, point `DB` ? `khelsutra-access`.)

You'll get a `khelsutra-site.*.workers.dev` URL when it finishes.

Step 5 ? Move the domain to the new Worker

The new **`khelsutra-site`** Worker ? **Settings ? Domains & Routes ? Add ? Custom domain** ? `khelsutra.guru` (then again for `www.khelsutra.guru`).
- If it says **"already in use,"** go to the **old `hidden-king-khelsutra`** Worker ? Settings ? Domains ? **remove** `khelsutra.guru` + `www` ? then add them on the new Worker.
- Once khelsutra.guru is on the new Worker, you can **delete** the old `hidden-king-khelsutra` Worker.

When the deploy finishes and the domain's moved, just say "deployed" ? and I'll verify end-to-end: confirm the form is live, POST a real test submission, and check it returns `200`. Then you'll see the row land in your D1 console (`SELECT * FROM access_requests;`).

Go for Step 4 ? and if any screen looks different from what I described, tell me what you see and I'll steer.

449. user (2026-06-20T06:22:23.918Z)

deployed and please verify

450. assistant (2026-06-20T06:22:59.343Z)

Verifying end-to-end now ? form live, the API round-trip, and assets.

451. user (2026-06-20T06:23:07.182Z)

=== 1) Is the FORM live on khelsutra.guru? ===

```
        1 <title>KhelSutra ? every rally indexed, the dead time gone</title>

=== 2) POST a real test submission -> /api/request-access (Worker + D1) ===

  -> HTTP 000
curl: (43) A libcurl function was given a bad argument

=== 3) GET /api/request-access (expect 405, writes nothing) ===

  -> HTTP 000
curl: (43) A libcurl function was given a bad argument

=== 4) assets ===
  /og.png -> HTTP/1.1 200 OK
  /capture -> HTTP/1.1 200 OK
```

452. assistant (2026-06-20T06:24:21.627Z)

The form check came back **empty** (the form markers aren't on the live page) and the API probe errored on a curl quirk ? so something's off. Let me diagnose cleanly: is the new build serving, and is the domain pointing at it?

453. user (2026-06-20T06:24:27.918Z)

```
######## https://khelsutra.guru ########
  form markers: NONE (old / form-less page)
  title: <title>KhelSutra ? every rally indexed, the dead time gone</title>
  GET /api/request-access -> HTTP/1.1 404 Not Found
######## https://khelsutra-site.avi-dullu.workers.dev ########
  form markers: NONE (old / form-less page)
  title:
  GET /api/request-access -> HTTP/1.1 404 Not Found
```

454. assistant (2026-06-20T06:25:20.910Z)

Honest result: **not live yet** ? and I can see exactly why. The deploy didn't fully land:

- **`khelsutra.guru` is still the OLD page** ? no form, and `GET /api/request-access` ? **404**. A 404 on `/api` means our Worker (which owns that route) isn't behind the domain. So **the domain is still pointing at the old `hidden-king-khelsutra` Worker** (Step 5 didn't take).
- **The new Worker build also looks wrong** ? its `*.workers.dev` had no title, no form, and `/api` ? 404 too. That pattern says **the build deployed from the wrong folder** ? i.e. the **root directory wasn't set to `site/`**, so Workers Builds didn't find `wrangler.toml`/`public/` and shipped the wrong thing.

So two things to fix (both quick). To pinpoint the first, tell me what you see:

A) Fix the build root ? in **Workers & Pages ? your new project ? Settings ? Builds (or "Build configuration")**:
- What is **Root directory** set to? It needs to be **`site/`**.
- Set it to `site/`, **Save**, then **Retry deployment** (Deployments tab ? ? ? Retry, or push a commit).
- Also: is the latest **build green or failed**? (If failed, the logs will say why ? paste the last few lines.)

B) Move the domain ? once the new build is green and its `*.workers.dev` URL shows the form:
- New Worker ? **Settings ? Domains & Routes ? Add ? Custom domain** ? `khelsutra.guru` + `www`.
- If "already in use," remove them from the **old `hidden-king-khelsutra`** Worker first.

Two things to paste back so I can verify/steer:
1. The **new project's name** and its **`*.workers.dev` URL** (from the Workers & Pages list).
2. Its **Root directory** setting + whether the last **build succeeded or failed**.

This is a normal first-deploy hiccup ? once the root dir is `site/` and the domain's moved, the

form + D1 will light up and I'll confirm the live round-trip immediately.

455. user (2026-06-20T06:26:28.155Z)

Build command:None
Deploy command:npx wrangler deploy
Version command:npx wrangler versions upload
Root directory:site/

456. user (2026-06-20T06:26:50.422Z)

[Request interrupted by user]

457. user (2026-06-20T06:26:50.435Z)

"Update [site/wrangler.toml](https://github.com/avidullu/khelsutra/blob/master/site/wrangler.toml) in your repo to keep settings consistent. On Wrangler v3.109.0+, we will auto-generate a PR to fix this after the build

```
# site/wrangler.toml
```

`name = 'hidden-king-khelsutra'"`

458. assistant (2026-06-20T06:27:58.356Z)

That message is the key ? Cloudflare is telling us the git-connect attached to your **existing** Worker `hidden-king-khelsutra` (the one that already owns khelsutra.guru), but our `wrangler.toml` says `name = "khelsutra-site"`. That **name mismatch** is why the deploy didn't land on the live Worker.

This is actually the good case ? it means **no domain cutover is needed**. I just need to rename in `wrangler.toml` to match the connected Worker, push, and it'll deploy straight onto the domain. Doing that now.

459. user (2026-06-20T06:28:19.062Z)

Switched to a new branch 'fix/wrangler-name'
=== name line ===
  8:name = "hidden-king-khelsutra" # the git-connected Worker that owns khelsutra.guru ? keep in sync with Cloudflare
TESTS: GREEN
warning: in the working copy of 'site/wrangler.toml', LF will be replaced by CRLF the next time Git touches it
 * [new branch]   fix/wrangler-name -> fix/wrangler-name
https://github.com/avidullu/khelsutra/pull/5
Deleted branch fix/wrangler-name (was bf4c503).
master now: a1c0d89 fix(site): match wrangler name to the git-connected Worker (hidden-king-khelsutra) (#5)

460. assistant (2026-06-20T06:28:39.359Z)

PR #5 merged ? `master` now has the matching name, which **triggers a fresh git deploy onto `hidden-king-khelsutra`** (the Worker that owns khelsutra.guru). Let me check whether it's landed yet.

461. user (2026-06-20T06:28:43.575Z)

form markers on khelsutra.guru: NONE (still old build)
GET /api/request-access -> HTTP/1.1 404 Not Found

462. assistant (2026-06-20T06:29:21.368Z)

Still the old build ? expected: the new deploy was triggered only seconds ago by the merge, and Workers Builds takes a minute or two. Rather than have you keep checking, I'll **poll khelsutra.guru in the background** and verify automatically the moment the form goes live.

463. user (2026-06-20T06:29:23.731Z)

Command running in background with ID: bbv41mw8y. Output is being written to: C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\bbv41mw8y.output. You will be notified when it completes. To check interim output, use Read on that file path.

464. assistant (2026-06-20T06:29:37.416Z)

Here's what's happening: merging PR #5 (matching the Worker name) **triggered a fresh build** on the git-connected `hidden-king-khelsutra` Worker ? the one that already owns khelsutra.guru. So when it finishes, the form + API land **directly on the live domain, no cutover needed.**

I'm polling khelsutra.guru in the background and will **auto-verify the moment the form appears** (then I'll POST a real test submission so you can watch the row land in D1).

One thing worth a glance on your end while it builds ? it'll catch a problem the poll can't: in **Workers & Pages ? `hidden-king-khelsutra` ? Deployments** (or the **Builds** tab), the newest entry should be for commit **"fix(site): match wrangler name?"** and go **Building ? Success**. If it shows **Failed**, open it and paste me the last few log lines ? the name mismatch is fixed, but if anything else trips (e.g. the D1 binding), the logs will tell us and I'll sort it.

Sit tight ? I'll report back as soon as it's live. ??

465. user (2026-06-20T06:29:42.446Z)

<task-notification>
<task-id>bbv41mw8y</task-id>
<tool-use-id>toolu_01WhTDR6KhkDzHdm4KpXnJgU</tool-use-id>
<output-file>C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\bbv41mw8y.output</output-file>
<status>completed</status>
<summary>Background command "Background-poll khelsutra.guru until the form is live" completed (exit code 0)</summary>
</task-notification>

466. assistant (2026-06-20T06:29:49.618Z)

The poll exited 0 ? **the form is live!** Let me run the full end-to-end verification now: confirm the form, POST a real submission through the Worker into D1, and check routing.

467. user (2026-06-20T06:29:59.347Z)

```
=== poll result ===
FORM IS LIVE after ~12s
=== 1) form live on khelsutra.guru ===
        1 Coming next to the alpha
        1 Get into the KhelSutra alpha
        1 id="af"
        4 name="has_gpu"
        6 name="storage"
=== 2) POST real test submission (Worker -> D1) ===
  HTTP/1.1 200 OK
  {"ok":true}
=== 3) GET /api (expect 405, no DB write) + honeypot drop check ===
  GET  -> HTTP/1.1 405 Method Not Allowed
  bad-email POST -> HTTP/1.1 400 Bad Request
```

468. assistant (2026-06-20T06:30:34.907Z)

? **It's LIVE and the whole stack works** ? verified end-to-end against the real domain:

Check	Result
Form on khelsutra.guru	? live (`id="af"`, 4 GPU options, 6 storage options, "Coming next" section)
`POST /api/request-access`	? **`HTTP 200 {"ok":true}`** ? Worker validated + **inserted a row into D1**
`GET /api/request-access`	? `405` (routing correct, writes nothing)
Bad-email `POST`	? `400` (validation works in production)

The form deployed straight onto the live domain (`hidden-king-khelsutra`) the moment PR #5 merged ? so **your git auto-deploy is now working**: every merge to `master` ships automatically. (That name-mismatch fix was the last blocker; it self-deployed in ~12s.)

Go look in your D1 console ? run:
```sql
SELECT id, created_at, segment, email, has_gpu, storage, message FROM access_requests;
```
You'll see one row from my verification (`deploy-test@khelsutra.guru`, GPU=yes, storage=gdrive,bucket) ? **safe to delete.** That's your first real lead-capture row, structured exactly as designed.

Let me record this milestone in memory.

469. user (2026-06-20T06:30:43.188Z)

The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\khelsutra-domain-launch.md has been updated successfully. (file state is current in your context ? no need to Read it back)

470. assistant (2026-06-20T06:31:01.953Z)

You're fully live, cofounder. ? khelsutra.guru now:
- serves the segmented form (academy/individual + GPU + storage + free text),
- writes structured leads into **your own Cloudflare D1**,
- and **auto-deploys from git** ? merge to `master`, it ships.

All on the free tier, all in your account.

Small things left (all optional, none blocking)

- **Always Use HTTPS + HSTS** ? Cloudflare ? SSL/TLS ? Edge Certificates (the one security-scanner flag from earlier).
- **Delete the orphan `khelsutra-site` Worker** if one got created during the earlier attempts (Workers & Pages ? if you see a `khelsutra-site` with no domain, delete it; keep `hidden-king-khelsutra`).
- Cosmetic: footer `/capture.html` ? `/capture`; paste the Web Analytics token.

The one I'd do next ? the portability seam

Remember the always-on-VM plan? The natural follow-up is the small refactor I set aside: **decouple the D1 call behind a storage adapter** + a `docs/PORTABILITY.md` migration runbook. Two nice properties:
1. It makes "if Cloudflare acts out ? move to a VM" a clean adapter swap, not a rewrite (exactly the design directive you gave).
2. It'll be the **first real feature change to ride your new git auto-deploy** ? so we prove that pipeline works on something beyond a config tweak.

Want me to build that now (one PR, tests green, then it auto-deploys), or park it and call it a session here? Either way ? that's a domain bought this morning turned into a live, lead-collecting, auto-deploying product by afternoon. ?

471. user (2026-06-20T06:33:19.750Z)

Yes, and I have an up and running Digital ocean droplet where another session is building a "CPU and Storage shm", I can give you ssh access and want you to drive this full work and testing live as well. Happy to do whateveris needed so you can unblock and go your way

472. assistant (2026-06-20T06:37:53.308Z)

I'm in ? and this is a great way to prove the portability for real, not just on paper. The plan: I'll finish the **portable host** (the same form logic running on a plain Node + SQLite server, dependency-free), verify it **locally first**, then deploy and test it **live on your droplet**.

Two things I'll be careful about, since it's a shared box with another active session: I'll **isolate** (its own directory + its own port, no system-wide clobbering) and **steer clear** of the "CPU and Storage" work ? which means I need to know where that lives before I touch anything.

Let me start building the portable host now (the part I can do + verify locally regardless of the droplet).

473. user (2026-06-20T06:37:58.441Z)

```
Switched to a new branch 'feat/portable-host'
branch: feat/portable-host  base: a1c0d89 fix(site): match wrangler name to the git-connected
Worker (hidden-king-khelsutra) (#5)
=== does Node 26 expose node:sqlite? ===
node:sqlite OK -> 42
```

474. assistant (2026-06-20T06:38:15.939Z)

`node:sqlite` works in Node 26 ? so the VM host can be **fully dependency-free** (no `npm install` on the droplet, just Node). Building the storage seam + refactor.

475. assistant (2026-06-20T06:38:21.557Z)

```
// Storage adapters for Request-Access submissions.
//
// PORTABILITY (see docs/PORTABILITY.md): the business logic (validate / toRow /
// handleRequest in index.js) is runtime-agnostic ? ALL database coupling lives here.
// The Cloudflare Worker uses d1Store(env.DB); an always-on VM uses sqliteStore(db)
// over node:sqlite (or better-sqlite3). Same schema.sql, same INSERT ? zero dialect
// change. To move hosts, swap the adapter, not the logic.

export const INSERT_SQL =
  `INSERT INTO access_requests
     (segment, name, email, city, org_name, courts, has_gpu, storage, message, consent, country,
user_agent)
   VALUES (?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?)`;

export const insertArgs = (r) => [
  r.segment, r.name, r.email, r.city, r.org_name, r.courts,
  r.has_gpu, r.storage, r.message, r.consent, r.country, r.user_agent,
];

// Cloudflare D1 (async): prepare(sql).bind(...args).run()
export function d1Store(DB) {
  return {
    async insertAccessRequest(r) {
      await DB.prepare(INSERT_SQL).bind(...insertArgs(r)).run();
    },
  };
}

// node:sqlite / better-sqlite3 (sync): prepare(sql).run(...args). For the VM host.
```

```
export function sqliteStore(db) {
  const stmt = db.prepare(INSERT_SQL);
  return {
    async insertAccessRequest(r) {
      stmt.run(...insertArgs(r));
    },
  };
}
```

476. user (2026-06-20T06:38:21.647Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\src\store.js (file state is
current in your context ? no need to Read it back)
```

477. assistant (2026-06-20T06:38:27.581Z)

```
Now refactoring `index.js` to use the adapter (the Worker logic becomes runtime-agnostic).
```

478. user (2026-06-20T06:38:29.068Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

479. user (2026-06-20T06:38:30.745Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

480. user (2026-06-20T06:38:32.842Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

481. user (2026-06-20T06:38:33.624Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

482. assistant (2026-06-20T06:38:39.372Z)

```
Now the Node host (`server.js`) ? dependency-free, same handler logic, SQLite via `node:sqlite`.
```

483. assistant (2026-06-20T06:38:50.786Z)

```
// Always-on VM host for the KhelSutra site ? dependency-free (Node 22+ only).
// Serves public/ and POST /api/request-access into SQLite (node:sqlite), using the
// SAME handler logic as the Cloudflare Worker (src/index.js). This is the portability
// fallback "if Cloudflare acts out" ? see docs/PORTABILITY.md.
//
//   PORT=8080 DB_PATH=/var/lib/khelsutra/access.db node server.js
//
import { createServer } from "node:http";
import { readFile } from "node:fs/promises";
import { join, extname, normalize } from "node:path";
import { DatabaseSync } from "node:sqlite";
import { handleRequest } from "./src/index.js";
import { sqliteStore } from "./src/store.js";

const PORT = Number(process.env.PORT) || 8080;
const ROOT = import.meta.dirname;
const PUBLIC = join(ROOT, "public");
const DB_PATH = process.env.DB_PATH || join(ROOT, "access.db");
```

```
const db = new DatabaseSync(DB_PATH);
db.exec(await readFile(join(ROOT, "schema.sql"), "utf8")); // CREATE TABLE IF NOT EXISTS ?
idempotent
const store = sqliteStore(db);

const TYPES = {
  ".html": "text/html; charset=utf-8", ".png": "image/png", ".svg": "image/svg+xml",
  ".css": "text/css", ".js": "text/javascript", ".ico": "image/x-icon", ".txt": "text/plain",
};

async function toRequest(req) {
  const url = `http://${req.headers.host || "localhost"}${req.url}`;
  let body;
  if (req.method !== "GET" && req.method !== "HEAD") {
    const chunks = [];
    for await (const c of req) chunks.push(c);
    body = Buffer.concat(chunks);
  }
  return new Request(url, { method: req.method, headers: req.headers, body });
}

async function send(res, response) {
  res.statusCode = response.status;
  for (const [k, v] of response.headers) res.setHeader(k, v);
  res.end(Buffer.from(await response.arrayBuffer()));
}

async function serveStatic(pathname, res) {
  const rel = pathname === "/" ? "index.html"
    : pathname === "/capture" ? "capture.html"
    : pathname.replace(/^\/+/, "");
  const file = normalize(join(PUBLIC, rel));
  if (!file.startsWith(PUBLIC)) { res.statusCode = 403; return res.end("forbidden"); }
  try {
    const data = await readFile(file);
    res.setHeader("content-type", TYPES[extname(file)] || "application/octet-stream");
    res.end(data);
  } catch {
    res.statusCode = 404;
    res.end("not found");
  }
}

createServer(async (req, res) => {
  try {
    const pathname = decodeURIComponent(new URL(`http://x${req.url}`).pathname);
    if (pathname === "/api/request-access") {
      if (req.method !== "POST") { res.statusCode = 405; return
res.end('{"ok":false,"error":"method not allowed"}'); }
      return await send(res, await handleRequest(await toRequest(req), store));
    }
    return await serveStatic(pathname, res);
  } catch (err) {
    res.statusCode = 500;
    res.end('{"ok":false,"error":"server error"}');
    console.error(err);
  }
}).listen(PORT, () => console.log(`khelsutra site listening on :${PORT}  (sqlite:
${DB_PATH})`));
```

484. user (2026-06-20T06:38:50.814Z)

File created successfully at: C:\Users\avidu\Projects\khelsutra\site\server.js (file state is current in your context ? no need to Read it back)

485. assistant (2026-06-20T06:38:57.621Z)

```
Now updating the tests ? switching to the store seam and adding a **real `node:sqlite` round-
trip** test (stronger than the fake DB: it proves the VM path actually persists).
```

486. user (2026-06-20T06:38:58.587Z)

```
1       // Unit tests for the Request-Access Worker logic. Run: node --test src/index.test.js
2       import { test } from "node:test";
3       import assert from "node:assert/strict";
4       import worker, { validate, toRow, isBot, handleRequest } from "./index.js";
5
6       function fakeDB() {
7         const inserts = [];
8         return {
9           inserts,
10          prepare(sql) {
11            return {
12              bind(...args) {
13                return { async run() { inserts.push({ sql, args }); return { success: true };
} };
14              },
15            };
16          },
17        };
18      }
19
20      function req(payload, headers = {}) {
21        return new Request("https://khelsutra.guru/api/request-access", {
22          method: "POST",
23          headers: { "content-type": "application/json", ...headers },
24          body: JSON.stringify(payload),
25        });
```

487. assistant (2026-06-20T06:39:23.754Z)

```
// Tests for the Request-Access logic + the storage adapters. Run: node --test src/
import { test } from "node:test";
import assert from "node:assert/strict";
import { DatabaseSync } from "node:sqlite";
import worker, { validate, toRow, isBot, handleRequest } from "./index.js";
import { d1Store, sqliteStore, insertArgs, INSERT_SQL } from "./store.js";

// A store spy (runtime-agnostic) ? what handleRequest depends on.
function fakeStore() {
  const rows = [];
  return { rows, async insertAccessRequest(r) { rows.push(r); } };
}
// A D1-shaped fake (prepare().bind().run()) ? for the Cloudflare adapter + fetch routing.
function fakeDB() {
  const inserts = [];
  return {
    inserts,
    prepare(sql) {
      return { bind(...args) { return { async run() { inserts.push({ sql, args }); } }; } };
    },
  };
}
function req(payload, headers = {}) {
  return new Request("https://khelsutra.guru/api/request-access", {
    method: "POST",
    headers: { "content-type": "application/json", ...headers },
    body: JSON.stringify(payload),
  });
}
```

```
test("validate requires a well-formed email", () => {
  assert.equal(validate({}).ok, false);
  assert.equal(validate({ email: "nope" }).ok, false);
  assert.equal(validate({ email: "a@b.co" }).ok, true);
});

test("isBot detects a filled honeypot only", () => {
  assert.equal(isBot({}), false);
  assert.equal(isBot({ company: "" }), false);
  assert.equal(isBot({ company: "spam" }), true);
});

test("toRow maps + sanitizes a valid payload", () => {
  const r = toRow(
    { segment: "academy", email: "a@b.co", storage: ["gdrive", "bucket", "evil"], has_gpu:
"yes", courts: "2-4", consent: true },
    { country: "IN", userAgent: "x" }
  );
  assert.equal(r.segment, "academy");
  assert.equal(r.storage, "gdrive,bucket"); // unknown "evil" dropped
  assert.equal(r.has_gpu, "yes");
  assert.equal(r.courts, "2-4");
  assert.equal(r.consent, 1);
  assert.equal(r.country, "IN");
});

test("toRow rejects invalid enum values", () => {
  const r = toRow({ segment: "hacker", has_gpu: "maybe", courts: "9", consent: 0 });
  assert.equal(r.segment, null);
  assert.equal(r.has_gpu, null);
  assert.equal(r.courts, null);
  assert.equal(r.consent, 0);
  assert.equal(r.storage, "");
});

test("handleRequest inserts a valid submission via the store", async () => {
  const store = fakeStore();
  const res = await handleRequest(req({ email: "a@b.co", segment: "individual", storage:
["gdrive"] }), store);
  assert.equal(res.status, 200);
  assert.deepEqual(await res.json(), { ok: true });
  assert.equal(store.rows.length, 1);
  assert.equal(store.rows[0].email, "a@b.co");
});

test("handleRequest drops a bot without inserting", async () => {
  const store = fakeStore();
  const res = await handleRequest(req({ email: "a@b.co", company: "spam" }), store);
  assert.equal(res.status, 200);
  assert.equal(store.rows.length, 0);
});

test("handleRequest rejects a bad email with 400", async () => {
  const store = fakeStore();
  const res = await handleRequest(req({ email: "nope" }), store);
  assert.equal(res.status, 400);
  assert.equal(store.rows.length, 0);
});

test("handleRequest rejects invalid JSON with 400", async () => {
  const store = fakeStore();
  const bad = new Request("https://khelsutra.guru/api/request-access", {
    method: "POST", headers: { "content-type": "application/json" }, body: "{not json",
  });
  const res = await handleRequest(bad, store);
```

```
  assert.equal(res.status, 400);
  assert.equal(store.rows.length, 0);
});

test("d1Store issues the bound INSERT (Cloudflare path)", async () => {
  const db = fakeDB();
  await d1Store(db).insertAccessRequest(toRow({ email: "a@b.co", segment: "individual" }));
  assert.equal(db.inserts.length, 1);
  assert.equal(db.inserts[0].sql, INSERT_SQL);
  assert.equal(db.inserts[0].args[2], "a@b.co"); // email is the 3rd column
});

test("sqliteStore persists a real row (node:sqlite round-trip ? the VM path)", async () => {
  const db = new DatabaseSync(":memory:");
  db.exec(`CREATE TABLE access_requests (id INTEGER PRIMARY KEY AUTOINCREMENT, created_at TEXT,
    segment TEXT, name TEXT, email TEXT, city TEXT, org_name TEXT, courts TEXT, has_gpu TEXT,
    storage TEXT, message TEXT, consent INTEGER, country TEXT, user_agent TEXT)`);
  const store = sqliteStore(db);
  await store.insertAccessRequest(toRow(
    { segment: "academy", email: "coach@club.in", storage: ["gdrive", "bucket"], has_gpu: "yes",
consent: true },
    { country: "IN" }
  ));
  const got = db.prepare("SELECT email, segment, has_gpu, storage, consent, country FROM
access_requests").get();
  assert.equal(got.email, "coach@club.in");
  assert.equal(got.segment, "academy");
  assert.equal(got.has_gpu, "yes");
  assert.equal(got.storage, "gdrive,bucket");
  assert.equal(got.consent, 1);
  assert.equal(got.country, "IN");
});

test("fetch routes: POST /api hits the handler, other paths delegate to ASSETS", async () => {
  const apiRes = await worker.fetch(req({ email: "a@b.co" }), { DB: fakeDB() });
  assert.equal(apiRes.status, 200);

  let assetCalled = false;
  const env = { DB: fakeDB(), ASSETS: { fetch: async () => { assetCalled = true; return new
Response("ok"); } } };
  const assetRes = await worker.fetch(new Request("https://khelsutra.guru/"), env);
  assert.equal(assetCalled, true);
  assert.equal(await assetRes.text(), "ok");
});

test("fetch rejects non-POST to the API with 405", async () => {
  const res = await worker.fetch(new Request("https://khelsutra.guru/api/request-access"), { DB:
fakeDB() });
  assert.equal(res.status, 405);
});
```

488. user (2026-06-20T06:39:23.828Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.test.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

489. assistant (2026-06-20T06:40:11.443Z)

```
Tests + the real VM-path proof. First the unit suite (now includes a real `node:sqlite` round-
trip).
```

490. user (2026-06-20T06:40:13.660Z)

```
  ? validate requires a well-formed email (0.9845ms)
```

```
  ? isBot detects a filled honeypot only (0.1355ms)
  ? toRow maps + sanitizes a valid payload (0.2211ms)
  ? toRow rejects invalid enum values (0.1329ms)
  ? handleRequest inserts a valid submission via the store (26.4361ms)
  ? handleRequest drops a bot without inserting (0.644ms)
  ? handleRequest rejects a bad email with 400 (0.7525ms)
  ? handleRequest rejects invalid JSON with 400 (0.6285ms)
  ? d1Store issues the bound INSERT (Cloudflare path) (0.3313ms)
  ? sqliteStore persists a real row (node:sqlite round-trip ? the VM path) (1.0888ms)
  ? fetch routes: POST /api hits the handler, other paths delegate to ASSETS (0.9883ms)
  ? fetch rejects non-POST to the API with 405 (0.2292ms)
  ? tests 12
  ? pass 12
  ? fail 0
```

491. assistant (2026-06-20T06:40:48.779Z)

```
**12/12 green** ? including the real `node:sqlite` round-trip (the VM path persists for real).
Now the end-to-end proof: I'll actually **run `server.js` locally** and exercise the full HTTP
round-trip (form served + POST ? SQLite) before it ever touches the droplet.
```

492. user (2026-06-20T06:40:56.599Z)

```
1) GET / ? form served by the Node host?
   Get into the KhelSutra alpha
   id="af"
   name="has_gpu"
2) GET /capture ->
   HTTP/1.1 200 OK
3) GET /og.png ->
   HTTP/1.1 200 OK
4) POST /api/request-access -> {"ok":true}
5) GET /api (expect 405) ->
   HTTP/1.1 405 Method Not Allowed
6) row persisted to SQLite?
    [{"email":"vm-test@khelsutra.guru","segment":"academy","org_name":"Test
Club","courts":"2-4","has_gpu":"yes","storage":"gdrive,local","consent":1}]
server log:
   khelsutra site listening on :8123  (sqlite:
C:/Users/avidu/AppData/Local/Temp/ks_smoke_1097.db)
```

493. assistant (2026-06-20T06:41:23.801Z)

```
**The portable host works end-to-end, locally** ? `server.js` (Node + `node:sqlite`, zero deps)
served the form, took the POST, and **persisted the exact row to SQLite**:
```
{"email":"vm-test@khelsutra.guru","segment":"academy","org_name":"Test
Club","courts":"2-4","has_gpu":"yes","storage":"gdrive,local","consent":1}
```
Same handler logic as the Cloudflare Worker ? just a different adapter. That's portability
*proven*, not asserted. Let me document it and open the PR.
```

494. assistant (2026-06-20T06:41:38.250Z)

Portability ? Cloudflare now, always-on VM if needed

```
> **Status:** design + proven (2026-06-20). **Decision:** keep the site/form host-portable so
> moving off Cloudflare to an owner-run always-on VM ("if Cloudflare acts out") is an **adapter
> swap, not a rewrite.**
```

The shape that makes it portable

```
- **Logic is runtime-agnostic.** `src/index.js` uses Web-standard `Request`/`Response` and pure
  helpers (`validate`, `toRow`, `isBot`, `handleRequest`) ? no host-specific calls.
```

- **All DB coupling is in adapters** (`src/store.js`): `d1Store` (Cloudflare D1) and `sqliteStore`
  (node:sqlite / better-sqlite3). `handleRequest(request, store)` takes the store by injection.
- **One schema, one SQL.** `schema.sql` and the `INSERT` are plain SQLite, so D1 and a VM's SQLite
  file are byte-for-byte compatible ? zero dialect change. (Postgres later = minor placeholder tweaks.)

Two hosts, same code

	Cloudflare (now)	Always-on VM (fallback)
Entry	`src/index.js` (Worker `fetch`)	`server.js` (`node:http`)
Static	`env.ASSETS`	`server.js` serves `public/`
Storage	`d1Store(env.DB)` ? D1	`sqliteStore(db)` ? `node:sqlite` file
Country	`cf-ipcountry` header	optional (GeoIP / null)
Deploy	git ? Workers Builds	`node server.js` under systemd
TLS/DNS	Cloudflare	keep Cloudflare in front, **or** Caddy/Let's Encrypt

Proven, not asserted: `node --test src/` is 12 green incl. a real `node:sqlite` round-trip, and
`server.js` was smoke-tested locally ? it served the form and persisted a real submission to SQLite.

VM deploy runbook ("if Cloudflare acts out")

Node 22+ only ? **no `npm install`** (node:sqlite is built in).
```
# on the VM:
git clone ? && cd khelsutra/site        # or scp the site/ folder
PORT=8080 DB_PATH=/var/lib/khelsutra/access.db node server.js
# put it under systemd (auto-restart) + a reverse proxy (Caddy/nginx) for TLS, then point DNS at the VM. ~1 hour; logic + schema unchanged.
```

Honest notes

- "Cloudflare acts out" usually means **Workers/D1 pricing or limits, not DNS** ? so the cheapest
  move is to keep Cloudflare DNS/CDN in front and shift only the **compute** to the VM (point a
  route/origin at the box). Full exit (own DNS + Caddy TLS) is also supported.
- You can run **both** (Cloudflare primary + VM warm standby) since the schema is identical; a
  periodic `access_requests` export keeps them in sync.
- `server.js` is intentionally dependency-free; if you later want Postgres or a framework (Hono runs
  on Workers *and* Node), it slots in behind the same `store` + handler.

495. user (2026-06-20T06:41:38.306Z)

File created successfully at: C:\Users\avidu\Projects\khelsutra\docs\PORTABILITY.md (file state is current in your context ? no need to Read it back)

496. assistant (2026-06-20T06:41:59.546Z)

All proven. Committing the portable-host PR (12 tests green, smoke-tested) and merging it ? which also **auto-deploys to Cloudflare**, proving your pipeline on a real code change. I'll re-verify the live form right after.

497. user (2026-06-20T06:42:19.216Z)

warning: in the working copy of 'site/src/index.test.js', LF will be replaced by CRLF the next time Git touches it
warning: in the working copy of 'docs/PORTABILITY.md', LF will be replaced by CRLF the next time Git touches it

```
warning: in the working copy of 'site/server.js', LF will be replaced by CRLF the next time Git
touches it
warning: in the working copy of 'site/src/store.js', LF will be replaced by CRLF the next time
Git touches it
=== staged ===
  A     docs/PORTABILITY.md
  A     site/server.js
  M     site/src/index.js
  M     site/src/index.test.js
  A     site/src/store.js
=== origin/master..HEAD ===
  92077ec feat(site): portable host ? runtime-agnostic store adapter + Node/SQLite server
 * [new branch]      feat/portable-host -> feat/portable-host
https://github.com/avidullu/khelsutra/pull/6
Deleted branch feat/portable-host (was 92077ec).
master now: 8c607eb feat(site): portable host ? runtime-agnostic store adapter + Node/SQLite
server (#6)
```

498. assistant (2026-06-20T06:43:11.213Z)

PR #6 merged ? `master` at `8c607eb`, which **auto-deploys the refactor to Cloudflare** (behavior-identical, 12 tests green). Quick health check that the live site's still up:

499. user (2026-06-20T06:43:14.928Z)

```
live khelsutra.guru form present: 1 (1 = yes)
live GET /api -> HTTP/1.1 405 Method Not Allowed
```

500. assistant (2026-06-20T06:43:42.858Z)

? Live Cloudflare site is healthy post-merge (form present, `/api` ? 405). And the **portable host is done and proven** ? `server.js` ran locally, served the form, and persisted a real submission to SQLite, with 12 tests green. PR #6 merged.

Now I'm ready to **deploy + test it live on your droplet.** Since it's a shared box with another active session, I'll work **isolated** (its own folder + a free port, no `sudo`, nothing system-wide) and stay clear of the "CPU and Storage" work. To do that safely I need three things from you:

1. SSH access ? how do you connect? e.g. `ssh root@<droplet-ip>` or an alias from your `~/.ssh/config`. I'll run `ssh` **from this machine using your existing key** ? **don't paste a private key here.** (If this machine can't reach the droplet yet, tell me and we'll sort the key first.)

2. Coordination ? what's the other session's **"CPU and Storage"** work, and **where does it live** (directory, any ports/services it's using)? I'll avoid those paths and ports entirely.

3. Where to run it ? a directory I can own (e.g. `~/khelsutra-site`) and a **free port** (I'll default to something like `8787` unless you say otherwise). Do I have **sudo**, or should I stay non-root?

My proposed first step (zero system changes):
- Clone the repo (or I `scp` the `site/` folder) to my isolated dir on the droplet.
- Run `PORT=8787 DB_PATH=./access.db node server.js` (Node 22+ only ? no installs).
- `curl` it from the droplet: GET the form, POST a test submission, confirm the row lands in SQLite ? i.e. the **same end-to-end proof I just did locally, but live on your VM.**

Then, once that round-trips, the optional **productionizing** (systemd auto-restart + a reverse proxy for TLS + a subdomain like `vm.khelsutra.guru`) is a clean follow-up ? but I'd want your nod before anything `sudo` or public-facing.

Drop me those three and I'll drive it. ?

501. user (2026-06-20T06:46:58.101Z)

can you not give an ssh key which I can add to droplet and you can ssh into it and drive all this?

502. assistant (2026-06-20T06:47:52.621Z)

Yes ? that's the clean way. I'll generate a **dedicated keypair on this machine**: the **private** key stays here (never leaves your machine ? `ssh` just uses it locally via my Bash tool), and you add the **public** key to the droplet. It's scoped to this one droplet and fully revocable (delete one line from `authorized_keys` when we're done).

503. user (2026-06-20T06:47:54.927Z)

Generated new keypair: /c/Users/avidu/.ssh/khelsutra_droplet

================= PUBLIC KEY ? add this to the droplet ==================
ssh-ed25519 AAAAC3NzaC1lZDI1NTE5AAAAIBp3aQSCMK9yTgayzTQRplgEmtLsZs9BjKPLlEV5GI9F claude-code-khelsutra-droplet
==

504. assistant (2026-06-20T06:48:18.288Z)

Done ? keypair generated, **private key stays on this machine**. Here's the public key to authorize on the droplet:

```
ssh-ed25519 AAAAC3NzaC1lZDI1NTE5AAAAIBp3aQSCMK9yTgayzTQRplgEmtLsZs9BjKPLlEV5GI9F claude-code-khelsutra-droplet
```

Add it to the droplet ? easiest is to paste this on the droplet (in the DigitalOcean web console, or however you already get in), **as the user I'll connect as**:
```bash
mkdir -p ~/.ssh && chmod 700 ~/.ssh
echo "ssh-ed25519 AAAAC3NzaC1lZDI1NTE5AAAAIBp3aQSCMK9yTgayzTQRplgEmtLsZs9BjKPLlEV5GI9F claude-code-khelsutra-droplet" >> ~/.ssh/authorized_keys
chmod 600 ~/.ssh/authorized_keys
```

Then give me **three quick things** and I'll take it from there:
1. **Droplet IP** (or hostname).
2. **SSH user** I should connect as ? `root`, or a non-root user like `ubuntu`/`deploy`? (The key goes in *that* user's `authorized_keys`, so add it as that user.)
3. **The other session's work** ? what's the "CPU and Storage" thing, and any **directories/ports it's using** so I steer clear. (I'll also inspect read-only once I'm in ? `ss -tlnp`, `ls` ? before touching anything.)

The moment you confirm, I'll test the connection (`ssh ? "echo connected; uname -a"`), then drive the full deploy + live round-trip test on the droplet ? isolated dir, free port, no `sudo` unless you green-light it.

Security note: this key only grants what you just authorized; to revoke me later, delete that one line from `~/.ssh/authorized_keys` on the droplet. ?

505. user (2026-06-20T06:59:31.327Z)

168.144.126.176 root@ you can log and verify

506. assistant (2026-06-20T07:00:11.168Z)

On it ? connecting and doing a **read-only recon first** (so I can see what the other session is running and pick an isolated dir + free port before touching anything).

507. user (2026-06-20T07:00:17.969Z)

```
Warning: Permanently added '168.144.126.176' (ED25519) to the list of known hosts.
=== CONNECTED ===
root
claude-do-rally-test
Linux 6.8.0-71-generic x86_64
PRETTY_NAME="Ubuntu 24.04.3 LTS"
=== node / npm ===
NO node
no npm
=== listening TCP ports ===
State  Recv-Q Send-Q Local Address:Port Peer Address:PortProcess
LISTEN 0      4096   127.0.0.53%lo:53        0.0.0.0:*    users:(("systemd-
resolve",pid=710,fd=15))
LISTEN 0      4096      127.0.0.54:53        0.0.0.0:*    users:(("systemd-
resolve",pid=710,fd=17))
LISTEN 0      4096       127.0.0.1:9000      0.0.0.0:*    users:(("minio",pid=13366,fd=6))
LISTEN 0      5          127.0.0.1:9023      0.0.0.0:*    users:(("gcp-storage-
emu",pid=14945,fd=3))
LISTEN 0      4096         0.0.0.0:22        0.0.0.0:*
users:(("sshd",pid=8337,fd=3),("systemd",pid=1,fd=132))
LISTEN 0      4096               *:9000            *:*    users:(("minio",pid=13366,fd=7))
LISTEN 0      4096               *:9001            *:*    users:(("minio",pid=13366,fd=9))
LISTEN 0      4096            [::]:22           [::]:*
users:(("sshd",pid=8337,fd=4),("systemd",pid=1,fd=133))
LISTEN 0      4096           [::1]:9000         [::]:*    users:(("minio",pid=13366,fd=8))
=== top processes (by mem) ===
USER         PID %CPU %MEM COMMAND
root       13366  0.1  3.4 minio
root       14945  0.1  0.7 gcp-storage-emu
root       12102  0.0  0.6 multipathd
root         994  0.0  0.5 unattended-upgr
do-agent    7697  0.0  0.4 do-agent
root        8241  0.0  0.3 droplet-agent
root           1  0.1  0.3 systemd
root       12329  0.0  0.3 udisksd
systemd+     710  0.0  0.3 systemd-resolve
root       12353  0.0  0.3 ModemManager
root        8319  0.0  0.3 systemd-journal
=== other logins / sessions ===
root     pts/0        2026-06-20 06:34 (162.243.190.66)
tmux:
none
screen:
No Sockets found in /run/screen/S-root.

=== dirs: /root /opt /srv /home ===
total 2552
drwx------  9 root root    4096 Jun 20 06:57 .
drwxr-xr-x 22 root root    4096 Jun 20 05:56 ..
drwxr-xr-x  2 root root    4096 Jun 20 06:28 .aws
-rw-------  1 root root      45 Jun 20 06:23 .bash_history
-rw-r--r--  1 root root    3106 Apr 22  2024 .bashrc
drwx------  4 root root    4096 Jun 20 06:04 .cache
-rw-r--r--  1 root root       0 Jun 20 05:56 .cloud-locale-test.skip
drwxr-xr-x  2 root root    4096 Jun 20 06:36 .keys
drwx------  3 root root    4096 Jun 20 06:14 .minio
-rw-r--r--  1 root root     161 Apr 22  2024 .profile
drwx------  2 root root    4096 Jun 20 06:38 .ssh
-rw-------  1 root root     794 Jun 20 06:36 .viminfo
-rw-r--r--  1 root root     241 Jun 20 06:17 .wget-hsts
-rw-r--r--  1 root root     366 Jun 20 06:54 do_gcs_config.json
-rw-r--r--  1 root root     833 Jun 20 06:52 do_gcs_setup.sh
-rw-r--r--  1 root root     359 Jun 20 06:54 do_gcs_upload.py
-rw-r--r--  1 root root    1586 Jun 20 06:52 do_gcs_validate.py
-rw-r--r--  1 root root    1002 Jun 20 06:54 do_gcs_worker.sh
```

```
-rw-r--r--  1 root root     664 Jun 20 06:54 do_gcs_wverify.py
--- /opt ---
total 12
drwxr-xr-x  3 root root    4096 Jun 20 05:56 .
drwxr-xr-x 22 root root    4096 Jun 20 05:56 ..
drwxr-xr-x  5 root nogroup 4096 Jun 20 05:56 digitalocean
--- /srv ---
total 8
drwxr-xr-x  2 root root 4096 Aug  5  2025 .
drwxr-xr-x 22 root root 4096 Jun 20 05:56 ..
--- /home ---
total 8
drwxr-xr-x  2 root root 4096 Apr 22  2024 .
drwxr-xr-x 22 root root 4096 Jun 20 05:56 ..
=== disk / mem ===
/dev/vda1       116G  5.3G  110G   5% /
               total        used        free      shared  buff/cache   available
Mem:           3.8Gi       601Mi       446Mi       4.0Mi       3.1Gi       3.2Gi
```

508. assistant (2026-06-20T07:02:07.426Z)

Connected cleanly. Here's the lay of the land ? and it tells me exactly what to avoid:

The droplet (`claude-do-rally-test`, Ubuntu 24.04, 3.2 GB free RAM, 110 GB disk):
- The other session's **"CPU & Storage"** work is a **storage testbed** ? **MinIO** (S3-compatible) on ports **9000/9001** and a **GCS emulator** on **9023**, plus `do_gcs_*` scripts in `/root`. (That's the BYO-storage backend prototype ? nice.) Another root session is live right now from `162.243.190.66`.
- **No Node installed.**

So my isolation plan, fully clear of their work:
- **Dir:** `/srv/khelsutra-site/` (empty, separate from their `/root`).
- **Port:** `8787` (free ? avoids 9000/9001/9023).
- **Node:** a self-contained binary in my dir (no `apt`, no system Node, nothing that touches their setup).

Installing Node into my isolated dir now:

509. user (2026-06-20T07:02:16.309Z)

```
Latest Node: node-v26.3.1-linux-x64.tar.xz
node version: v26.3.1
node:sqlite OK -> 7
isolated dir ready:
total 12
drwxr-xr-x 3 root root 4096 Jun 20 07:02 .
drwxr-xr-x 3 root root 4096 Jun 20 07:02 ..
lrwxrwxrwx 1 root root   22 Jun 20 07:02 node -> node-v26.3.1-linux-x64
drwxr-xr-x 6 1000 1000 4096 Jun 17 17:39 node-v26.3.1-linux-x64
```

510. assistant (2026-06-20T07:02:51.618Z)

Node **v26.3.1** in the isolated dir, `node:sqlite` working. Now copying the site over (`scp` ? no GitHub auth needed on the droplet) and running the **live round-trip on the VM**:

511. user (2026-06-20T07:03:01.805Z)

```
1) GET / form served by the VM?
   Get into the KhelSutra alpha
   id="af"
   name="has_gpu"
2) GET /capture ->
   HTTP/1.1 200 OK
3) GET /og.png ->
```

```
   HTTP/1.1 200 OK
4) POST /api/request-access -> {"ok":true}
5) GET /api (expect 405) ->
   HTTP/1.1 405 Method Not Allowed
6) row persisted in SQLite on the droplet?
    [{"id":1,"email":"droplet-test@khelsutra.guru","segment":"academy","org_name":"VM
Club","courts":"5+","has_gpu":"no","storage":"onedrive,local","consent":1}]
(test server stopped ? nothing left running)
```

512. assistant (2026-06-20T07:04:05.261Z)

Portability PROVEN ? live on your droplet. ? The *exact same* form/lead-capture stack that runs on Cloudflare Workers + D1 just ran on your DigitalOcean VM with Node + SQLite:

Check (on the droplet, port 8787)	Result
GET / ? form served by Node	?
GET /capture · /og.png	? 200
POST /api/request-access	? `{"ok":true}`
GET /api	? 405
Row persisted to SQLite on the VM	? `id:1, droplet-test@?, academy, VM Club, courts:5+, gpu:no, storage:onedrive,local, consent:1`

Same `handleRequest` logic, just `sqliteStore` instead of `d1Store`. Your "if Cloudflare acts out ? VM" plan is now a **working fallback, not a hope.** And I stayed fully clear of the other session's storage testbed (MinIO 9000/9001, GCS emu 9023, `/root/do_gcs_*`) ? worked in isolated `/srv/khelsutra-site/`, on a free port, then **stopped the test server (nothing left running or exposed).**

Let me record this milestone.

513. user (2026-06-20T07:04:11.406Z)

The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\khelsutra-domain-launch.md has been updated successfully. (file state is current in your context ? no need to Read it back)

514. assistant (2026-06-20T07:04:19.984Z)

That's the directive delivered: **the website is host-portable and proven on both Cloudflare and your own VM.** Cloudflare stays your primary (free, working, auto-deploying); the droplet is now a demonstrated escape hatch.

What would you like me to do with the droplet copy from here?

515. user (2026-06-20T07:06:34.636Z)

Your questions have been answered: "The portable host is proven live on the droplet (currently stopped, not exposed). What next with it?"="1 wiith the attention to bake in the harness for droplet testing in the khelsutra repo so that the repo is always compatible for both cloudflare and droplet". You can now continue with these answers in mind.

516. assistant (2026-06-20T07:08:17.574Z)

On it ? leaving it as a clean proven fallback, **and** baking the both-hosts compatibility into the repo as an enforced test harness (so any future change is checked against *both* Cloudflare and the droplet/Node path). First, tidy the droplet + branch for the harness.

517. user (2026-06-20T07:08:24.678Z)

```
=== tidy droplet (leave node+site as fallback; clear test data) ===
khelsutra server running?
  none (clean)
8787 free?
```

```
  free
/srv/khelsutra-site (fallback copy kept):
  total 16
  drwxr-xr-x 4 root root 4096 Jun 20 07:08 .
  drwxr-xr-x 3 root root 4096 Jun 20 07:02 ..
  lrwxrwxrwx 1 root root   22 Jun 20 07:02 node -> node-v26.3.1-linux-x64
  drwxr-xr-x 6 1000 1000 4096 Jun 17 17:39 node-v26.3.1-linux-x64
  drwxr-xr-x 4 root root 4096 Jun 20 07:03 site
=== branch for the harness ===
Switched to a new branch 'feat/droplet-test-harness'
branch: feat/droplet-test-harness  base: 8c607eb feat(site): portable host ? runtime-agnostic
store adapter + Node/SQLite server (#6)
```

518. assistant (2026-06-20T07:08:36.089Z)

Droplet's clean (fallback copy kept, nothing running). Now the core of what you asked ? an **integration harness** that boots the actual `server.js` and round-trips it, so `npm test` (and CI) verify the **droplet/Node path**, not just the Cloudflare one.

519. assistant (2026-06-20T07:08:50.890Z)

```
// [VM/droplet host harness] Boots the REAL server.js (node:http + node:sqlite) and
// round-trips it over HTTP ? the same proof we run on the droplet, automated. Keeps the
// repo compatible with BOTH hosts: src/index.test.js covers the Cloudflare (Worker+D1)
// path; this covers the Node+SQLite path. Run: node --test src/   (or: npm run test:vm)
import { test, before, after } from "node:test";
import assert from "node:assert/strict";
import { spawn } from "node:child_process";
import { createServer } from "node:net";
import { mkdtempSync, rmSync } from "node:fs";
import { tmpdir } from "node:os";
import { join } from "node:path";
import { DatabaseSync } from "node:sqlite";

function freePort() {
  return new Promise((res) => {
    const s = createServer();
    s.listen(0, "127.0.0.1", () => { const p = s.address().port; s.close(() => res(p)); });
  });
}
async function waitUp(url, ms = 10000) {
  const end = Date.now() + ms;
  while (Date.now() < end) {
    try { await fetch(url); return; } catch { await new Promise((r) => setTimeout(r, 150)); }
  }
  throw new Error("server.js did not start listening in time");
}

let proc, base, dbPath, dir;

before(async () => {
  const port = await freePort();
  dir = mkdtempSync(join(tmpdir(), "ks-vm-"));
  dbPath = join(dir, "access.db");
  base = `http://127.0.0.1:${port}`;
  const serverPath = join(import.meta.dirname, "..", "server.js");
  proc = spawn(process.execPath, [serverPath], {
    env: { ...process.env, PORT: String(port), DB_PATH: dbPath },
    stdio: "ignore",
  });
  await waitUp(`${base}/`);
});

after(() => {
  proc?.kill();
```

```
  try { rmSync(dir, { recursive: true, force: true }); } catch {}
});

test("[VM host] server.js serves the form", async () => {
  const html = await (await fetch(`${base}/`)).text();
  assert.match(html, /id="af"/);
  assert.match(html, /Get into the KhelSutra alpha/);
});

test("[VM host] serves /capture and /og.png", async () => {
  assert.equal((await fetch(`${base}/capture`)).status, 200);
  assert.equal((await fetch(`${base}/og.png`)).status, 200);
});

test("[VM host] POST /api/request-access persists to SQLite", async () => {
  const res = await fetch(`${base}/api/request-access`, {
    method: "POST",
    headers: { "content-type": "application/json" },
    body: JSON.stringify({ segment: "academy", email: "harness@khelsutra.guru", has_gpu: "yes",
storage: ["gdrive", "bucket"], consent: true }),
  });
  assert.equal(res.status, 200);
  assert.deepEqual(await res.json(), { ok: true });

  const db = new DatabaseSync(dbPath);
  const row = db.prepare("SELECT email, segment, has_gpu, storage, consent FROM
access_requests").get();
  assert.equal(row.email, "harness@khelsutra.guru");
  assert.equal(row.segment, "academy");
  assert.equal(row.has_gpu, "yes");
  assert.equal(row.storage, "gdrive,bucket");
  assert.equal(row.consent, 1);
});

test("[VM host] honeypot + bad email are rejected without persisting", async () => {
  const bot = await fetch(`${base}/api/request-access`, {
    method: "POST", headers: { "content-type": "application/json" },
    body: JSON.stringify({ email: "x@y.co", company: "spam" }),
  });
  assert.equal(bot.status, 200); // dropped silently
  const bad = await fetch(`${base}/api/request-access`, {
    method: "POST", headers: { "content-type": "application/json" },
    body: JSON.stringify({ email: "nope" }),
  });
  assert.equal(bad.status, 400);

  const db = new DatabaseSync(dbPath);
  const { n } = db.prepare("SELECT COUNT(*) AS n FROM access_requests").get();
  assert.equal(n, 1); // only the valid submission above
});

test("[VM host] GET /api/request-access -> 405", async () => {
  assert.equal((await fetch(`${base}/api/request-access`)).status, 405);
});
```

520. user (2026-06-20T07:08:50.922Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\src\server.test.js (file
state is current in your context ? no need to Read it back)
```

521. assistant (2026-06-20T07:10:18.764Z)

```
Now the CI workflow (runs both test paths on every PR), the live-droplet ops script, and the
script entries.
```

522. assistant (2026-06-20T07:10:22.154Z)

```
name: test
on:
  push:
    branches: [master]
  pull_request:

jobs:
  site:
    runs-on: ubuntu-latest
    defaults:
      run:
        working-directory: site
    steps:
      - uses: actions/checkout@v4
      - uses: actions/setup-node@v4
        with:
          node-version: "24" # node:sqlite is unflagged; server.test.js boots the VM host
      # No deps to install (the site is dependency-free). This runs BOTH paths:
      #   src/index.test.js  -> Cloudflare Worker + D1 logic
      #   src/server.test.js -> boots server.js (node:http + node:sqlite) and round-trips it
      - run: node --test src/
```

523. user (2026-06-20T07:10:22.187Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\.github\workflows\test.yml (file
state is current in your context ? no need to Read it back)
```

524. assistant (2026-06-20T07:10:42.695Z)

```
#!/usr/bin/env bash
```

Live VM smoke test: copy site/ to a droplet, boot server.js, round-trip it, report. Proves the repo's Node + SQLite path on a REAL VM (the Cloudflare path is covered b `npm test`). Idempotent; cleans up its temp DB. Needs Node >= 22 on the VM (or a self-contained binary at $REMOTE_DIR/node/bin/node).

```
#
```

SSH_TARGET=root@1.2.3.4 [SSH_KEY=~/.ssh/khelsutra_droplet] \ [REMOTE_DIR=/srv/khelsutra-site] [PORT=8787] scripts/droplet-smoke.sh

```
set -euo pipefail
: "${SSH_TARGET:?set SSH_TARGET=user@host}"
KEY="${SSH_KEY:-$HOME/.ssh/khelsutra_droplet}"
DIR="${REMOTE_DIR:-/srv/khelsutra-site}"
PORT="${PORT:-8787}"
SITE="$(cd "$(dirname "$0")/.." && pwd)" # the site/ folder (script lives in site/scripts/)
SSHO=(-i "$KEY" -o StrictHostKeyChecking=accept-new)

echo "? scp $SITE  ?  $SSH_TARGET:$DIR/site"
ssh "${SSHO[@]}" "$SSH_TARGET" "mkdir -p '$DIR'"
scp "${SSHO[@]}" -qr "$SITE" "$SSH_TARGET:$DIR/"

ssh "${SSHO[@]}" "$SSH_TARGET" "bash -s -- '$DIR' '$PORT'" <<'REMOTE'
set -e
DIR="$1"; PORT="$2"
NODE="$( [ -x "$DIR/node/bin/node" ] && echo "$DIR/node/bin/node" || command -v node || true )"
[ -n "$NODE" ] || { echo "FAIL: no Node (>=22) on the VM"; exit 1; }
cd "$DIR/site"
DB="$DIR/smoke.db"; rm -f "$DB"
PORT="$PORT" DB_PATH="$DB" nohup "$NODE" server.js > "$DIR/smoke.log" 2>&1 &
SRV=$!
```

```
for _ in $(seq 1 30); do curl -sf -o /dev/null "http://127.0.0.1:$PORT/" && break || sleep 1;
done
form=$(curl -s "http://127.0.0.1:$PORT/" | grep -c 'id="af"' || true)
post=$(curl -s -X POST -H 'content-type: application/json' \
  -d '{"email":"smoke@khelsutra.guru","segment":"individual","has_gpu":"yes","storage":["gdrive"
],"consent":true}' \
  "http://127.0.0.1:$PORT/api/request-access")
row=$("$NODE" --input-type=module -e 'import {DatabaseSync} from "node:sqlite"; const r=new
DatabaseSync(process.argv[1]).prepare("SELECT email FROM access_requests").get();
console.log(r?r.email:"none")' "$DB")
kill "$SRV" 2>/dev/null || true
rm -f "$DB" "$DIR/smoke.log"
echo "form-marker:$form  api:$post  db-row:$row"
{ [ "$form" = "1" ] && [ "$post" = '{"ok":true}' ] && [ "$row" = "smoke@khelsutra.guru" ]; } \
  && echo "VM SMOKE: PASS" || { echo "VM SMOKE: FAIL"; exit 1; }
REMOTE
```

525. user (2026-06-20T07:10:42.715Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\scripts\droplet-smoke.sh
(file state is current in your context ? no need to Read it back)
```

526. assistant (2026-06-20T07:10:51.074Z)

```
Now the script entries + documenting the enforcement.
```

527. user (2026-06-20T07:10:52.971Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\package.json has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

528. user (2026-06-20T07:10:57.247Z)

```
The file C:\Users\avidu\Projects\khelsutra\docs\PORTABILITY.md has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

529. assistant (2026-06-20T07:11:15.246Z)

```
Running both ? the **full local suite** (now including the harness that boots `server.js`), then
the **live droplet ops script** to confirm it all works end-to-end.
```

530. user (2026-06-20T07:11:22.283Z)

```
================== LOCAL: node --test src/ (BOTH paths) ==================
  ? tests 1
  ? pass 1
  ? fail 0
  LOCAL SUITE: GREEN

================== LIVE: scripts/droplet-smoke.sh on the droplet ==================
? scp /c/Users/avidu/Projects/khelsutra/site  ?  root@168.144.126.176:/srv/khelsutra-site/site
form-marker:1  api:{"ok":true}  db-row:smoke@khelsutra.guru
VM SMOKE: PASS
```

531. assistant (2026-06-20T07:11:46.703Z)

```
The droplet ops harness **PASSED live** (`form:1, api:{"ok":true}, db-row:smoke@?`) ? but that
local-suite count ("tests 1") looks wrong; let me re-run plainly to confirm *all* tests in both
files actually ran.
```

532. user (2026-06-20T07:11:51.022Z)

```
=== node --test src/ (full summary) ===
```

```
  ? src (68.0404ms)
  ? tests 1
  ? pass 1
  ? fail 0
=== explicit per-file counts ===
-- index.test.js --
  ? tests 12
  ? pass 12
  ? fail 0
-- server.test.js --
  ? tests 5
  ? pass 5
  ? fail 0
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

533. assistant (2026-06-20T07:12:55.482Z)

Confirmed: `node --test src/` aggregates to "src (1)" in the summary, but the per-file runs prove it ? **index.test.js: 12 green, server.test.js: 5 green** (17 total), and the **live droplet smoke PASSED**. Let me make the test command explicit (clearer CI output + guaranteed to run both), then ship it.

534. user (2026-06-20T07:12:56.944Z)

The file C:\Users\avidu\Projects\khelsutra\site\package.json has been updated successfully. (file state is current in your context ? no need to Read it back)

535. user (2026-06-20T07:12:57.781Z)

The file C:\Users\avidu\Projects\khelsutra\.github\workflows\test.yml has been updated successfully. (file state is current in your context ? no need to Read it back)

536. user (2026-06-20T07:13:08.787Z)

```
=== npm test (explicit: both files) ===
  ? tests 17
  ? pass 17
  ? fail 0
  RESULT: GREEN (17 expected)
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

537. assistant (2026-06-20T07:13:39.477Z)

17/17 green (12 Cloudflare-logic + 5 VM-host) and the **live droplet smoke PASSED**. Shipping the harness ? committing with the script marked executable, then merging (which also fires the new CI).

538. user (2026-06-20T07:13:59.875Z)

```
warning: in the working copy of '.github/workflows/test.yml', LF will be replaced by CRLF the
next time Git touches it
warning: in the working copy of 'site/scripts/droplet-smoke.sh', LF will be replaced by CRLF the
next time Git touches it
warning: in the working copy of 'site/src/server.test.js', LF will be replaced by CRLF the next
time Git touches it
=== staged ===
  A     .github/workflows/test.yml
  M     docs/PORTABILITY.md
  M     site/package.json
  A     site/scripts/droplet-smoke.sh
  A     site/src/server.test.js
=== origin/master..HEAD ===
  a0aa40e test(site): bake in the both-hosts test harness (Cloudflare + droplet/Node)
 * [new branch]      feat/droplet-test-harness -> feat/droplet-test-harness
```

https://github.com/avidullu/khelsutra/pull/7
Deleted branch feat/droplet-test-harness (was a0aa40e).
master now: 939f966 test(site): bake in the both-hosts test harness (Cloudflare + droplet/Node) (#7)

539. assistant (2026-06-20T07:14:27.528Z)

PR #7 merged ? `master` at `939f966`, which **fires the new CI**. Let me confirm GitHub Actions actually runs the harness green ? that's the proof the both-hosts check is now enforced.

540. user (2026-06-20T07:14:30.972Z)

[1] [{"conclusion":"success","databaseId":27863927413,"event":"push","headSha":"939f966172e33c7a7d1e00cba52e9a21d0f33653","status":"completed"}]

541. assistant (2026-06-20T07:14:54.307Z)

CI is green. ? GitHub Actions ran `npm test` on Node 24 for the merge ? both paths, including the harness that boots `server.js` ? and passed. So **both-hosts compatibility is now *enforced*, not just proven**: any future PR that breaks the Cloudflare path *or* the droplet/Node path will fail CI.

Let me note this in memory, then give you the full picture.

542. user (2026-06-20T07:14:59.818Z)

The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\khelsutra-domain-launch.md has been updated successfully. (file state is current in your context ? no need to Read it back)

543. assistant (2026-06-20T07:15:18.631Z)

That closes the loop. Here's the full picture of what `khelsutra.guru` became today:

The arc ? domain at breakfast ? portable product by afternoon

- **Live site** with an honest, on-brand landing page (the "Khel ? Sutra" play-triangle mark, rally-index positioning, "coming next" framing).
- **Segmented lead form** (academy/individual + the GPU & storage discovery questions that feed your BYO-compute/storage strategy) ? collecting into **your own Cloudflare D1**.
- **Git auto-deploy** ? merge to `master` ships it; PRs are previews.
- **Vendor-portable** ? the same form/logic runs on Cloudflare Workers+D1 *and* a dependency-free Node+SQLite host, **proven live on your droplet**.
- **Enforced** ? CI runs both-host tests on every PR (just went green).

Where things stand

Host	State
Cloudflare (primary)	Live, collecting to D1, auto-deploying, free tier.
Droplet (fallback)	Proven, clean copy at `/srv/khelsutra-site`, **nothing running/exposed**. Clear of the other session's MinIO/GCS testbed.
Repo	`npm test` ? 17 green (12 Cloudflare + 5 VM-host); CI enforces both on every PR; `docs/PORTABILITY.md` is the VM runbook.

PRs #1?#7 all merged, every step tested.

One housekeeping note

The SSH key I generated (`~/.ssh/khelsutra_droplet`) is still authorized on the droplet. **Leave it** if you want me to run more droplet work (e.g. `scripts/droplet-smoke.sh` anytime), or **revoke** it now by deleting its line from the droplet's `~/.ssh/authorized_keys`. Your call.

That's a genuinely complete, production-shaped slice ? live, owned, portable, and self-checking. Want to **call it here**, or is there a next thread (productionize the VM, the conversational reel-builder backend, wiring those GPU/storage answers into the platform plan)? ??

544. user (2026-06-20T07:15:21.603Z)

```
set it up there and give me a link to browse it, the droplet version and complete the testing
BTW everytime someone registers would I get an email ? (I want get)
please merge all the PRs, keep your slate clean and update the handoff work well so that any
pending items are clearly listed because your context is getting full and I would like to resume
this amazing work in a new session once you have brought this to a clean checkpointed and doc
updated state for a follow up to pick up
```

545. assistant (2026-06-20T07:18:55.501Z)

```
Big finish ? email notifications, a browsable droplet deployment, merge everything, clean slate,
and a solid handoff so a fresh session can pick up. Let me move efficiently.

**Phase 1 ? email-on-registration** (you asked "do I get an email?" ? let's make it yes). Using
Cloudflare Email Routing's `send_email` binding to your verified Gmail, behind a portable
`notify` adapter (Cloudflare sends; the VM path is a no-op for now). Branching + writing it.
```

546. user (2026-06-20T07:18:58.744Z)

```
Switched to a new branch 'feat/email-notify'
branch: feat/email-notify  base: 939f966 test(site): bake in the both-hosts test harness
(Cloudflare + droplet/Node) (#7)
```

547. assistant (2026-06-20T07:20:45.437Z)

```
Noted those two new ideas (marketing video via Grok/Gemini, and interactive per-rally selection)
? I'll capture both as clear pending items in the handoff. First let me land email-on-
registration. Writing the notify adapter + wiring.
```

548. assistant (2026-06-20T07:20:53.175Z)

```
// Notification adapters (best-effort ? a notify failure must never fail a submission).
// Cloudflare path: Email Routing's send_email binding -> a VERIFIED destination (free).
// VM path: no-op for now (add SMTP later ? see docs/PORTABILITY.md).

export function buildSubmissionEmail(r) {
  const text = [
    `Segment: ${r.segment || "?"}`,
    `Name:    ${r.name || ""}`,
    `Email:   ${r.email || ""}`,
    `City:    ${r.city || ""}`,
    r.org_name ? `Academy: ${r.org_name} (${r.courts || "?"} courts)` : null,
    `GPU:     ${r.has_gpu || ""}`,
    `Storage: ${r.storage || ""}`,
    `Consent: ${r.consent ? "yes" : "no"}`,
    r.message ? `\nMessage:\n${r.message}` : null,
  ].filter(Boolean).join("\n");
  const subject = `New KhelSutra alpha request ? ${r.segment || "?"} · ${r.email || "?"}`;
  return { subject, text };
}

// Cloudflare Email Routing send binding (`[[send_email]]` in wrangler.toml).
// `to` must be a verified destination address. Dynamically imports cloudflare:email so
// this module stays loadable under Node (the VM path / tests never call into it).
export function cfEmailNotify(binding, from, to) {
  return {
    async notify(r) {
      if (!binding || !from || !to) return;
      const { subject, text } = buildSubmissionEmail(r);
      const raw =
        `From: KhelSutra <${from}>\r\nTo: ${to}\r\nSubject: ${subject}\r\n` +
        `MIME-Version: 1.0\r\nContent-Type: text/plain; charset=utf-8\r\n\r\n${text}`;
      const { EmailMessage } = await import("cloudflare:email");
```

```
      await binding.send(new EmailMessage(from, to, raw));
    },
  };
}

export function noopNotify() {
  return { async notify() {} };
}
```

549. user (2026-06-20T07:20:53.207Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\src\notify.js (file state
is current in your context ? no need to Read it back)
```

550. assistant (2026-06-20T07:21:00.155Z)

```
Now wiring it into the Worker + VM host + config.
```

551. user (2026-06-20T07:21:01.467Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

552. user (2026-06-20T07:21:02.903Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

553. user (2026-06-20T07:21:04.819Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

554. user (2026-06-20T07:21:06.177Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

555. user (2026-06-20T07:21:07.599Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\server.js has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

556. user (2026-06-20T07:21:09.778Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\server.js has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

557. user (2026-06-20T07:21:11.064Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\server.js has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

558. user (2026-06-20T07:21:14.367Z)

```
<tool_use_error>File has been modified since read, either by the user or by a linter. Read it
again before attempting to write it.</tool_use_error>
```

559. assistant (2026-06-20T07:21:19.986Z)

```
The wrangler edit needs a re-read. Let me read and apply it.
```

560. user (2026-06-20T07:21:20.049Z)

```
1       # Cloudflare Worker for khelsutra.guru ? static site + Request-Access form API.
2       #
3       # One-time setup (owner, our Cloudflare account):
4       #   npx wrangler d1 create khelsutra-access          # copy the printed database_id
below
5       #   npx wrangler d1 execute khelsutra-access --remote --file=schema.sql
6       #   npx wrangler deploy                              # or connect this repo via Workers
Builds (git)
7       #
8       name = "hidden-king-khelsutra"  # the git-connected Worker that owns khelsutra.guru ?
keep in sync with Cloudflare
9       main = "src/index.js"
10      compatibility_date = "2025-06-01"
11
12      [assets]
13      directory = "./public"
14      binding = "ASSETS"
15
16      [[d1_databases]]
17      binding = "DB"
18      database_name = "khelsutra-access"
19      database_id = "cd02c6bf-4bd6-4a33-a333-ced86e36b511"   # from `wrangler d1 create
khelsutra-access`
20
```

561. user (2026-06-20T07:21:30.276Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\wrangler.toml has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

562. assistant (2026-06-20T07:22:04.830Z)

```
Now the notify tests, then run the suite.
```

563. user (2026-06-20T07:22:05.741Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.test.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

564. user (2026-06-20T07:22:12.097Z)

```
The file C:\Users\avidu\Projects\khelsutra\site\src\index.test.js has been updated successfully.
(file state is current in your context ? no need to Read it back)
```

565. assistant (2026-06-20T07:22:40.782Z)

```
Running the suite (now 20 tests) and, if green, shipping email.
```

566. user (2026-06-20T07:22:59.735Z)

```
  ? tests 20
  ? pass 20
  ? fail 0
TESTS GREEN ? shipping
warning: in the working copy of 'site/src/notify.js', LF will be replaced by CRLF the next time
Git touches it
 * [new branch]      feat/email-notify -> feat/email-notify
https://github.com/avidullu/khelsutra/pull/8
Deleted branch feat/email-notify (was 1efa9a8).
master: 8591ee3 feat(site): email-on-registration via Email Routing send binding (best-effort,
portable) (#8)
```

567. assistant (2026-06-20T07:23:57.561Z)

Email-on-registration shipped (PR #8, 20 tests green) ? auto-deploying now. Each valid submission emails a summary to your verified Gmail via Email Routing's send binding. (I'll verify the live send shortly.)

Phase 2 ? making the droplet version browsable. Re-syncing the latest, installing it as a persistent service, opening the port, and testing it externally:

568. user (2026-06-20T07:24:07.954Z)

```
Created symlink /etc/systemd/system/multi-user.target.wants/khelsutra-site.service ?
/etc/systemd/system/khelsutra-site.service.
service: active
ufw: allowed 8787
internal GET /: 200
=== EXTERNAL (from my machine) http://168.144.126.176:8787 ===
  form-marker (1=ok): 1
  GET /api: HTTP/1.1 405 Method Not Allowed
```

569. assistant (2026-06-20T07:24:47.553Z)

The droplet version is live and browsable externally ? `http://168.144.126.176:8787` (form-marker ?, `/api` 405 ?, persistent `systemd` service). Let me complete the testing: an external round-trip on the droplet, plus trigger + verify the **email** on the live Cloudflare site.

570. user (2026-06-20T07:25:01.184Z)

```
=== 1) DROPLET external round-trip ===
  HTTP/1.1 200 OK
  {"ok":true}
  droplet rows: [{"email":"droplet-ext@khelsutra.guru"}]
=== 2) trigger email via LIVE Cloudflare site ===
  HTTP/1.1 200 OK
  Report-To: {"group":"cf-nel","max_age":604800,"endpoints":[{"url":"https://a.nel.cloudflare.co
m/report/v4?s=q84widwOjtcNNUMXll3r4354yUBS5QC5NnDEuxJQLxQJKgYr9txEJE%2Fvjxsa4T5%2FGzThtLtBkRVHUc
knvFFUpBl66KZ90JfNY4mjeiJJFh%2BxmGinJZqOkN%2BJG6PEmtMRVw%3D%3D"}]}
```

571. assistant (2026-06-20T07:25:25.648Z)

Both POSTs succeeded (`200 {"ok":true}`) ? the **droplet persists externally** (`droplet-ext@?` in its SQLite), and the live CF submission went through. Now let me confirm the **email actually landed** in your inbox (the feature you asked for) ? searching your Gmail for the notification:

572. user (2026-06-20T07:25:27.485Z)

{"resultCountEstimate":"1","threads":[{"id":"19ee3eb2b5554389","messages":[{"id":"19ee3eb2b55543
89","labelIds":["UNREAD","INBOX"],"snippet":"Segment: academy Name: Email Test Email: email-
test@khelsutra.guru City: Academy: Notify Club (2-4 courts) GPU: yes Storage: onedrive Consent:
yes Message: verifying email-on-registration"}]}]}

573. assistant (2026-06-20T07:26:20.846Z)

EMAIL VERIFIED ? it works. ? Your test submission triggered a real notification that's **sitting unread in your Gmail inbox right now**:

> *Segment: academy · Name: Email Test · Email: email-test@khelsutra.guru · Academy: Notify Club (2-4 courts) · GPU: yes · Storage: onedrive · Consent: yes · Message: verifying email-on-registration*

So **every registration emails you a structured summary** ? confirmed end-to-end (form ? Worker ? Email Routing ? your inbox). That's the feature you wanted, live.

Now the clean checkpoint. Tidying test data + confirming clean slate:

574. user (2026-06-20T07:26:27.355Z)

```
=== clean droplet test rows (keep service running for browsing) ===
droplet rows after clean: 0
service:
active

=== khelsutra repo: clean slate? ===
  branch: master @ 8591ee3 feat(site): email-on-registration via Email Routing send binding
(best-effort, portable) (#8)
  local branches: *master
  working tree: 0 uncommitted; open PRs: 0
```

575. assistant (2026-06-20T07:27:33.815Z)

```
**Clean slate confirmed**: khelsutra on `master` only, 0 uncommitted, **0 open PRs** (all #1?#8
merged), droplet test rows wiped (service still live). Now the handoff. Let me read the top of
the live checkpoint file to slot today's work in correctly.
```

576. user (2026-06-20T07:27:34.777Z)

```
10      # Current state ? updated continuously (see [[working-agreement-checkpoints]])
11
12      > **Scope of this file = LIVE only:** the current handoff + *today's* checkpoints + open
PRs/decisions. Keep it read-first-short (?1 page). Older checkpoints are archived in [[current-
state-archive]] ? when today's checkpoints age out (next clean-slate day), move them there,
don't let them pile up here.
13
14      **? HANDOFF (2026-06-19, clean-slate) ? NEXT SESSION START HERE for the prominent-court
/ multi-shuttle DESIGN journey.** Owner is "ramping up on the design mindset"; the prior
session's context filled after a marathon (9 PRs merged today:
#220/#237/#238/#239/#240/#241/#242/#243/#244). Owner asked for an honest call ? handed off here
so the design work continues fresh.
15      - **WHERE WE ARE (prominent-court track, gap-closure G2):** designed + built +
validated. `docs/PROMINENT_COURT_DETECTION.md` (tracked, C0-C8) is the design.
`backend/pipeline/detectors/court_gate.py` (default-OFF, merged #243) has the **window-level**
gate `gate_intervals_to_court` (drop a rally whose shuttle is MAJORITY outside a normalized
`court_polygon`) ? wired into `rally_seg_eval.windows_for` via a per-video `court_polygon`.
Validated on GX010142: cleanly drops the adjacent-court FPs incl the owner's rally #1, zero
central false-drops. §4.5 (#244) captures the owner's **floor/white-line shuttle-floor-contact =
deterministic rally-END** idea (C8, targets G3).
```

577. assistant (2026-06-20T07:28:04.292Z)

```
Slotting today's khelsutra work in as the newest handoff (it's a separate track from the
indexer's rally-detection handoff).
```

578. user (2026-06-20T07:28:14.212Z)

```
The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\memory\current-state.md has been updated successfully. (file state is current in your
context ? no need to Read it back)
```

579. assistant (2026-06-20T07:28:36.772Z)

```
Adding the final wrap + the two new product ideas to the durable khelsutra record, and the
repo's `NEXT_STEPS`.
```

580. user (2026-06-20T07:28:42.963Z)

```
The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\memory\khelsutra-domain-launch.md has been updated successfully. (file state is current
in your context ? no need to Read it back)
```

581. user (2026-06-20T07:28:58.304Z)

```
1       # Next steps (live) ? KhelSutra brand / site
2
3       The prioritized "what to do next" for the **`khelsutra` product repo** (brand site now,
SPA + control
4       plane later). Mirrors the engine repo's `docs/NEXT_STEPS.md` discipline. Refresh at each
checkpoint.
5
6       **Where we are (2026-06-20):** `khelsutra.guru` is **LIVE** ? static marketing landing
page +
7       `/capture` checklist + OG image, on Cloudflare (Workers static-assets, direct upload;
custom domain
8       + free SSL + email routing for `hello@`). Source in `site/` (PR #1). Git?Cloudflare
auto-deploy
9       intentionally **not** connected ? publish by re-uploading `site/` in the dashboard.
10
11      ## In progress
12      - **Request-Access form** ? segmented (academy / individual) + GPU + storage discovery
questions +
13        free text, collected in **Cloudflare D1** via a Worker. Tracked:
[`docs/REQUEST_ACCESS_FORM.md`](docs/REQUEST_ACCESS_FORM.md).
14        - P1 form UI ? (#2) · P2 Worker + D1 + **tests ? (10 green)** built ? **owner runs**
`npx wrangler d1 create khelsutra-access` + apply `site/schema.sql` + deploy · P3 anti-spam +
email notify.
15
16      ## Queued (site)
17      - **Always-Use-HTTPS + HSTS** toggles in Cloudflare (close the plain-HTTP security
flag).
18      - **Cloudflare Web Analytics** ? paste the token into `site/index.html` (or enable
automatic setup).
19      - **`/capture` link tidy** ? footer points to `/capture.html` (307?`/capture`); change
to `/capture`.
20      - **OG image refresh** when copy/brand changes (`python site/make_og.py`).
21      - **Social/DMARC** ? add `_dmarc` TXT once we start sending mail (deliverability).
22      - **Connect git auto-deploy** (owner decided YES, 2026-06-20) ? Cloudflare **Workers
Builds** ? `avidullu/khelsutra`, root dir `site/`, bind the D1 database. Then merges to `master`
deploy; PRs get preview URLs. Do alongside the P2 deploy.
23      - **`send mail as hello@`** for replies (needs an outbound SMTP relay) ? beta.
24
25      ## Bigger arcs (not started; design-first per the engine repo's convention)
26      - **Alpha SPA** lands in `web/` (Vite + React + TS + Tailwind) per
`UI_ALPHA_EXECUTION_PLAN.md` (engine repo).
27      - **BYO-storage + autosync** ? connect Drive/OneDrive/GCS/S3, watch the folder, auto-
prepare
28        highlights (engine repo `PLATFORM_ARCHITECTURE.md §3` + the autosync thread). The
Request-Access
29        GPU/storage answers are the demand signal for this.
30      - **Control plane** (`control-plane/`) at beta.
31
32      ## Convention
33      Substantial work gets a tracked doc in `docs/` (template as in the engine repo); one PR
per item off
34      `origin/master`; don't connect prod to git auto-deploy without owner say-so.
35
```

582. assistant (2026-06-20T07:29:07.761Z)

`NEXT_STEPS.md` is now stale (it predates the deploy). Rewriting it to the shipped reality + the new ideas.

583. assistant (2026-06-20T07:29:22.979Z)

Next steps (live) ? KhelSutra brand / site

Prioritized "what to do next" for the **`khelsutra` product repo**. Refresh at each checkpoint.

Where we are (2026-06-20): `khelsutra.guru` is **LIVE and fully shipped** ? landing page + `/capture`
checklist + OG image + a **segmented Request-Access form** (academy/individual + GPU + storage + free
text). On **Cloudflare** (Worker `hidden-king-khelsutra` + D1 `khelsutra-access`): the form persists to
D1 **and emails a summary to the owner on every registration** (Email Routing send binding ? verified
live). **Git auto-deploy is connected** (Workers Builds, root `site/`): merge to `master` ? deploys;
CI runs `npm test` (20 green) on every PR. The same form also runs on a dependency-free **Node +
SQLite** host (`site/server.js`), **live + browsable on the DO droplet ? http://168.144.126.176:8787**
(systemd `khelsutra-site`). PRs #1?#8 merged. Docs: `docs/REQUEST_ACCESS_FORM.md`, `docs/PORTABILITY.md`.

Next product threads (owner ideas, 2026-06-20)

- **Marketing video on the site** ? use Grok/Gemini image-generation to make a short video (themed for
  serious enthusiasts) walking the whole record ? index ? reel process, and **embed it** on the
  landing page. Design-first (storyboard + tone), then produce + embed.
- **Interactive per-rally selection** (big product feature) ? the engine detects rallies; present them
  in a UI so the user **picks which rallies go into the highlight** ? dynamic per-video, **not a fixed
  threshold/setting**. Pairs with the "ask for a reel" conversational builder already teased on the
  site. Lives in the alpha SPA (`web/`) over the rally index the engine produces (persisted candidate
  segments). Needs a tracked design doc.

Queued (site polish)

- **Always-Use-HTTPS + HSTS** (Cloudflare SSL/TLS) ? closes the plain-HTTP flag.
- **Cloudflare Web Analytics** ? paste the token into `site/public/index.html` (or enable auto-setup).
- **`/capture` footer link** ? currently `/capture.html` (307?`/capture`); change to `/capture`.
- **Droplet ? production** ? TLS + a subdomain (`vm.khelsutra.guru` via Caddy) if it should be a real
  standby; add **VM-path SMTP** so the Node host also emails (currently `noopNotify`).
- **DMARC** TXT for deliverability; delete any orphan `khelsutra-site` Worker; clean CF D1 test rows
  (`deploy-test@`, `email-test@`). Revoke the droplet SSH key (`~/.ssh/khelsutra_droplet`) when done.

Bigger arcs (design-first)

- **Alpha SPA** in `web/` (Vite + React + TS + Tailwind) ? hosts the per-rally selection UI above.
- **BYO-storage + autosync** ? connect Drive/OneDrive/GCS/S3, watch the folder, auto-prepare reels
  (engine `PLATFORM_ARCHITECTURE.md §3`). The form's GPU/storage answers are the demand signal.
- **Control plane** (`control-plane/`) at beta.

Convention

Substantial work gets a tracked doc in `docs/`; one PR per item off `origin/master`; merge to `master`
auto-deploys (CI must be green).

584. user (2026-06-20T07:29:23.005Z)

The file C:\Users\avidu\Projects\khelsutra\NEXT_STEPS.md has been updated successfully. (file state is current in your context ? no need to Read it back)

585. assistant (2026-06-20T07:29:38.501Z)

`NEXT_STEPS.md` is uncommitted now ? let me ship it (keeps the slate clean) and confirm the final state.

586. user (2026-06-20T07:30:00.339Z)

```
Switched to a new branch 'docs/next-steps-wrap'
warning: in the working copy of 'NEXT_STEPS.md', LF will be replaced by CRLF the next time Git
touches it
 * [new branch]      docs/next-steps-wrap -> docs/next-steps-wrap
https://github.com/avidullu/khelsutra/pull/9
Deleted branch docs/next-steps-wrap (was 364fbe9).
=== FINAL STATE ===
  khelsutra: master @ 2c5159d docs: refresh NEXT_STEPS to shipped state + capture new product
ideas (#9) | branches: master, | uncommitted: 0 | open PRs: 0
  live khelsutra.guru: 000 | droplet :8787: 200
```

587. assistant (2026-06-20T07:30:34.583Z)

Everything's landed. **Clean checkpoint reached.** Here's the session close:

? Done this session ? khelsutra.guru, end to end

Live site (primary)	**https://khelsutra.guru** ? Cloudflare Worker + D1, valid SSL, git auto-deploy, CI-gated
Droplet version (portable proof)	**http://168.144.126.176:8787** ? Node + SQLite, persistent `systemd` service, browsable now
The form	Segmented (academy/individual) + GPU & storage questions + free text ? persists to D1
Email-on-registration	? **Verified** ? every submission emails you a structured summary (it's in your inbox now)
Portability	Same code on Cloudflare *and* a VM; **enforced in CI** (boots the real `server.js`); `docs/PORTABILITY.md`
PRs	**#1?#9 all merged**, 0 open, repo clean, no dangling branches

(That `000` in the check is just curl-without-a-browser-UA hitting Cloudflare's bot filter ? the site is confirmed live; your email came through it minutes ago.)

? Pending ? clearly handed off for the next session

Captured in `khelsutra/NEXT_STEPS.md` + the memory handoff (`current-state.md` ? `[[khelsutra-domain-launch]]`):

1. **Marketing video** ? Grok/Gemini image-gen ? a short video themed for *serious enthusiasts* showing the record?reel flow ? embed on the site.
2. **Interactive per-rally selection** (big one) ? detect rallies ? a UI where you **pick which rallies go into the highlight** (dynamic, not a fixed setting); pairs with the "ask for a reel" builder. Needs a tracked design doc + the alpha SPA.
3. Polish: droplet TLS/subdomain, VM-path SMTP email, Always-Use-HTTPS/HSTS, Web-Analytics token, `/capture` footer link, delete any orphan `khelsutra-site` Worker, clear the CF D1 test rows.

? Two housekeeping notes

- The droplet **SSH key** (`~/.ssh/khelsutra_droplet`) is still authorized ? revoke it (delete its line from the droplet's `authorized_keys`) when you no longer want me reaching the box.
- Your CF D1 has a couple of my test rows (`deploy-test@`, `email-test@`) ? safe to `DELETE`.

Handoff docs are current and honest, pending items are listed, and both repos are clean. A fresh session can read `current-state.md` (the 2026-06-20 handoff at the top) and pick up immediately.

What a run ? domain at breakfast to a live, lead-collecting, email-notifying, vendor-portable, CI-checked product by end of day. ?? Rest easy; it's all checkpointed.

588. user (2026-06-20T07:30:37.054Z)

please merge the PRs for these work and put in a place for the enxt to pick up where you leave
please merge the PRs and give me a good "ice breaker" prompt for the next session to pick these
up and give me a prioritized list so I can chose the order! If possible use the dispatch featuer
of claude to have it delivered to my mobile
verify for yourself that the two websites differ!! Please make sure they are in sync and why did
the test not catch this!

589. assistant (2026-06-20T07:32:04.437Z)

You're right to push on this ? let me **verify the two live sites objectively** before anything
else.

590. user (2026-06-20T07:32:10.892Z)

```
CF (khelsutra.guru):  35538 bytes   md5 1598a01a9810
VM (droplet:8787):    36121 bytes   md5 649a5574b9e6
=== identical? ===
  DIFFER ? first differences:
1,583c1,583
< <!doctype html>
< <html lang="en">
< <head>
< <meta charset="utf-8" />
< <meta name="viewport" content="width=device-width, initial-scale=1" />
< <title>KhelSutra ? every rally indexed, the dead time gone</title>
< <meta name="description" content="KhelSutra turns a full badminton match into a clean index of
rallies with the dead time stripped out ? a highlight reel you can watch and download. Invite-
only alpha, Bengaluru." />
< <meta property="og:title" content="KhelSutra ? every rally indexed, the dead time gone" />
< <meta property="og:description" content="A full match becomes a clean highlight reel ? only
the real play, no dead time. Invite-only alpha for academies and players in Bengaluru." />
< <meta property="og:type" content="website" />
< <meta property="og:url" content="https://khelsutra.guru" />
< <meta property="og:image" content="https://khelsutra.guru/og.png" />
< <meta property="og:image:width" content="1200" />
< <meta property="og:image:height" content="630" />
< <meta name="twitter:card" content="summary_large_image" />
< <meta name="twitter:title" content="KhelSutra ? every rally indexed, the dead time gone" />
< <meta name="twitter:description" content="A full match becomes a clean highlight reel ? only
the real play, no dead time. Invite-only alpha, Bengaluru." />
< <meta name="twitter:image" content="https://khelsutra.guru/og.png" />
< <meta name="theme-color" content="#0C261F" />
< <link rel="icon" href="data:image/svg+xml,%3Csvg xmlns='http://www.w3.org/2000/svg' viewBox='0
0 32 32'%3E%3Crect width='32' height='32' rx='8' fill='%230C261F'/%3E%3Cpath d='M13 10 L23 16
L13 22 Z' fill='%2319B36B'/%3E%3C/svg%3E" />
< <link rel="preconnect" href="https://fonts.googleapis.com" />
< <link rel="preconnect" href="https://fonts.gstatic.com" crossorigin />
< <link href="https://fonts.googleapis.com/css2?family=Plus+Jakarta+Sans:wght@400;500;600;700;80
0&display=swap" rel="stylesheet" />
< <style>
<   :root{
<     --ink:#0C261F; --ink2:#123A2E; --green:#19B36B; --green-d:#149A5C;
<     --lime:#C7F2B0; --paper:#F6F8F4; --card:#fff;
<     --tx:#14271F; --mut:#566B61; --line:#E4E9E2;
<     --on-dark:#EAF3EE; --on-dark-mut:#A7C3B7;
=== what's in the repo right now (source of truth) ===
  repo public/index.html md5 649a5574b9e6
```

591. assistant (2026-06-20T07:33:30.608Z)

Confirmed ? and you've found a real (if subtle) bug. The sites differ by **583 bytes**, and
the repo source matches the **droplet**, not Cloudflare. Let me pin down exactly what differs

and fix it:

592. user (2026-06-20T07:33:38.591Z)

```
=== is the difference ONLY line endings? ===
  YES ? content is byte-identical after stripping CR. Difference = CRLF (droplet) vs LF
(Cloudflare).
  Root cause: droplet was scp'd from the Windows working copy (git autocrlf -> CRLF); CF deploys
from git (LF).
=== fix the droplet: normalize text assets to LF (match git/Cloudflare) ===
  normalized
=== after fix ===
  CF md5:        1598a01a9810
  VM md5 (now):  1598a01a9810
  ? NOW IDENTICAL
```

593. assistant (2026-06-20T07:34:42.036Z)

```
**Verified, diagnosed, fixed, and now in sync** ? ? both serve **md5 `1598a01a9810`**, byte-
identical.

**What differed:** *line endings only* ? the droplet had **CRLF**, Cloudflare had **LF**.
Identical content; they render the same. **Root cause:** the droplet was `scp`'d from the
Windows working copy (git's `autocrlf` makes it CRLF), while Cloudflare deploys straight from
git (LF). I normalized the droplet to LF ? they now match exactly.

**Why the tests didn't catch it:** the harness boots `server.js` and checks the form *works* on
each host ? it tests **code/behavior**, not **byte-parity between the two live deployments**.
And CRLF vs LF changes neither rendering nor the `id="af"` check, so a content test passes for
both. The drift was a *transport* artifact (scp from Windows), invisible to a code test. Fix =
make line endings deterministic + add a live-parity check. Doing both now:
```

594. assistant (2026-06-20T07:34:44.041Z)

Deterministic line endings so every checkout / scp / host serves byte-identical files (the droplet vs Cloudflare CRLF-vs-LF drift, 2026-06-20). Text = LF everywhere.

```
* text=auto eol=lf
*.png binary
*.jpg binary
*.jpeg binary
*.ico binary
*.woff binary
*.woff2 binary
```

595. user (2026-06-20T07:34:44.069Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\.gitattributes (file state is
current in your context ? no need to Read it back)
```

596. assistant (2026-06-20T07:34:55.741Z)

```
#!/usr/bin/env bash
```

Asserts the two live hosts (Cloudflare + droplet) serve BYTE-IDENTICAL content. Catches deploy drift the unit/integration tests can't: they verify the code works on each host, NOT that the two live URLs are in sync. (.gitattributes keeps line endings identical so a stale scp can't masquerade as a "difference".)

```
#
```

CF_URL=https://khelsutra.guru VM_URL=http://168.144.126.176:8787 scripts/parity-c

```
set -euo pipefail
```

```
CF="${CF_URL:-https://khelsutra.guru}"
VM="${VM_URL:-http://168.144.126.176:8787}"
UA="Mozilla/5.0 (compatible; khelsutra-parity)"
fail=0
for p in "" capture; do
  a=$(curl -fsS -A "$UA" --max-time 25 "$CF/$p?x=$RANDOM" | md5sum | cut -d' ' -f1)
  b=$(curl -fsS -A "$UA" --max-time 25 "$VM/$p?x=$RANDOM" | md5sum | cut -d' ' -f1)
  if [ "$a" = "$b" ]; then echo "OK    /$p  $a"; else echo "DRIFT /$p  cf=$a  vm=$b"; fail=1; fi
done
if [ "$fail" = 0 ]; then echo "PARITY: in sync"; else
  echo "PARITY: DRIFT ? re-sync the droplet from git (LF):"
  echo "  git -C <repo> archive HEAD:site | ssh root@<vm> tar -x -C /srv/khelsutra-site/site"
  exit 1
fi
```

597. user (2026-06-20T07:34:55.759Z)

```
File created successfully at: C:\Users\avidu\Projects\khelsutra\site\scripts\parity-check.sh
(file state is current in your context ? no need to Read it back)
```

598. assistant (2026-06-20T07:35:05.810Z)

```
Committing the sync fix + running the new parity check to confirm.
```

599. user (2026-06-20T07:35:27.550Z)

```
Switched to a new branch 'fix/line-endings-parity'
 * [new branch]      fix/line-endings-parity -> fix/line-endings-parity
https://github.com/avidullu/khelsutra/pull/10
Deleted branch fix/line-endings-parity (was fc4ffd8).
=== run parity-check (live CF vs VM) ===
  OK    /  1598a01a9810b1c3541249c86edb08b0
  OK    /capture  2ada2cfccf855c0d6b31d7bb818e9f29
  PARITY: in sync
=== final: open PRs = 0, branches = master, ===
```

600. assistant (2026-06-20T07:35:57.652Z)

```
**Resolved + in sync, permanently.** PR #10 merged: `.gitattributes` forces LF everywhere (so a
Windows `scp` can't reintroduce CRLF), and **`scripts/parity-check.sh`** now asserts the two
live hosts are byte-identical. Ran it live:
```
OK   /   1598a01a9810?   OK   /capture 2ada2cfc?   PARITY: in sync
```
**All 10 PRs merged, 0 open, clean slate.** Let me get the mobile dispatch tool, then hand you
the ice-breaker + prioritized list.
```

601. assistant (2026-06-20T07:36:07.671Z)

```
KhelSutra: live + portable, email-on-signup verified, both sites byte-in-sync, 10 PRs merged,
clean checkpoint. Ice-breaker + prioritized next steps are ready in this chat.
```

602. user (2026-06-20T07:36:07.684Z)

```
Mobile push not sent (Remote Control inactive).
```

603. assistant (2026-06-20T07:36:45.718Z)

```
**Mobile dispatch:** the push didn't reach your phone ? **Remote Control is inactive** (that's
the pairing between this session and the Claude mobile app). Connect Remote Control and future
`PushNotification`s land on your phone. For now, here it is in chat. ?
```

? Where this leaves you (clean checkpoint)

khelsutra.guru **live** (Cloudflare Worker + D1, git auto-deploy, CI-gated) · portable Node+SQLite twin **live + browsable** at `http://168.144.126.176:8787` · **both byte-in-sync** (parity-check) · **email-on-registration verified** in your inbox · **PRs #1?#10 merged, 0 open, repo clean** · handoff docs current (`current-state.md` 2026-06-20 block ? `khelsutra-domain-launch.md` + `khelsutra/NEXT_STEPS.md`).

? Prioritized pickup list ? choose your order

#	Item	Value / size
1	**Interactive per-rally selection** ? detect rallies ? UI to pick which go in the highlight (dynamic, not a fixed setting)	? Highest product value · Big (needs design doc + SPA)
2	**Marketing video** ? Grok/Gemini image-gen ? "serious enthusiast" video of the record?reel flow ? embed on site	? High visibility · Medium
3	**Conversational reel-builder backend** ? ship the "ask for a reel over 10s" you teased as *coming*	? Product · Medium
4	**Alpha SPA scaffold** (`web/`, Vite+React+TS) ? the home for #1 and #3	? Enabler · Medium
5	**Productionize the droplet** ? TLS + `vm.khelsutra.guru`, VM-path SMTP email, auto-sync	? Infra · Medium
6	**Site polish** ? Always-HTTPS/HSTS, analytics token, `/capture` link, DMARC, cleanup test rows/orphan Worker	? Quick wins

(1, 3, 4 are one arc ? the interactive reel app. 2 + 6 are independent quick-ish wins.)

? Ice-breaker prompt ? paste this into a fresh session

> Resume the **KhelSutra** brand-site work. Repo `avidullu/khelsutra` (site in `site/`; engine = the badminton-highlight-indexer). Get current: read the **2026-06-20 handoff** at the top of memory `current-state.md` ? `khelsutra-domain-launch.md` + `khelsutra/NEXT_STEPS.md`. State: khelsutra.guru is LIVE (Cloudflare Worker + D1; merge to `master` auto-deploys; CI `npm test`=20 green); a portable Node+SQLite twin runs at the droplet `http://168.144.126.176:8787` (both byte-in-sync via `scripts/parity-check.sh`); each registration emails me + persists to D1. ? The droplet is a shared storage testbed ? don't touch ports 9000/9001/9023 or `/root`. **Today let's pick up: `<#1 interactive per-rally selection>` (or your choice).** Start with a tracked design doc per the repo convention before building.

Swap in whichever number you want first. Everything's checkpointed and honest ? rest easy, and what a build today. ??

604. user (2026-06-22T09:19:25.292Z)

hello. can you sync to remote head and identify any session relevant low hanging fruits which you could complete right now?

605. assistant (2026-06-22T09:20:09.104Z)

Welcome back ? but my context predates a lot of what's happened in this repo (org consolidation, multilingual, per-rally MVP, PRs up to ~#31). So first rule: **sync and re-read the real state, don't trust stale memory.** Let me pull to remote head and survey what's current.

606. user (2026-06-22T09:20:16.992Z)

```
=== remote + branch ===
origin  https://github.com/avidullu/khelsutra.git (fetch)
origin  https://github.com/avidullu/khelsutra.git (push)
branch: master
=== local vs remote ===
```

master...origin/master

```
?? .claude/
?? site/access.db
=== uncommitted? ===
?? .claude/
?? site/access.db
```

```
=== ff-pull ===
Already up to date.
=== recent history ===
e2570e4 docs: archive the Friendly & Multilingual tracker (engineering complete) (#93)
0bef61e docs(friendly-multilingual): close out the chapter ? log #89/#90/#91 + strengthen fitNav
tests (#92)
9f77ddc test(site): browser layout-regression CI guard + auto-collapse the nav to fit every
locale (#91)
4415fe8 feat(site): localize remaining home sections (6 langs) + smooth the responsive header
(#90)
26c9da1 feat(site): localize the hero demo mockup + #how feature block (6 languages) (#89)
7ae9cee feat(web): P2c first-run setup wizard (3 steps from /api/capabilities, 6 languages)
(#88)
cf59fbf feat(site): add es/da/id to the marketing site (parity with the SPA's 6 languages) (#87)
17242a9 feat(site): localize own-model body content to kn/hi (#86)
7b04f02 deploy(kit): align with the live-verified droplet deploy (+ the gemini-default fix)
(#85)
5e5a2ac docs(friendly-multilingual): close out the chapter (§9 DoD + UI/UX + B7 deferred) (#84)
4c0f35c fix(site): UI/UX ? compact language switcher + header overflow fix + localized hero
(#83)
0fae364 feat(web+site): A2 ? localized WhatsApp share for the reel (SPA + site) (#82)
1e52cfd deploy(droplet): tailored alpha kit for the CPU droplet (engine api. side) (#81)
64076a5 test(web): add a coverage floor (CI-gated) + i18n regression tests for the #77 audit
(#80)
8328721 docs(tracker): refresh the §7 milestone prose to merged reality (A6/A7/#77 landed) (#79)
```

607. assistant (2026-06-22T09:20:55.494Z)

Synced ? clean working tree on `master`, up to date with `origin/master` (latest is #93; lots landed since my last context). Two untracked artifacts (`site/access.db`, `.claude/`) and a `/capture` footer link in NEXT_STEPS are the candidate fruits. Let me verify them against the *actual current* code before claiming anything's pending.

608. user (2026-06-22T09:20:59.959Z)

```
=== 1) footer capture link ? still .html? ===
393:href="/capture"
620:href="/capture"
=== 2) .gitignore + are the artifacts ignored? ===
  node_modules/
  dist/
  .env
  .env.*
  *.local
  .DS_Store
  Thumbs.db
  vm-state-backups/
  check-ignore site/access.db: NOT-IGNORED
  check-ignore .claude: NOT-IGNORED
=== 3) what's in .claude/ ? ===
  total 10
  drwxr-xr-x 1 avidu 197609   0 Jun 22 11:12 .
  drwxr-xr-x 1 avidu 197609   0 Jun 22 11:41 ..
  -rw-r--r-- 1 avidu 197609 384 Jun 22 11:12 launch.json
  -rw-r--r-- 1 avidu 197609  76 Jun 21 22:38 settings.local.json
  drwxr-xr-x 1 avidu 197609   0 Jun 21 21:16 worktrees
=== 4) access.db size (runtime artifact) ===
  -rw-r--r-- 1 avidu 197609 12288 Jun 20 20:15 /c/Users/avidu/Projects/khelsutra/site/access.db
```

609. user (2026-06-22T09:23:10.895Z)

This session is being continued from a previous conversation that ran out of context. The summary below covers the earlier portion of the conversation.

Summary:
1. Primary Request and Intent:
   The session spanned multiple distinct requests. Most recent and active intent (the khelsutra brand-site track):
   - Review PR #245 (rally-detection-fixes proposal) thoroughly and reply on it (done; user then closed it).
   - Declutter `current-state.md` memory (sharded into live + archive) and the docs/ folder (classified ? almost nothing archivable).
   - Build a prioritized pending-items list and a landing page for newly-purchased **khelsutra.guru** (3-year registration).
   - Make the landing page "amazing," run visual validations, remove the shuttle logo ("we will do more sports later"), do "something witty with the characters in the name."
   - Position the page around: input match ? segments indexed as rallies ? kill dead time ? conversational reel-building. Then user said move the conversational builder to "where it's going" (lead with index+dead-time which is real today).
   - Set up the domain on Cloudflare (user uses **GoDaddy** registrar), give a browsable link, set up email CTA.
   - Check the landing page into the existing **`khelsutra`** brand repo (NOT the indexer), connect git auto-deploy.
   - Update the **Request Access** CTA to a real **form** serving both academies AND individual enthusiasts, asking "do you have GPU machines for processing" and "how would you store your videos," collecting responses (chose Cloudflare Worker + D1 via AskUserQuestion) plus a free-text message.
   - Iterate with "the same diligence as the badminton highlights indexer repository," add docs + next steps.
   - "design keeping in [an always-on VM] plan" (portability if Cloudflare "acts out").
   - Deploy/test the portable version LIVE on the user's DigitalOcean droplet via SSH; "bake in the harness for droplet testing in the khelsutra repo so that the repo is always compatible for both cloudflare and droplet."
   - Email-on-registration: "everytime someone registers would I get an email? (I want get)."
   - "set it up there and give me a link to browse it, the droplet version and complete the testing."
   - "please merge all the PRs, keep your slate clean and update the handoff work well so that any pending items are clearly listed."
   - Capture two NEW product ideas: (1) a Grok/Gemini image-gen marketing video (themed for serious enthusiasts) embedded on the site; (2) interactive per-rally selection (engine detects rallies ? UI lets user pick which go into the highlight; dynamic, NOT a fixed setting).
   - FINAL message: "please merge the PRs and give me a good 'ice breaker' prompt for the next session to pick these up and give me a prioritized list so I can chose the order! If possible use the dispatch feature of claude to have it delivered to my mobile. verify for yourself that the two websites differ!! Please make sure they are in sync and why did the test not catch this!"

2. Key Technical Concepts:
   - Cloudflare Workers (static assets via `env.ASSETS`), D1 (SQLite-as-a-service), Email Routing (inbound forwarding + outbound `send_email` binding), Workers Builds (git-connected deploy), Pages, custom domains, Web Analytics.
   - Runtime-portable design: Web-standard Request/Response handler + pure logic + injected storage adapter (`d1Store` for Cloudflare, `sqliteStore` for VM) + injected notify adapter (`cfEmailNotify`/`noopNotify`).
   - `node:sqlite` (DatabaseSync, built-in, unflagged in Node 22.13+/24+, confirmed in Node 26).
   - `node:test` test runner; node:http; systemd; ufw; scp/ssh; git autocrlf / `.gitattributes` (CRLF vs LF).
   - GoDaddy ? Cloudflare nameserver migration; Cloudflare Email Routing destination verification.
   - Repo conventions (CLAUDE.md): branch off `origin/master`, ONE PR per work item, no stacking, stage explicit paths, re-verify HEAD before commit, never push to default branch directly, commit trailer `Co-Authored-By: Claude Opus 4.8 <noreply@anthropic.com>`, PR bodies end with `? Generated with [Claude Code](https://claude.com/claude-code)`.
   - Memory/handoff conventions (session-handoff.md, current-state.md sharding with prune rule).

3. Files and Code Sections (khelsutra repo = `C:\Users\avidu\Projects\khelsutra`, the brand repo `avidullu/khelsutra`):
   - **`site/public/index.html`** ? the landing page (segmented Request-Access form, "Khel ? Sutra" play-triangle logo, today-positioning "Every rally, indexed. The dead time, gone.",

"Coming next to the alpha" section). Form posts JSON to `/api/request-access` with mailto fallback.
   - **`site/public/capture.html`** ? printable 1-page capture checklist; served at `/capture`.
   - **`site/public/og.png`** ? 1200×630 OG image (generated by `site/make_og.py`, Pillow + Windows fonts; writes to `public/og.png`).
   - **`site/src/index.js`** ? the Cloudflare Worker. Key exported functions: `validate`, `toRow`, `isBot`, `handleRequest(request, store, notify)`, default `{ fetch }`. The fetch wires: `return handleRequest(request, d1Store(env.DB), cfEmailNotify(env.NOTIFY, env.NOTIFY_FROM, env.NOTIFY_TO));`. handleRequest: validate?honeypot?toRow?`await store.insertAccessRequest(r)`?`try { await notify?.notify(r); } catch {}`?`json({ok:true})`.
   - **`site/src/store.js`** ? storage adapters. Exports `INSERT_SQL`, `insertArgs`, `d1Store(DB)` (uses `DB.prepare(INSERT_SQL).bind(...insertArgs(r)).run()`), `sqliteStore(db)` (uses `db.prepare(INSERT_SQL); stmt.run(...insertArgs(r))`).
   - **`site/src/notify.js`** ? `buildSubmissionEmail(r)` (returns {subject, text}), `cfEmailNotify(binding, from, to)` (dynamic `import("cloudflare:email")` ? `binding.send(new EmailMessage(from, to, raw))`), `noopNotify()`.
   - **`site/server.js`** ? dependency-free Node VM host (node:http + node:sqlite). Serves `public/` + handles `POST /api/request-access` via `handleRequest(await toRequest(req), store, notify)` where `store = sqliteStore(db)`, `notify = noopNotify()`. Env: PORT (default 8080), DB_PATH. Uses `import.meta.dirname`.
   - **`site/src/index.test.js`** ? 12 Cloudflare-logic tests + 3 notify tests (validate, isBot, toRow, handleRequest+store, d1Store, sqliteStore real round-trip, fetch routing, 405, buildSubmissionEmail, notify-called/best-effort, no-notify-on-bot/bad).
   - **`site/src/server.test.js`** ? VM-host integration harness: spawns real `server.js` (free port + temp DB), round-trips via fetch, verifies SQLite persistence. 5 tests.
   - **`site/wrangler.toml`** ? `name = "hidden-king-khelsutra"` (must match the git-connected Worker), `main = "src/index.js"`, `[assets] directory = "./public" binding = "ASSETS"`, `[[d1_databases]] binding = "DB" database_name = "khelsutra-access" database_id = "cd02c6bf-4bd6-4a33-a333-ced86e36b511"`, `[[send_email]] name = "NOTIFY" destination_address = "avi.dullu@gmail.com"`, `[vars] NOTIFY_FROM = "noreply@khelsutra.guru" NOTIFY_TO = "avi.dullu@gmail.com"`.
   - **`site/schema.sql`** ? `access_requests` table (id, created_at, segment, name, email, city, org_name, courts, has_gpu, storage, message, consent, country, user_agent).
   - **`site/package.json`** ? scripts: `"test": "node --test src/index.test.js src/server.test.js"`, `test:cf`, `test:vm`, `deploy`, `dev`, `og`.
   - **`.github/workflows/test.yml`** ? runs `npm test` on push to master + PRs, Node 24, working-directory `site`.
   - **`site/scripts/droplet-smoke.sh`** ? live scp+boot+round-trip against a VM (SSH_TARGET env). Executable.
   - **`site/scripts/parity-check.sh`** ? asserts CF and VM serve byte-identical content (md5 of `/` and `/capture`). Executable.
   - **`.gitattributes`** (NEW, PR #10) ? `* text=auto eol=lf` + binary rules for png/jpg/ico/woff.
   - **`docs/REQUEST_ACCESS_FORM.md`**, **`docs/PORTABILITY.md`**, **`NEXT_STEPS.md`** (repo docs, updated to shipped state + new ideas).
   - **Memory files** (`C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\`): `khelsutra-domain-launch.md` (full durable khelsutra record, repeatedly appended), `current-state.md` (has a 2026-06-20 KHELSUTRA handoff block prepended at top), `current-state-archive.md`, `MEMORY.md`, `working-agreement-checkpoints.md`.

4. Errors and fixes:
   - **wrangler.toml name mismatch:** git-connect attached to existing Worker `hidden-king-khelsutra` but wrangler.toml said `khelsutra-site` ? deploys missed the live site (khelsutra.guru 404 on /api). Fixed by PR #5 renaming `name = "hidden-king-khelsutra"`. NO domain cutover needed (good case).
   - **Preview served wrong dir:** copied updated index.html to `artifacts/site_deploy/` but the preview server (port 8099) served `artifacts/landing/`. Fixed by copying to `artifacts/landing/index.html`.
   - **Screenshot tool flakiness:** preview_screenshot timed out after cross-page navigation; worked around with DOM checks (preview_eval/inspect) and curl.
   - **`node --test src/` aggregated to "tests 1":** directory mode showed misleading count; switched to explicit file list `node --test src/index.test.js src/server.test.js` ? clear 17?20 count.
   - **`curl -w '%{content_type}'/'%{size_download}'` gave "(43) bad argument"** on the Windows curl build; worked around using `-i | head -1` and `-o /dev/null -w '%{http_code}'`.

- **WebFetch 403 on khelsutra.guru** (Cloudflare bot-block) ? used `curl` with a browser User-Agent.
- **CRITICAL ? line-ending drift (user-reported):** CF (LF, md5 1598a01a9810) vs droplet (CRLF, md5 649a5574b9e6); user said "verify for yourself that the two websites differ!! ... why did the test not catch this!" Diagnosis: droplet scp'd from Windows working copy (autocrlf?CRLF), CF from git (LF); content byte-identical after stripping CR. Fixed: normalized droplet to LF (now md5-match), added `.gitattributes` (force LF) + `site/scripts/parity-check.sh` (PR #10). Explained the tests check code/behavior per host, not byte-parity between live deployments, and CRLF/LF doesn't affect functionality or `id="af"` checks.

5. Problem Solving:
Shipped khelsutra.guru end-to-end: domain (GoDaddy?Cloudflare nameservers), landing page, segmented form ? D1, email-on-registration (verified live in inbox via Gmail MCP search showing the structured summary), git auto-deploy, CI. Proved runtime portability live on the DigitalOcean droplet (Node+SQLite at http://168.144.126.176:8787, persistent systemd service). Resolved the line-ending drift permanently. All 10 PRs merged, repo clean (0 open PRs, only master branch). Cleaned droplet test rows; CF D1 still has test rows `deploy-test@`/`email-test@` (noted for user to delete).

6. All user messages:
- "Hello, We have been collaborating on this project and we got a contributor sending us https://github.com/avidullu/badminton-highlight-indexer/pull/245 .. can you do a thorough review and very carefully assess the changes being proposed and reply on the PR accordingly and tell me whether we should merge and fund the project being proposed?"
- "Feel free to post comments on the PR and also push back if needed and appreciate when deserved"
- "So between you and me, I should just ask to close this PR?"
- "Done. Closed it. Now can you tell me why is current-state.md so complex? Can you not shard it or simplify it?"
- "we can do that later for sure and appreciate the house cleaning! Can you help me cleanup the docs/ folder and identify all the docs which are 'partially or completed' projects/ideas/single efforts which we can move to docs/archives now?"
- "There is too much clutter and I want to only maintain active items and want to get a prioritized list of pending items (UI or desktop application, Setting up a domain now that i have bought khelsutra.guru for 3 years! and its a very dumb page so I want to add some good html there etc.)"
- "I want to see the list of incomplete/partial docs so I can take them to completion / thanks for understanding my excitement and pleae make this an amzing page and run vidual validations or gemini calls if you need! / I will be back with you to do godaddy login and do it tonightCofounder! Lets go LIVE!!"
- "Can you refine this to show a feature that 'we can take an input match and give you segments of the video indexed as rallies and killing all the "dead time", you can arrange and stictch them using a conversational input form like give me a reel of all videos which lasted over 10 seconds etc.' .. Do you see the PMF for this use case so maybe make that the index upon?"
- "can you load the website tab in this window?"
- "and oh I verified the email working when my sister clicked on request access / should we link the github with cloudflare so that the deployments are quick and safe?"
- "yes, LGTM! And after its ive you will checkin the PR?"
- "BTW we are also planning on owning an always on VM to host the website eventually if cloudflare acts out, so design keeping in that plan"
- "cool so these were my questions, now I am ready to follow your steps and deploy!"
- "Yes, and I have an up and running Digital ocean droplet where another session is building a 'CPU and Storage shm', I can give you ssh access and want you to drive this full work and testing live as well. Happy to do whateveris needed so you can unblock and go your way"
- "can you not give an ssh key which I can add to droplet and you can ssh into it and drive all this?"
- "168.144.126.176 root@ you can log and verify"
- "set it up there and give me a link to browse it, the droplet version and complete the testing / BTW everytime someone registers would I get an email ? (I want get) / please merge all the PRs, keep your slate clean and update the handoff work well so that any pending items are clearly listed because your context is getting full and I would like to resume this amazing work in a new session once you have brought this to a clean checkpointed and doc updated state for a follow up to pick up"
- "if you can drop an AI that we should use a grok/gemini image generation to make a video

(maybe themed towards serious enthusiasts) and embed a video in the website highlighting the lentire process and also we need to add the ability to 'be able to operate on individual rallies and decide which ones to pick for highlighting' implying the tool will detect rallies and should show to the user in a way that user can pick and chose the ones to select for highlighting and it should always not be from a fixed setting." (the two NEW product ideas)

- "please merge the PRs for these work and put in a place for the enxt to pick up where you leave / please merge the PRs and give me a good 'ice breaker' prompt for the next session to pick these up and give me a prioritized list so I can chose the order! If possible use the dispatch featuer of claude to have it delivered to my mobile / verify for yourself that the two websites differ!! Please make sure they are in sync and why did the test not catch this!"

Security/constraints stated by user/context (preserve): avoid the droplet's other-session storage testbed (MinIO ports 9000/9001, GCS emulator port 9023, `/root/do_gcs_*`); SSH key `~/.ssh/khelsutra_droplet` is revocable (delete its line from droplet `~/.ssh/authorized_keys`); do not paste private keys in chat; paid LLMs never a training data source; MIT/Apache/BSD-only licensing; form/lead data is contact-only, never a training source; user email is avi.dullu@gmail.com.

7. Pending Tasks (captured in NEXT_STEPS.md + the current-state.md 2026-06-20 handoff):

- Provide an **ice-breaker prompt** for the next session and a **prioritized list** (user wants to choose order) ? NOT yet delivered in text.
- **Marketing video** ? Grok/Gemini image-gen ? short video themed for serious enthusiasts showing record?reel ? embed on the site.
- **Interactive per-rally selection** (big feature) ? engine detects rallies ? UI to pick which go into the highlight (dynamic, not a fixed setting); needs a design doc + alpha SPA (`web/`).
- Conversational reel-builder backend (the "ask for a reel over 10s" feature, currently teased as "coming next").
- Productionize droplet (TLS + subdomain `vm.khelsutra.guru` via Caddy; VM-path SMTP email since it's currently `noopNotify`; auto-sync mechanism).
- Site polish: Always-Use-HTTPS/HSTS toggle, paste Web Analytics token into `site/public/index.html`, change footer `/capture.html`?`/capture`, add DMARC TXT, delete any orphan `khelsutra-site` Worker, clean CF D1 test rows (`deploy-test@`, `email-test@`).
- Revoke the droplet SSH key when done.

8. Current Work:

Immediately before the summary request, I was completing the FINAL user message's asks. I had: verified the two sites differed (line endings only, CRLF droplet vs LF Cloudflare), fixed it (normalized droplet to LF; both now md5 `1598a01a9810`), shipped PR #10 (`.gitattributes` + `site/scripts/parity-check.sh`) and merged it, ran `parity-check.sh` live confirming "PARITY: in sync" for `/` and `/capture`, and confirmed 0 open PRs / only master branch. I then loaded the `PushNotification` tool and sent: "KhelSutra: live + portable, email-on-signup verified, both sites byte-in-sync, 10 PRs merged, clean checkpoint. Ice-breaker + prioritized next steps are ready in this chat." The tool returned "Mobile push not sent (Remote Control inactive)." I had NOT yet written out the ice-breaker prompt or the prioritized list in plain text to the user.

9. Optional Next Step:

Deliver, as plain text (no tools), the two remaining items from the user's final message: a ready-to-paste **ice-breaker prompt** for the next session AND a **prioritized list** of next-session tasks so the user can choose the order. Also note that the mobile dispatch could not be delivered (Remote Control inactive) per the PushNotification result. Direct quote of the outstanding request: "give me a good 'ice breaker' prompt for the next session to pick these up and give me a prioritized list so I can chose the order!" The prioritized list should draw from the Pending Tasks (interactive per-rally selection; marketing video; conversational reel-builder backend; productionize droplet; site polish; alpha SPA scaffold) and the ice-breaker should point the next session to read the 2026-06-20 handoff at the top of `current-state.md` ? `[[khelsutra-domain-launch]]` + `khelsutra/NEXT_STEPS.md`, note the repo is `avidullu/khelsutra` (live at khelsutra.guru + droplet http://168.144.126.176:8787, both byte-in-sync, CI-gated), and reiterate the constraint to avoid the droplet's MinIO/GCS storage testbed (ports 9000/9001/9023, `/root/do_gcs_*`).

If you need specific details from before compaction (like exact code snippets, error messages, or content you generated), read the full transcript at: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675.jsonl
Continue the conversation from where it left off without asking the user any further questions. Resume directly ? do not acknowledge the summary, do not recap what was happening, do not

preface with "I'll continue" or similar. Pick up the last task as if the break never happened.

610. assistant (2026-06-22T09:24:43.203Z)

Good ? the verify-first rule already paid off:

- **`/capture` footer link** (NEXT_STEPS line 44, "currently `/capture.html` ? change to `/capture`") is **already done** ? both links in `index.html` (lines 393, 620) are `/capture`. The doc is stale, not the code.
- Most other "Queued" items are **owner-account actions**, not code I can complete: Cloudflare HTTPS/HSTS toggle, Web Analytics token, DMARC TXT, deleting the orphan Worker / D1 test rows, revoking the SSH key.

That leaves two genuine **repo-hygiene fruits I can finish right now**, both verified against current state:

1. **Untracked local artifacts aren't ignored** ? `site/access.db` (12 KB runtime SQLite, created by `server.js`) and `.claude/` (machine-specific: `launch.json` with absolute `C:/...` paths, `settings.local.json`, `worktrees/`) both show as `??` and `git check-ignore` confirms **NOT-IGNORED** ? one stray `git add -A` commits them. `.gitignore` should cover them.
2. **NEXT_STEPS.md is stale** ? header still says "PRs #1?#8 merged / npm test (20 green)" (we're at #93, CI now runs site+web+layout), and the `/capture` item is done.

Let me confirm nothing's tracked under those patterns, then branch off `origin/master`.

611. user (2026-06-22T09:24:46.781Z)

```
=== tracked .db files (should be none) ===
  [end]
=== tracked under .claude/ (should be none) ===
  [end]
=== branch off origin/master ===
branch 'chore/repo-hygiene-ignores' set up to track 'origin/master'.
Switched to a new branch 'chore/repo-hygiene-ignores'
HEAD: e2570e4 docs: archive the Friendly & Multilingual tracker (engineering complete) (#93)
branch: chore/repo-hygiene-ignores
```

612. assistant (2026-06-22T09:24:50.724Z)

Clean ? nothing tracked under those patterns, and I'm on a fresh branch off `origin/master` (#93). Let me read the full `.gitignore` then make both edits.

613. user (2026-06-22T09:24:51.366Z)

```
1       node_modules/
2       dist/
3       .env
4       .env.*
5       *.local
6       .DS_Store
7       Thumbs.db
8       vm-state-backups/
9
```

614. user (2026-06-22T09:24:55.727Z)

The file C:\Users\avidu\Projects\khelsutra\.gitignore has been updated successfully. (file state is current in your context ? no need to Read it back)

615. user (2026-06-22T09:25:13.199Z)

The file C:\Users\avidu\Projects\khelsutra\NEXT_STEPS.md has been updated successfully. (file state is current in your context ? no need to Read it back)

616. user (2026-06-22T09:25:17.673Z)

```
The file C:\Users\avidu\Projects\khelsutra\NEXT_STEPS.md has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

617. assistant (2026-06-22T09:25:21.095Z)

Now verify the site suite stays green (it shouldn't be affected, but the merge gate requires it), then commit explicit paths and open the PR.

618. user (2026-06-22T09:25:27.495Z)

```
? [VM host] GET /api/request-access -> 405 (1.1325ms)
? [VM host · N8] a Kannada browser (Accept-Language: kn) gets the kn-rendered page (2.7632ms)
? [VM host · N8] ?lang= overrides the browser language (2.0768ms)
? [VM host · N8] the persisted cookie beats Accept-Language (returning user keeps their choice)
(1.633ms)
? [VM host · N8] unsupported/absent language falls back to en (never a locale we can't serve)
(3.3689ms)
? [VM host · N8] non-localized assets are NOT rewritten (only the 3 marketing pages negotiate)
(1.7548ms)
? [VM host] serves the shared catalogs the shim fetches (/locales/*/common.json as JSON)
(2.9452ms)
? submit shows an obvious in-flight state (disabled + aria-busy + 'Sending?') (345.4406ms)
? on failure the button is restored ? never stuck on a frozen 'Sending?' (66.0443ms)
? after sign-up BOTH videos appear and a ? closes the reveal (52.6453ms)
? A2: after sign-up the reveal offers a WhatsApp share of the reel (wa.me, safe, localizable)
(46.9898ms)
? the auto-close HOLDS while the reel is playing (no mid-watch cut-off) (42.8889ms)
? tests 86
? suites 0
? pass 86
? fail 0
? cancelled 0
? skipped 0
? todo 0
? duration_ms 2944.9983
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

619. assistant (2026-06-22T09:25:33.382Z)

86/86 green (the suite grew from 20 ? 86 since my last context ? another reason the "20 green" claim was stale). Re-verifying HEAD, then committing explicit paths and pushing.

620. user (2026-06-22T09:25:36.436Z)

```
Exit code 1
=== HEAD/branch re-verify ===
chore/repo-hygiene-ignores
e2570e4 docs: archive the Friendly & Multilingual tracker (engineering complete) (#93)
=== only my commits vs origin/master (should be empty pre-commit) ===
=== stage explicit paths ===
M  .gitignore
M  NEXT_STEPS.md
?? .claude/
?? site/access.db
=== confirm artifacts now ignored (not staged) ===
```

621. assistant (2026-06-22T09:25:50.853Z)

Caught a real bug ? and this is exactly why I verified: `git check-ignore` returned exit 1 (NOT ignored), and the artifacts still show `??`. **`.gitignore` doesn't support trailing inline comments** ? my `*.db   # ...` was parsed as a literal pattern (with spaces), matching nothing. Comments must be on their own line. Fixing:

622. user (2026-06-22T09:25:52.429Z)

```
The file C:\Users\avidu\Projects\khelsutra\.gitignore has been updated successfully. (file state
is current in your context ? no need to Read it back)
```

623. user (2026-06-22T09:25:57.728Z)

```
=== now ignored? (expect both paths echoed = ignored) ===
site/access.db
.claude
  -> IGNORED ?
=== status (artifacts should be GONE from untracked) ===
M  .gitignore
M  NEXT_STEPS.md
```

624. assistant (2026-06-22T09:26:05.073Z)

```
Now both are properly ignored and the untracked artifacts no longer appear. Committing and
pushing.
```

625. user (2026-06-22T09:26:14.646Z)

```
[chore/repo-hygiene-ignores 6db63c5] chore(repo): ignore local runtime artifacts + correct stale
NEXT_STEPS claims
 2 files changed, 11 insertions(+), 4 deletions(-)
=== commit landed; only-mine vs origin/master ===
6db63c5 chore(repo): ignore local runtime artifacts + correct stale NEXT_STEPS claims
=== push ===
remote:      https://github.com/avidullu/khelsutra/pull/new/chore/repo-hygiene-ignores
remote:
branch 'chore/repo-hygiene-ignores' set up to track 'origin/chore/repo-hygiene-ignores'.
To https://github.com/avidullu/khelsutra.git
 * [new branch]      chore/repo-hygiene-ignores -> chore/repo-hygiene-ignores
```

626. assistant (2026-06-22T09:26:17.679Z)

```
Pushed cleanly (only my commit ahead of `origin/master`). Opening the PR.
```

627. user (2026-06-22T09:26:28.525Z)

```
https://github.com/avidullu/khelsutra/pull/94
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

628. assistant (2026-06-22T09:26:37.644Z)

```
PR **[#94](https://github.com/avidullu/khelsutra/pull/94)** opened. Watching CI ? it'll run site
+ web + layout jobs (the layout job installs Chromium, ~2?3 min). I'll merge once it's green,
per the established khelsutra flow.
```

629. user (2026-06-22T09:27:42.593Z)

```
site    pass    15s
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333264
Refreshing checks status every 20 seconds. Press Ctrl+C to quit.

Workers Builds: hidden-king-khelsutra   pass    0
https://dash.cloudflare.com/213c12cf60de9041dd8a3521aa206472/workers/services/view/hidden-king-
khelsutra/production/builds/17b1d667-eceb-42fd-a2f1-bd1a6b394eb5
layout  pending 0
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333244
web     pending 0
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333290
site    pass    15s
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333264
```

```
Workers Builds: hidden-king-khelsutra   pass    0
https://dash.cloudflare.com/213c12cf60de9041dd8a3521aa206472/workers/services/view/hidden-king-
khelsutra/production/builds/17b1d667-eceb-42fd-a2f1-bd1a6b394eb5
layout  pass    58s
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333244
site    pass    15s
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333264
web     pass    48s
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333290
Workers Builds: hidden-king-khelsutra   pass    0
https://dash.cloudflare.com/213c12cf60de9041dd8a3521aa206472/workers/services/view/hidden-king-
khelsutra/production/builds/17b1d667-eceb-42fd-a2f1-bd1a6b394eb5
layout  pass    58s
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333244
site    pass    15s
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333264
web     pass    48s
https://github.com/avidullu/khelsutra/actions/runs/27942890042/job/82680333290
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

630. assistant (2026-06-22T09:27:49.815Z)

```
All four checks green ? `site`, `web`, `layout`, and the Cloudflare Workers Builds preview.
Merging (squash, per the repo's one-commit-per-PR convention) and deleting the branch.
```

631. user (2026-06-22T09:28:01.397Z)

```
Updating e2570e4..50b2bd7
Fast-forward
 .gitignore    | 6 ++++++
 NEXT_STEPS.md | 9 +++++----
 2 files changed, 11 insertions(+), 4 deletions(-)
=== back to master, sync ===
HEAD: 50b2bd7 chore(repo): ignore local runtime artifacts + correct stale NEXT_STEPS claims
(#94)
=== open PRs / branches (expect clean) ===
    chore/web-coverage-floor
    docs/multilingual-expansion-proposal
    feat/own-model-kn-hi
    feat/p2c-setup-wizard
    feat/site-es-da-id
    fix/kn-hi-spa-console-native-review
  * master
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

632. assistant (2026-06-22T09:28:29.672Z)

```
Done ? clean slate. Here's the session-relevant sweep:
```

? What I synced and completed

```
**Synced to remote head** ? `khelsutra` was already up to date with `origin/master` at **#93**
(lots had landed since my last context: org consolidation, 6-language localization, per-rally
MVP, suite grown 20 ? **86 tests**). Now at **#94** after this fix.

**Two repo-hygiene fruits found, verified, and shipped** ? PR
**[#94](https://github.com/avidullu/khelsutra/pull/94)** merged, all CI green (site + web +
layout + Cloudflare preview), master fast-forwarded, **0 open PRs**:
1. **`.gitignore`** now covers `*.db` (the runtime `site/access.db`) and `.claude/` (machine-
specific config with absolute `C:/?` paths) ? both were untracked and *not* ignored, so a stray
`git add -A` could have committed them.
2. **`NEXT_STEPS.md`** honesty pass ? dropped a queued item and fixed a stale status line
(below).
```

? Verify-first paid off twice

- The NEXT_STEPS "**change `/capture.html` ? `/capture`**" item was **already done** ? both footer links in `index.html` are `/capture`. Doc was stale, not the code. (Also corrected "PRs #1?#8 / 20 green" ? through #93 / 86 green / 3 CI jobs.)
- My first `.gitignore` edit used **trailing inline comments** (`*.db  # ?`) ? `.gitignore` has no such syntax, so it silently matched nothing. `git check-ignore` caught it (exit 1 = not ignored) before commit; fixed to own-line comments and re-verified.

? What I did *not* touch (owner-account, not code)

The rest of the "site polish" queue needs your Cloudflare/DNS/host access, not a code change: Always-HTTPS + HSTS toggle, Web Analytics token, DMARC TXT, delete the orphan `khelsutra-site` Worker, clean the D1 test rows, droplet TLS/SMTP, and revoking the droplet SSH key.

? One optional follow-up

You have **6 stale local branches** (all merged & deleted on remote): `chore/web-coverage-floor`, `docs/multilingual-expansion-proposal`, `feat/own-model-kn-hi`, `feat/p2c-setup-wizard`, `feat/site-es-da-id`, `fix/kn-hi-spa-console-native-review`. Safe to prune for a tidy local workspace ? want me to delete them?

633. user (2026-06-22T09:30:03.557Z)

yes, can you please cleanup the stale branches and do a quick doc sweep to ensure the docs are all updated and that should be enough for the session! Thanks and appreciate the hard work!

634. assistant (2026-06-22T09:30:42.721Z)

Appreciate it ? let me do this right. First, gather the facts: confirm each stale branch maps to a **merged** PR (squash-merges mean `git branch -d` can't auto-detect them, so I'll verify via GitHub before any delete), and enumerate the full doc set for the sweep.

635. user (2026-06-22T09:30:46.948Z)

```
=== local branches (non-master) ===
chore/web-coverage-floor
docs/multilingual-expansion-proposal
feat/own-model-kn-hi
feat/p2c-setup-wizard
feat/site-es-da-id
fix/kn-hi-spa-console-native-review
=== merged-PR head branches (authoritative) ===
chore/annotator-org-links       #23     2026-06-20T14:54:26Z
chore/org-move-done     #27     2026-06-20T16:04:40Z
chore/repo-hygiene-ignores      #94     2026-06-22T09:27:56Z
chore/site-crawl-hygiene        #71     2026-06-21T17:50:54Z
chore/web-coverage-floor        #80     2026-06-21T19:48:02Z
chore/wire-d1-id        #4      2026-06-20T06:15:54Z
ci/regression-guards    #12     2026-06-20T08:52:16Z
claude/badminton-academy-flyer-w3obpr   #40     2026-06-21T04:39:55Z
deploy/alpha-cloudflared-kit    #65     2026-06-21T15:46:19Z
deploy/droplet-alpha-kit        #81     2026-06-21T22:13:02Z
deploy/kit-live-verified        #85     2026-06-22T01:19:40Z
docs/annotate-and-train #15     2026-06-20T12:58:39Z
docs/archive-friendly-multilingual      #93     2026-06-22T06:11:23Z
docs/correct-bucket-ingest-staleness    #55     2026-06-21T09:35:45Z
docs/dry-run-cuj-plan   #53     2026-06-21T08:24:43Z
docs/dry-run-pathb-done #57     2026-06-21T13:34:58Z
docs/dry-run-result     #54     2026-06-21T08:25:54Z
docs/friendly-multilingual-project      #41     2026-06-21T06:20:39Z
docs/friendly-multilingual-spa-done     #49     2026-06-21T07:19:29Z
docs/hygiene-refresh    #59     2026-06-21T14:26:56Z
docs/local-appliance-architecture       #13     2026-06-20T12:58:35Z
docs/multilingual-expansion-proposal    #58     2026-06-21T14:28:44Z
docs/multilingual-header-closeout       #92     2026-06-22T05:55:20Z
docs/next-steps-mvp-complete    #32     2026-06-20T17:35:58Z
```

```
docs/next-steps-p3      #29     2026-06-20T16:53:35Z
docs/next-steps-wrap    #9      2026-06-20T07:29:55Z
docs/per-rally-selection        #11     2026-06-20T08:49:18Z
docs/t11-retention-done #69     2026-06-21T17:39:20Z
docs/tracker-prose-freshness    #79     2026-06-21T19:34:22Z
feat/a1-plain-language-copy     #60     2026-06-21T14:39:54Z
feat/a2-whatsapp-share  #82     2026-06-21T22:19:05Z
feat/a3-recording-l10n  #74     2026-06-21T18:56:55Z
feat/a4-mobile-touch-targets    #61     2026-06-21T15:05:27Z
feat/a5-friendly-build-errors   #62     2026-06-21T15:22:13Z
feat/a6-access-plain-language   #75     2026-06-21T19:04:22Z
feat/a7-capture-diagram #76     2026-06-21T19:15:02Z
feat/a8-status-chips    #63     2026-06-21T15:37:27Z
feat/a9-tap-filters     #68     2026-06-21T17:26:10Z
feat/access-worker-d1   #3      2026-06-20T06:00:00Z
feat/b3-site-i18n       #70     2026-06-21T18:03:56Z
feat/demo-reel-video    #17     2026-06-20T14:05:02Z
feat/droplet-test-harness       #7      2026-06-20T07:13:56Z
feat/email-notify       #8      2026-06-20T07:22:56Z
feat/flyer-nav-ownmodel-links   #44     2026-06-21T06:31:19Z
feat/flyer-ui-screenshots       #52     2026-06-21T07:38:26Z
feat/link-academy-flyers        #43     2026-06-21T06:10:06Z
feat/lw2-locale-passthrough     #73     2026-06-21T18:25:33Z
feat/marketing-landing-page     #1      2026-06-20T03:07:26Z
feat/ml-closeout        #84     2026-06-22T01:05:10Z
feat/mlx-phaseA-es-da-id        #67     2026-06-21T17:08:18Z
feat/own-model-backlink #26     2026-06-20T15:25:35Z
feat/own-model-kn-hi    #86     2026-06-22T01:51:22Z
feat/own-model-page     #22     2026-06-20T14:48:53Z
feat/own-model-reel     #24     2026-06-20T15:03:49Z
feat/own-model-tool-shots       #25     2026-06-20T15:17:38Z
feat/p2c-setup-wizard   #88     2026-06-22T03:57:00Z
feat/per-rally-p2-grid  #21     2026-06-20T14:42:08Z
feat/per-rally-p3-query-bar     #28     2026-06-20T16:51:26Z
feat/per-rally-p4-footer        #30     2026-06-20T17:04:46Z
feat/per-rally-p5-build #31     2026-06-20T17:15:53Z
feat/portable-host      #6      2026-06-20T06:42:15Z
feat/portable-vm-provisioning   #16     2026-06-20T13:10:10Z
feat/reel-on-signup     #18     2026-06-20T14:15:01Z
feat/reel-random-start  #20     2026-06-20T14:28:03Z
feat/request-access-form        #2      2026-06-20T05:53:06Z
feat/site-es-da-id      #87     2026-06-22T03:37:47Z
feat/site-gated-tool-demo       #34     2026-06-20T17:57:14Z
feat/site-hero-mockup-i18n      #89     2026-06-22T04:34:28Z
feat/site-home-sections-i18n    #90     2026-06-22T05:16:09Z
feat/site-layout-ci-guard       #91     2026-06-22T05:45:08Z
feat/site-recording-gist        #39     2026-06-21T02:54:49Z
feat/site-reveal-watchable      #37     2026-06-21T02:40:50Z
feat/site-serve-mp4     #35     2026-06-20T17:59:28Z
feat/site-submit-loading        #36     2026-06-21T02:10:44Z
feat/spa-bucket-ingest  #56     2026-06-21T10:15:55Z
feat/t4-in-ui-upload    #66     2026-06-21T16:47:52Z
feat/ui-header-cleanup  #83     2026-06-22T00:59:37Z
feat/web-duration-first #33     2026-06-20T17:44:41Z
feat/web-hi-catalog     #48     2026-06-21T07:13:50Z
feat/web-i18n-extract-strings   #45     2026-06-21T06:55:52Z
feat/web-i18n-scaffold  #42     2026-06-21T06:42:14Z
feat/web-kn-catalog     #47     2026-06-21T07:08:21Z
feat/web-noto-fonts     #46     2026-06-21T07:01:24Z
feat/web-offline-guard  #50     2026-06-21T07:26:50Z
feat/web-query-sample-prompt    #51     2026-06-21T07:30:39Z
feat/web-spa-scaffold   #14     2026-06-20T09:55:04Z
fix/archive-lf-crossplatform    #19     2026-06-20T14:19:48Z
fix/deploy-alpha-runbook-correctness    #78     2026-06-21T19:34:18Z
fix/kn-hi-spa-console-native-review     #77     2026-06-21T19:16:09Z
```

```
fix/line-endings-parity #10     2026-06-20T07:35:21Z
fix/wrangler-name       #5      2026-06-20T06:28:15Z
spa/api-base-env        #64     2026-06-21T15:41:03Z
test/site-ux-behavioral #38     2026-06-21T02:50:37Z
=== any OPEN PRs? ===
[end-open]
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

636. assistant (2026-06-22T09:30:53.563Z)

All 6 stale branches confirmed as **merged** PRs (#80, #58, #86, #88, #87, #77), and **0 open PRs**. Deleting them (force-delete needed since squash-merges leave `git branch -d` unable to see the merge), and enumerating the doc set in parallel.

637. user (2026-06-22T09:30:58.297Z)

```
=== delete 6 merged local branches ===
Deleted branch chore/web-coverage-floor (was ba5e14a).
Deleted branch docs/multilingual-expansion-proposal (was 3297596).
Deleted branch feat/own-model-kn-hi (was afc4c9a).
Deleted branch feat/p2c-setup-wizard (was 9827259).
Deleted branch feat/site-es-da-id (was c0ef790).
Deleted branch fix/kn-hi-spa-console-native-review (was 50ccbba).
=== remaining branches (expect only master) ===
  * master
=== prune stale remote-tracking refs ===
 * [pruned] origin/feat/site-hero-mockup-i18n
 * [pruned] origin/feat/site-home-sections-i18n
 * [pruned] origin/feat/site-layout-ci-guard
[pruned]
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

638. assistant (2026-06-22T09:31:01.165Z)

Branches clean ? only `master` remains. Now the doc sweep. Let me enumerate the full doc set so I know what the workflow must cover.

639. user (2026-06-22T09:31:04.862Z)

```
=== all tracked markdown docs (excl node_modules) ===
NEXT_STEPS.md
README.md
deploy/DEPLOY_ALPHA.md
deploy/droplet/README.md
docs/ANNOTATE_AND_TRAIN.md
docs/DRY_RUN_CUJ.md
docs/LOCAL_APPLIANCE_ARCHITECTURE.md
docs/MULTILINGUAL_EXPANSION.md
docs/PER_RALLY_SELECTION.md
docs/PORTABILITY.md
docs/README.md
docs/REQUEST_ACCESS_FORM.md
docs/archives/FRIENDLY_AND_MULTILINGUAL.md
site/README.md
web/README.md

=== docs/ tree ===
docs/ANNOTATE_AND_TRAIN.md
docs/DRY_RUN_CUJ.md
docs/LOCAL_APPLIANCE_ARCHITECTURE.md
docs/MULTILINGUAL_EXPANSION.md
docs/PER_RALLY_SELECTION.md
docs/PORTABILITY.md
docs/README.md
```

```
docs/REQUEST_ACCESS_FORM.md
docs/archives/FRIENDLY_AND_MULTILINGUAL.md

=== word counts (size sense) ===
   197  docs/archives/FRIENDLY_AND_MULTILINGUAL.md
   196  docs/LOCAL_APPLIANCE_ARCHITECTURE.md
   173  docs/PER_RALLY_SELECTION.md
   150  deploy/DEPLOY_ALPHA.md
   115  docs/MULTILINGUAL_EXPANSION.md
   113  deploy/droplet/README.md
   106  docs/REQUEST_ACCESS_FORM.md
    99  docs/ANNOTATE_AND_TRAIN.md
    93  docs/DRY_RUN_CUJ.md
    80  docs/PORTABILITY.md
    79  NEXT_STEPS.md
    73  site/README.md
    32  docs/README.md
    30  web/README.md
    21  README.md
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

640. assistant (2026-06-22T09:31:57.806Z)

15 docs. Ultracode is on, so I'll run this as a **fan-out audit** ? one verifier per doc that checks every falsifiable claim against the *live* repo (paths, routes, PR numbers, counts, status markers), with the hard rule that a claim is "correct" only with `file:line`/command evidence, else "unverifiable". Then I synthesize and apply the safe fixes myself in one PR.

641. user (2026-06-22T09:32:17.850Z)

```
Workflow launched in background. Task ID: wqjgke5dp
Summary: Audit every KhelSutra doc for staleness against the live repo, returning structured
findings
Transcript dir: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\subagents\workflows\wf_81ded7b8-667
Script file: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-sweep-
wf_81ded7b8-667.js
(Edit this file with Write/Edit and re-invoke Workflow with {scriptPath:
"C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-sweep-
wf_81ded7b8-667.js"} to iterate without resending the script.)
Run ID: wf_81ded7b8-667
To resume after editing the script: Workflow({scriptPath: "C:\Users\avidu\.claude\projects\C--
Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-sweep-
wf_81ded7b8-667.js", resumeFromRunId: "wf_81ded7b8-667"}) ? completed agents return cached
results.

You will be notified when it completes. Use /workflows to watch live progress.
```

642. assistant (2026-06-22T09:32:25.318Z)

`args` arrived as a string, not an array. Simplest robust fix: inline the doc list in the script itself. Re-running.

643. user (2026-06-22T09:32:45.978Z)

```
Workflow launched in background. Task ID: w3dbzrpby
Summary: Audit every KhelSutra doc for staleness against the live repo, returning structured
findings
Transcript dir: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\subagents\workflows\wf_f9255361-69d
Script file: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
```

indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-sweep-wf_f9255361-69d.js
(Edit this file with Write/Edit and re-invoke Workflow with {scriptPath: "C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-sweep-wf_f9255361-69d.js"} to iterate without resending the script.)
Run ID: wf_f9255361-69d
To resume after editing the script: Workflow({scriptPath: "C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-sweep-wf_f9255361-69d.js", resumeFromRunId: "wf_f9255361-69d"}) ? completed agents return cached results.

You will be notified when it completes. Use /workflows to watch live progress.

644. assistant (2026-06-22T09:32:55.976Z)

Doc-sweep workflow is running (15 verifiers in parallel against the live repo). Branch cleanup is already done ? only `master` remains locally, stale remote-tracking refs pruned. I'll synthesize the audit findings and apply the safe doc fixes when it completes.

645. user (2026-06-22T09:32:56.902Z)

waiting on workflow wf_f9255361-69d (15 doc auditors). Branch cleanup complete.

646. assistant (2026-06-22T09:33:00.584Z)

Branch cleanup is finished and the doc audit is in flight ? I'll pick up the findings and apply fixes the moment the 15 verifiers report back.

647. user (2026-06-22T09:42:58.321Z)


```
<task-notification>
<task-id>w3dbzrpby</task-id>
<tool-use-id>toolu_01AxCUnWQdQ2RjTfW8i5KhyU</tool-use-id>
<output-file>C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\w3dbzrpby.output</output-file>
<status>completed</status>
<summary>Dynamic workflow "Audit every KhelSutra doc for staleness against the live repo, returning structured findings" completed</summary>
<result>{"perDoc":[{"doc":"NEXT_STEPS.md","overallStatus":"minor-staleness","summary":"NEXT_STEPS.md is overwhelmingly accurate against master: routes, file paths, the Worker/D1 names, the 6-language localization, the 3-step wizard, the per-rally PR numbers (#11/#14/#21/#28/#30/#31), the appliance/per-rally doc statuses, and every still-pending \"Queued\" item all check out against code and git. The only real staleness is the \"PRs through #93 merged\" count, which lags by one because PR #94 (the very commit that wrote \"through #93\") merged afterward; plus a low-severity GCS-artifact filename mismatch. Not an archive candidate ? it is the live next-steps tracker.","findings":[{"claim":"PRs through #93 merged (master).","docLocation":"NEXT_STEPS.md:13","verdict":"stale","evidence":"git log --oneline -1 = '50b2bd7 chore(repo): ... correct stale NEXT_STEPS claims (#94)'; total 94 commits/PRs on master. Ironically #94 is the commit that updated this line to '#93', then merged itself, so the doc now lags one PR behind master.","suggestedFix":"Update to 'PRs through #94 merged (master).'","severity":"medium"},{"claim":"Where we are (2026-06-20): ... khelsutra.guru is LIVE and fully shipped (landing + /capture checklist + OG image + segmented Request-Access form).","docLocation":"NEXT_STEPS.md:5-14","verdict":"correct","evidence":"site/server.js:98 maps /capture-&gt;capture.html; site/public/og.png + capture.html tracked; index.html:191-193 textareas + lines 537-581 academy/individual/gpu/storage fields. Header date 2026-06-20 is slightly behind the latest work (PR #94 merged 2026-06-22) but the substance holds.","suggestedFix":"Optionally bump the 'Where we are' date to 2026-06-22 to match the latest merged work.","severity":"low"},{"claim":"On Cloudflare (Worker hidden-king-khelsutra + D1 khelsutra-access): the form persists to D1 and emails a summary via an Email Routing send binding.","docLocation":"NEXT_STEPS.md:7-9","verdict":"correct","evidence":"site/wrangler.toml: name = \"hidden-king-khelsutra\"; [[d1_databases]] database_name = \"khelsutra-access\"; [[send_email]] name = \"NOTIFY\". site/src/index.js:80 handles POST /api/request-access -&gt;
```

insert into D1 binding DB.","suggestedFix":"","severity":"low"},{"claim":"CI gates every PR ? three jobs: full site npm test, web/ SPA build + coverage floor, and a real-browser layout check.","docLocation":"NEXT_STEPS.md:9-12","verdict":"correct","evidence":"site/package.json scripts include test (node --test), test:layout (layout.browser.mjs); web has a coverage floor (#80). PR #94 message confirms 'three jobs: site npm test = 86 green, web SPA build+coverage, real-browser layout'. Non-browser site suite ran green (61 tests in the 4 non-layout files).","suggestedFix":"","severity":"low"},{"claim":"The form also runs on a dependency-free Node + SQLite host (site/server.js), live on the DO droplet -> http://168.144.126.176:8787 (systemd khelsutra-site).","docLocation":"NEXT_STEPS.md:11-13","verdict":"correct","evidence":"site/server.js exists, uses node:sqlite (header comment line 2). IP 168.144.126.176 + port 8787 + 'khelsutra-site' all appear in docs/LOCAL_APPLIANCE_ARCHITECTURE.md:48 ('/srv/khelsutra-site, node :8787') and deploy refs. Live-status is external/unverifiable from the repo, but the host code and identifiers are correct.","suggestedFix":"","severity":"low"},{"claim":"Docs: docs/REQUEST_ACCESS_FORM.md, docs/PORTABILITY.md.","docLocation":"NEXT_STEPS.md:13-14","verdict":"correct","evidence":"git ls-files: both docs/REQUEST_ACCESS_FORM.md and docs/PORTABILITY.md are tracked.","suggestedFix":"","severity":"low"},{"claim":"DONE ? Friendly & Multilingual (ARCHIVED 2026-06-22): localized to all 6 languages (en/kn/hi/es/da/id); full record docs/archives/FRIENDLY_AND_MULTILINGUAL.md.","docLocation":"NEXT_STEPS.md:17-21","verdict":"correct","evidence":"docs/archives/FRIENDLY_AND_MULTILINGUAL.md is tracked; git log shows PR #93 'docs: archive the Friendly & Multilingual tracker'. web/src/i18n/index.ts:25 SUPPORTED_LOCALES = [en,kn,hi,es,da,id]; PR #87 added es/da/id to the marketing site.","suggestedFix":"","severity":"low"},{"claim":"Interactive per-rally selection MVP/M0 COMPLETE; Done: P0 design (#11), P1 scaffold (#14), P2 grid+hook (#21), P3 parseQuery query bar (#28), P4 SelectionFooter (#30), P5 real load+download (#31). Tracked doc docs/PER_RALLY_SELECTION.md (IN PROGRESS).","docLocation":"NEXT_STEPS.md:32-41","verdict":"correct","evidence":"git log: cb74ce5 #11 'tracked design ... RallyPicker'; aac7e0c #14 'P1 scaffold'; 62598e7 #21 'P2 selectable rally grid + hook'; d537d19 #28 'P3 deterministic parseQuery() query bar'; aff8de8 #30 'P4 SelectionFooter + min_shot_count warning'; 8bf200d #31 'P5 load real rallies + Build->compile->preview->download'. docs/PER_RALLY_SELECTION.md status header = 'IN PROGRESS ? MVP/M0 (P1-P5) COMPLETE'.","suggestedFix":"","severity":"low"},{"claim":"P2c first-run setup wizard ? 3-step SetupWizard from /api/capabilities, 6 languages.","docLocation":"NEXT_STEPS.md (implied via track) and verified for the merged #88","verdict":"correct","evidence":"web/src/components/SetupWizard.tsx:22 'const STEPS = 3'; line 1 comment 'Three quick steps'; line 3 'Localized (en/kn/hi/es/da/id)'; web/src/api/client.ts:139 GET /capabilities. git: 7ae9cee #88.","suggestedFix":"","severity":"low"},{"claim":"Single-box appliance ? Tracked design docs/LOCAL_APPLIANCE_ARCHITECTURE.md (DRAFT, owner-gated).","docLocation":"NEXT_STEPS.md:52-58","verdict":"correct","evidence":"docs/LOCAL_APPLIANCE_ARCHITECTURE.md status header: 'Status: DRAFT ? owner-gated (auth choice, retention, cost ceiling, GPU host, P5 bridge)'.","suggestedFix":"","severity":"low"},{"claim":"Queued: Always-Use-HTTPS + HSTS (Cloudflare) and Cloudflare Web Analytics token paste into site/public/index.html ? both still pending.","docLocation":"NEXT_STEPS.md:44-45","verdict":"correct","evidence":"No Strict-Transport-Security/HSTS header found in site/*.js or wrangler.toml. site/public/index.html:629 still has data-cf-beacon token 'REPLACE_WITH_CF_TOKEN' (commented-out block). Both correctly listed as not-yet-done.","suggestedFix":"","severity":"low"},{"claim":"Droplet -> production: add VM-path SMTP so the Node host also emails (currently noopNotify).","docLocation":"NEXT_STEPS.md:46-47","verdict":"correct","evidence":"site/server.js:14 imports noopNotify; line 24 'const notify = noopNotify(); // VM host: no email yet (add SMTP later)'.","suggestedFix":"","severity":"low"},{"claim":"Org consolidation DONE; khelsutra (this site repo) transfer is an owner step still PENDING (unchecked).","docLocation":"NEXT_STEPS.md:65-73","verdict":"correct","evidence":"git remote -v shows origin = https://github.com/avidullu/khelsutra.git (NOT under Khelsutra org) ? consistent with the '? owner step pending' marker. Other repos already moved: docs/DRY_RUN_CUJ.md:49 links github.com/Khelsutra/badminton-highlight-indexer#289.","suggestedFix":"","severity":"low"},{"claim":"A real-UI CUJ screencast exists -> promo_cuj.mp4/.gif in gs://khelsutra-rally-corpus/promo/.","docLocation":"NEXT_STEPS.md:29-31","verdict":"unverifiable","evidence":"No GCS access to confirm the object. The companion doc docs/DRY_RUN_CUJ.md:85 names the actual artifacts 'khelsutra-bucket-cuj-screencast.mp4' and ':60 khelsutra-i18n-screencast.mp4' in the same gs://khelsutra-rally-corpus/promo/ prefix; 'promo_cuj.mp4' appears nowhere else in the repo, suggesting a possible filename mismatch.","suggestedFix":"Reconcile the filename with docs/DRY_RUN_CUJ.md (likely 'khelsutra-bucket-cuj-screencast.mp4') or confirm 'promo_cuj.mp4' is

a real object in the bucket.","severity":"low"},{"claim":"Control plane (control-plane/) at beta.","docLocation":"NEXT_STEPS.md:62","verdict":"correct","evidence":"control-plane/ directory does not exist in the repo (test -d failed), but this line sits under 'Bigger arcs (design-first)', i.e. a forward-looking/aspirational item, not
... (truncated 120095 chars, full result in C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\w3dbzrpby.output)</result>
<usage><agent_count>15</agent_count><subagent_tokens>999898</subagent_tokens><tool_uses>302</tool_uses><duration_ms>612293</duration_ms></usage>
</task-notification>

648. assistant (2026-06-22T09:43:03.324Z)

Audit complete (15 docs, 302 verification tool-uses). The result was truncated ? let me read the full output to get every stale finding.

649. user (2026-06-22T09:43:07.119Z)

```
Exit code 1
=== file size ===
149267 C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\w3dbzrpby.output
=== STALE findings only (doc / loc / severity / claim -> fix) ===
total stale: 0

=== per-doc overallStatus ===
[eval]:14
for(const d of o.perDoc){console.log("  "+d.overallStatus.padEnd(22)+d.doc);}
                 ^

TypeError: o.perDoc is not iterable
    at [eval]:14:18
    at runScriptInThisContext (node:internal/vm:219:10)
    at node:internal/process/execution:483:12
    at [eval]-wrapper:6:24
    at runScriptInContext (node:internal/process/execution:481:60)
    at evalFunction (node:internal/process/execution:315:30)
    at evalTypeScript (node:internal/process/execution:327:3)
    at node:internal/main/eval_string:71:3

Node.js v26.1.0
```

650. assistant (2026-06-22T09:43:14.679Z)

Output file isn't the bare return object ? let me inspect its actual structure.

651. user (2026-06-22T09:43:17.964Z)

```
first 200 chars: "{\n  \"summary\": \"Audit every KhelSutra doc for staleness against the live repo, returning structured findings\",\n  \"agentCount\": 15,\n  \"logs\": [\n    \"audited 15/15 docs; 44 stale findings; archive-candi"
top-level keys: [ 'summary', 'agentCount', 'logs', 'result' ]
types: {
  summary: 'string',
  agentCount: 'number',
  logs: 'array[1]',
  result: 'object'
}
```

652. assistant (2026-06-22T09:43:21.421Z)

The real payload is under `result`, and the log says **44 stale findings**. Let me extract them all, grouped by severity.

653. user (2026-06-22T09:43:26.159Z)

```
LOG: audited 15/15 docs; 44 stale findings; archive-candidates: none

=== per-doc status ===
  minor-staleness       NEXT_STEPS.md
  significant-staleness README.md
  minor-staleness       deploy/DEPLOY_ALPHA.md
  fresh                 deploy/droplet/README.md
  significant-staleness docs/ANNOTATE_AND_TRAIN.md
  minor-staleness       docs/DRY_RUN_CUJ.md
  significant-staleness docs/LOCAL_APPLIANCE_ARCHITECTURE.md
  minor-staleness       docs/MULTILINGUAL_EXPANSION.md
  significant-staleness docs/PER_RALLY_SELECTION.md
  minor-staleness       docs/PORTABILITY.md
  minor-staleness       docs/README.md
  minor-staleness       docs/REQUEST_ACCESS_FORM.md
  fresh                 docs/archives/FRIENDLY_AND_MULTILINGUAL.md
  minor-staleness       site/README.md
  minor-staleness       web/README.md

############ HIGH (11) ############

? README.md  @README.md:9
  CLAIM: Status: pre-alpha scaffold. SPA lands in `web/` per execution-plan milestone M1.
  EVID : web/src/ has 51 tracked source files incl. web/src/App.tsx; `git log --oneline -- web/`
= 31 commits; merged PRs show the SPA is feature-complete: P2c first-run wizard (#88), in-UI
file upload T4 (#66), per-rally selection MVP P0?P5 (#11/#14/#21/#28/#30/#31), 6-language
catalogs (#67). NEXT_STEPS.md:7 states 'khelsutra.guru is LIVE and fully shipped'.
UI_ALPHA_EXECUTION_PLAN.md:36 defines M1 only as the initial 'Vite+React+TS+Tailwind scaffold',
which is long past.
  FIX  : Replace with a current status, e.g.: 'Status: alpha ? the web/ SPA is built and
multilingual (6 languages), with the per-rally selection MVP and first-run setup wizard shipped;
khelsutra.guru / app.khelsutra.guru are live (Cloudflare Pages + cloudflared tunnel + Access).
See NEXT_STEPS.md.'

? docs/ANNOTATE_AND_TRAIN.md  @docs/ANNOTATE_AND_TRAIN.md:76
  CLAIM: P5 'Capture + commit real screenshots ... replace placeholders' ? Status ? (Todo),
Gated Owner.
  EVID : PR #25 (1906c82) 'show the real annotator screenshots' committed
site/public/img/annotate-enable.png (133712 bytes) and annotate-tool.png (153133 bytes); own-
model.html:111-112 references them as the real VLC-UI shots (no people/footage). The annotator-
UI half of P5 is done and merged, not Todo.
  FIX  : Mark P5 ? for the annotator-UI screenshots (cite #25); if footage screenshots with
consented people are still pending, split that into its own row so the done part isn't hidden
under a blanket Todo.

? docs/ANNOTATE_AND_TRAIN.md  @docs/ANNOTATE_AND_TRAIN.md:5,71,91
  CLAIM: Doc is 'index[ed] in docs/README.md' (line 5), §7 P0 'indexed in docs/README.md' (?,
#15), and DoD 'doc indexed in docs/README.md'.
  EVID : grep -ni 'annotate|own.model|train.your|ANNOTATE_AND' docs/README.md returns NOT FOUND;
the docs/README.md 'Tracked work' list (lines 1-30) does not mention ANNOTATE_AND_TRAIN.md or
own-model. The claimed index entry does not exist.
  FIX  : Either add an ANNOTATE_AND_TRAIN.md entry to docs/README.md's 'Tracked work' list, or
correct P0/DoD to reflect that it is not yet indexed.

? docs/LOCAL_APPLIANCE_ARCHITECTURE.md  @docs/LOCAL_APPLIANCE_ARCHITECTURE.md:49,113,164
  CLAIM: §3 lists `GET /api/capabilities` under 'Absent (net-new): ... `[verified ? grep returns
none]`'; §4.6 Reuse map lists it under 'Net-new'; §9 T7 (capabilities + health + banner) is ?
Todo.
  EVID : engine origin/master backend/main.py:340 `@app.get("/api/capabilities")` (and
main.py:307 `/api/health`, main.py:334 `/api/version`) all exist on master. git grep
'/api/capabilities' origin/master returns a hit.
  FIX  : Move `/api/capabilities` out of 'Absent (net-new)' and out of §4.6 'Net-new'; mark the
engine half of T7 (capabilities+health) as shipped (?/?), leaving only the SPA banner pending.
```

? docs/LOCAL_APPLIANCE_ARCHITECTURE.md @docs/LOCAL_APPLIANCE_ARCHITECTURE.md:14,30,103,128
  CLAIM: TL;DR + Risk table: 'the engine has zero app-layer auth and CORS `*` `[verified backend/main.py:42]`' ? structural safety only because the engine binds 127.0.0.1.
  EVID : engine origin/master backend/main.py:82 uses `allow_origins=_cors_origins_from_env()` (default ['*'] but env-lockable) with `allow_credentials=False`; main.py:109 adds TrustedHostMiddleware (DNS-rebinding/Host allowlist); backend/api/auth.py + config/startup.py enforce_startup_checks + M9 `Cf-Access-Authenticated-User-Email`/`owner_id` now exist on master. CORS is no longer a hardcoded bare `*`, and an app-layer edge-auth integration now exists.
  FIX : Reword to: CORS default is `*` but now env-configurable (allow_credentials off), and the engine now ships a Host-header guard + an edge-auth integration (Cf-Access trust) + exposure-safety boot posture ? so 'zero app-layer auth' is no longer literally true. Re-anchor the line (main.py:42 no longer points at CORS).

? docs/LOCAL_APPLIANCE_ARCHITECTURE.md @docs/LOCAL_APPLIANCE_ARCHITECTURE.md:175,182,183
  CLAIM: §9 alpha-deploy block: 'the actual stand-up (tunnel + Pages + CF Access + key rotation) is the **owner** step'; §11 DoD box 1 (CUJ-1 e2e on droplet behind gate) and box 2 (raw :8000 refused, tested) unchecked.
  EVID : deploy/droplet/README.md:100-106 '? Live-verified on the droplet (2026-06-22)': engine booted on 168.144.126.176; suite 1313 passed; boot-posture fail-fast, Host-guard 400?200, path-jail, T4 upload?jail, and edge-auth 401-gate all exercised live; full DNS?Cloudflare?Access?tunnel?engine chain confirmed; app.khelsutra.guru up. Only Gemini-key rotation + a browser OTP login remain.
  FIX : Update §9 to note the alpha stand-up is LIVE (2026-06-22, app.khelsutra.guru, posture test exercised) with only key rotation + browser-OTP login outstanding; tick/qualify the corresponding DoD boxes.

? docs/PER_RALLY_SELECTION.md @docs/PER_RALLY_SELECTION.md:133 (P7 tracker row P6), :83-84 (§4.3 table)
  CLAIM: P6 (engine repo) ? GET /api/health + offline banner; /api/compile returns download_url ? Status ? (Todo, not built); §4.3 table marks GET /api/health as 'new' and the download_url addition as 'modify' (not yet done).
  EVID : backend/main.py:307 `@app.get("/api/health")` is fully implemented; backend/main.py compile (~line 769) returns `{"status":"success","filepath":out_path,"download_url":download_url}` where download_url = `/api/highlights/{video_id}/{out_name}`. Both P6 items are shipped in the engine.
  FIX : Flip P6 to ? (or split: health=?, download_url=?) and update §4.3 status from 'new'/'modify' to 'exists'. Note in the doc that the engine now returns download_url so the SPA no longer needs to hand-build the URL (though it still can).

? docs/PER_RALLY_SELECTION.md @docs/PER_RALLY_SELECTION.md:12, :44, :73, :116
  CLAIM: Compile returns a server-local filepath, NOT a URL, so the SPA constructs this download URL itself (repeated as a core caveat in §0 TL;DR, §3 Foundation, §4.2 data flow, D11/risk table, changelog).
  EVID : backend/main.py compile endpoint now returns `download_url` alongside `filepath`: `download_url = f"/api/highlights/{quote(req.video_id)}/{quote(out_name)}"; return {"status":"success","filepath":out_path,"download_url":download_url}`. The 'filepath not a URL' premise that the whole 'SPA must build the URL' design rests on is no longer the full picture.
  FIX : Update the caveat to: compile returns BOTH a server-local filepath and a ready-to-use download_url (/api/highlights/...). The SPA may use download_url directly; the hand-built-URL fallback is now optional, not required.

? docs/PER_RALLY_SELECTION.md @docs/PER_RALLY_SELECTION.md:135 (P8 row), :86 (§4.3)
  CLAIM: P8 ? GET /api/videos + identity scoping ? Status ? owner, ? (not built); §4.3 marks 'GET /api/videos' as 'new', noting 'only get_video(id) exists, database.py:418'.
  EVID : backend/main.py:374 `@app.get("/api/videos")` ? `db_instance.list_videos(limit=limit)`; backend/infrastructure/database.py:545 `def list_videos(...)`. The endpoint and the DB enumerate both exist (added per PACKAGING_PLAN.md P1). The doc's 'only get_video(id) exists' claim and the 'new' status are wrong; identity scoping may still be pending but the listing endpoint shipped.
  FIX : Mark the GET /api/videos endpoint as 'exists' in §4.3; in P8 note that the videos-list endpoint already shipped and only the identity/owner-scoping layer remains gated.

? docs/README.md @docs/README.md:16-18
  CLAIM: MULTILINGUAL_EXPANSION.md is a 'proposal to add 7 more locales (th/id/da/es/zh-Hans/ja/ko) ... DRAFT ? owner-gated (approval to start), [#58].'
  EVID : docs/MULTILINGUAL_EXPANSION.md:3 status header reads 'IN PROGRESS ? Phase A (es/da/id)

shipped (#67); Phases B (Thai) / C (CJK) owner-gated'. Git confirms PR #67 'feat(web): MULTILINGUAL Phase A ? Spanish/Danish/Indonesian SPA catalogs'. site/public/locales/ now contains es, da, id directories and web/src/i18n/index.ts:25 lists SUPPORTED_LOCALES = en,kn,hi,es,da,id. So es/da/id are no longer 'to add', and the doc is not a not-yet-started DRAFT.
  FIX : Update to: '**`MULTILINGUAL_EXPANSION.md`** ? `IN PROGRESS`: Phase A (es/da/id) shipped ([#67]); Thai (B) / CJK (C) remain owner-gated. Adds locales beyond the en/kn/hi foundation; peer of the archived `FRIENDLY_AND_MULTILINGUAL.md`.'

? web/README.md @web/README.md:30
  CLAIM: `src/App.tsx` ? the console shell (screens land in M2+).
  EVID : web/src/App.tsx is a complete screen, not a shell: it renders RallyGrid, QueryBar, TapFilters, SelectionFooter, LanguageSwitcher, IngestPanel, SetupWizard, CompileResult (imports App.tsx:3-16) and orchestrates listRallies/compileReel against the engine (App.tsx:80,150). Last touched by PR #88 'P2c first-run setup wizard' (git log -- web/src/App.tsx), long past M2. The README was last modified at PR #14 (git log -- web/README.md), so this line predates the entire screen build-out.
  FIX : Replace with something like: "`src/App.tsx` ? the operator console: per-rally selection (RallyGrid/QueryBar/TapFilters/SelectionFooter), in-UI ingest (IngestPanel), the P2c first-run wizard (SetupWizard) and reel compile, in 6 languages."

############ MEDIUM (17) ############

? NEXT_STEPS.md @NEXT_STEPS.md:13
  CLAIM: PRs through #93 merged (master).
  EVID : git log --oneline -1 = '50b2bd7 chore(repo): ... correct stale NEXT_STEPS claims (#94)'; total 94 commits/PRs on master. Ironically #94 is the commit that updated this line to '#93', then merged itself, so the doc now lags one PR behind master.
  FIX : Update to 'PRs through #94 merged (master).'

? README.md @README.md:4
  CLAIM: the web app (alpha SPA) now, the hosted control plane at beta
  EVID : This tagline pairs with line 9's 'pre-alpha scaffold', which is contradicted by the shipped, live SPA (see web/ evidence above and NEXT_STEPS.md:7). The 'alpha SPA now' framing is directionally right but undercut by the stale 'pre-alpha scaffold' status it sits beside.
  FIX : Keep 'web app (alpha SPA) now' but align the rest of the intro/status so it no longer describes the SPA as not-yet-landed.

? README.md @README.md:17
  CLAIM: `control-plane/` ? (beta) Cloud Run FastAPI control plane
  EVID : No control-plane/ directory exists: `ls control-plane` -> 'No such file or directory'; `git ls-files control-plane/` returns nothing. Top-level tracked dirs are deploy/, docs/, locales/, site/, web/ only.
  FIX : Mark it as a future/planned directory that does not yet exist (e.g. move it out of the 'Structure' list into a 'Planned (beta)' note), so a reader doesn't expect the path to be present.

? deploy/DEPLOY_ALPHA.md @deploy/DEPLOY_ALPHA.md:3
  CLAIM: header stale deploy live
  EVID : droplet/README.md:100 live-verified 2026-06-22
  FIX : set header to LIVE

? docs/ANNOTATE_AND_TRAIN.md @docs/ANNOTATE_AND_TRAIN.md:3
  CLAIM: Status: DRAFT ? owner-gated on scope (header, line 3); Last updated: 2026-06-20.
  EVID : git log --follow shows the marketing-page work merged: P0 doc (#15, 6b62ebf), own-model page P1-P4 (#22, b1110e1), org link (#23, 87520b8). Per §7, P0-P4 are all marked Done ? yet the header still reads DRAFT. Lifecycle (line 4) is DRAFT->IN PROGRESS->DONE; with P0-P4 shipped this is at least IN PROGRESS.
  FIX : Flip Status to IN PROGRESS (only owner-gated P5-P9 remain) and bump Last updated; or, if the owner considers the marketing surface complete, mark DONE and archive per the lifecycle note.

? docs/DRY_RUN_CUJ.md @docs/DRY_RUN_CUJ.md:3
  CLAIM: Status: DRAFT ? owner-gated · ... Goal for this session.
  EVID : docs/README.md:4 marks this doc '? DELIVERED (2026-06-21)'; DoD §6 boxes 1-3 are all

checked [x] (DRY_RUN_CUJ.md:47-49) and the work shipped (#53/#55/#56/#57). A DRAFT/'goal for this session' header no longer matches the completed, delivered state.
  FIX : Flip the header to Status: DONE / DELIVERED (2026-06-21) and drop 'Goal for this session'; note the only open box (droplet, §6 box 4) is owner-parked. Consider archiving once that box closes.

? docs/LOCAL_APPLIANCE_ARCHITECTURE.md @docs/LOCAL_APPLIANCE_ARCHITECTURE.md:45,96
  CLAIM: §3 / §4.3: 'Native WASB is in-flight on branch `feat/native-wasb-runner`: `native_wasb_runner.py` exists `[verified]`'.
  EVID : engine origin/master now contains backend/pipeline/detectors/native_wasb_runner.py and tests/test_native_wasb_runner.py (git ls-tree origin/master). The runner has merged to master; the `feat/native-wasb-runner` branch still exists but the file is no longer branch-only. (trajectory_hybrid._build_native_runner still raises NotImplementedError at trajectory_hybrid.py:61 ? that sub-claim is still correct.)
  FIX : Change 'in-flight on branch feat/native-wasb-runner' to 'native_wasb_runner.py is on master; trajectory_hybrid still refuses, fusion builds it; owner byte-parity gate still pending.'

? docs/LOCAL_APPLIANCE_ARCHITECTURE.md @docs/LOCAL_APPLIANCE_ARCHITECTURE.md:3
  CLAIM: Status header: 'Status: DRAFT ... Last updated: 2026-06-21' with §9 the source of truth.
  EVID : khelsutra git log shows merges through #94 (master HEAD 50b2bd7) and the alpha went live 2026-06-22, both after the doc's 2026-06-21 'last updated'. Multiple §9 rows and engine facts have moved since (T7 engine half, alpha deploy, native_wasb on master).
  FIX : Bump 'Last updated' and consider promoting DRAFT?IN PROGRESS; refresh §9 rows to current reality.

? docs/MULTILINGUAL_EXPANSION.md @docs/MULTILINGUAL_EXPANSION.md:87,16,68
  CLAIM: EX3 ('site/ marketing parity for the 7') is ? and 'Blocked on B3'; and §1/§4/§6 say site localization 'depends on FRIENDLY B3 landing first ? until then this is SPA-only.'
  EVID : FRIENDLY B3 shipped: git log 1b1fa7b 'feat(web+site): B3 ? trilingual marketing site via shared t()/data-i18n shim (#70)'. Site es/da/id parity also shipped: cf59fbf '#87'. git ls-files shows site/public/locales/{es,da,id}/common.json present. The dependency is satisfied and the Phase-A trio is no longer SPA-only.
  FIX : Update EX3 to reflect that FRIENDLY B3 landed (#70) and the Phase-A es/da/id site parity shipped (#87) ? mark the es/da/id portion ? (site has those catalogs) with only th/zh/ja/ko site parity remaining; drop the 'until then this is SPA-only' caveat for es/da/id in §1/§4/§6.

? docs/PER_RALLY_SELECTION.md @docs/PER_RALLY_SELECTION.md:130-132
  CLAIM: P3/P4/P5 tracker rows show PR = '(this PR)' as the merge reference.
  EVID : git log: P3 merged as #28 (d537d19), P4 as #30 (aff8de8), P5 as #31 (8bf200d). The '(this PR)' placeholders were never backfilled, so a reader cannot find the actual PRs.
  FIX : Replace '(this PR)' with [#28], [#30], [#31] for P3/P4/P5 respectively.

? docs/PER_RALLY_SELECTION.md @docs/PER_RALLY_SELECTION.md:38-46, :163 (§10 anchors)
  CLAIM: Engine line anchors in [verified ...] tags: config at backend/main.py:121-123; jobs :274-293; highlights download :307-328; query :331-340; compile + floor :342-396; play_segments database.py:124-135; dead correction cols :166-168; no list-videos :418; query read :427-475.
  EVID : Live engine: /api/config is at main.py:297 (not 121-123); /api/jobs/{job_id} at :637 (not 274-293); /api/highlights at :670 (not 307-328); /api/query at :694 (not 331-340); /api/compile at :705 (not 342-396); play_segments CREATE TABLE at database.py:186 (not 124-135); get_video at :536 (not 418). The endpoints/contracts still EXIST and behave as described, but every cited line number has drifted because the engine ships PRs independently.
  FIX : Re-run the anchors against current engine HEAD and update line numbers, or replace brittle line numbers with symbol names (e.g. main.py::get_config, database.py::list_videos). The sibling DRY_RUN_CUJ.md §8 already warns that engine-surface claims drift between sessions.

? docs/PORTABILITY.md @docs/PORTABILITY.md:64
  CLAIM: VM deploy runbook: Node 22+ only ? no npm install (node:sqlite is built in).
  EVID : Internally inconsistent with the doc's own provisioning section (line 40: Node >=24 'so node:sqlite is unflagged') and the repo: provision-vm.sh NODE_MAJOR default 24 + 'must be >= 24', .github/workflows/test.yml:17 node-version 24, server.js:1 also says 'Node 22+'. node:sqlite was flagged/experimental before Node ~22.5 and only stabilized later; a reader running the runbook on Node 22 may hit a flagged/unavailable node:sqlite. The authoritative

requirement in code is >=24.
  FIX : Change 'Node 22+ only' to 'Node 24+ only (node:sqlite is unflagged at >=24)' in both the runbook (line 64) and server.js:1, to match provision-vm.sh / CI.

? docs/README.md @docs/README.md:3-27
  CLAIM: The index lists every tracked doc in this repo (DRY_RUN_CUJ, FRIENDLY_AND_MULTILINGUAL, MULTILINGUAL_EXPANSION, LOCAL_APPLIANCE_ARCHITECTURE, PER_RALLY_SELECTION, REQUEST_ACCESS_FORM, PORTABILITY).
  EVID : git ls-files docs/ also returns docs/ANNOTATE_AND_TRAIN.md, which is NOT referenced anywhere in README.md. That doc's own header (docs/ANNOTATE_AND_TRAIN.md:5) states its tracking anchor is 'index in docs/README.md', so it expects to be listed here.
  FIX : Add a 'Tracked work' entry, e.g.: '**`ANNOTATE_AND_TRAIN.md`** ? annotate-your-footage ? train-your-own-model USP pages (rally-annotator + engine `worker train`). `DRAFT` ? owner-gated on scope.'

? docs/REQUEST_ACCESS_FORM.md @docs/REQUEST_ACCESS_FORM.md:58
  CLAIM: Frontend lives in `site/index.html` (the header/hero buttons scroll to #join; #join becomes the form).
  EVID : `ls site/index.html` -> exit 2 (no such file). The frontend is git-tracked as site/public/index.html (git ls-files), which contains the #join section (public/index.html:524) and the href="#join" CTAs (lines 269,282,284). The file was moved by PR #3 (git show f699847 --stat: `site/{ => public}/index.html`) on 2026-06-20 ? the same day as the doc's 'Last updated', so the path reference was never updated.
  FIX : Change `site/index.html` to `site/public/index.html` in §4 (the #join + form behavior it describes is otherwise correct).

? site/README.md @site/README.md:8-9
  CLAIM: Layout: src/ contains index.js (the Worker) and index.test.js (unit tests for the form logic) ? implies that is the whole src/ surface.
  EVID : git ls-files site/src/ returns 11 files: assets.test.js, deploy-config.test.js, i18n.shim.test.js, index.js, index.test.js, layout.browser.mjs, locales-drift.test.js, notify.js, server.js absent (it's at site/server.js), store.js, server.test.js, ux.dom.test.js. The README names only index.js + index.test.js and omits store.js, notify.js, and the 6 added test files.
  FIX : List the real src/ surface: src/index.js (shared handler/validation), src/store.js (D1 + SQLite stores), src/notify.js (email/no-op notifiers), and note the test files (index/server/assets/deploy-config/i18n.shim/locales-drift/ux.dom + layout.browser.mjs).

? site/README.md @site/README.md:1-12
  CLAIM: The site is just 'a static landing page + /capture checklist plus a small Worker' ? Cloudflare-Worker-only; server.js / VM host / droplet are not mentioned at all.
  EVID : site/server.js:1-14 is an 'Always-on VM host for the KhelSutra site ... the portability fallback if Cloudflare acts out ? see docs/PORTABILITY.md', sharing the same handleRequest from src/index.js. docs/PORTABILITY.md exists. The site test run shows a full '[VM host]' suite (POST persists to SQLite, serves /own-model, flyers, locales). The README has zero mention of server.js, the VM host, the droplet, or PORTABILITY.md.
  FIX : Add a short section: the same handler runs in two places ? the Cloudflare Worker (src/index.js, this README's focus) and an always-on VM host (server.js, the portability fallback; see docs/PORTABILITY.md), used by the alpha droplet.

? web/README.md @web/README.md:24-30
  CLAIM: Layout section lists only src/api/client.ts, src/api/types.ts, src/lib/log.ts, src/App.tsx as the SPA's structure.
  EVID : git ls-files web/ shows the SPA has grown a full tree the Layout omits: src/components/ (9 non-test components incl. RallyGrid, QueryBar, SetupWizard, IngestPanel, LanguageSwitcher), src/hooks/ (useIngestJob, useRallySelection, useUpload), src/i18n/ (6-locale catalog: en/kn/hi/es/da/id per grep of src/i18n/index.ts), and most of src/lib/ (query, selection, floor, format, errors, errorText). The README froze at the #14 scaffold (git log -- web/README.md) when only these four files existed.
  FIX : Add the components/, hooks/, i18n/ (6 languages) and remaining lib/ entries so the Layout reflects the shipped SPA, not just the P1 scaffold.

############ LOW (16) ############

? deploy/DEPLOY_ALPHA.md @deploy/DEPLOY_ALPHA.md:66

  CLAIM: host list too narrow
  EVID : engine.env.example adds loopback
  FIX : add localhost

? docs/ANNOTATE_AND_TRAIN.md @docs/ANNOTATE_AND_TRAIN.md:53,64
  CLAIM: SPA today: web/src/App.tsx is one static screen, no router (web/package.json); a correction screen needs react-router-dom.
  EVID : grep on web/package.json: no react-router dep (claim 'no router' still TRUE). But App.tsx is no longer 'one static screen' ? it now imports and conditionally renders SetupWizard (web/src/App.tsx:11,65,242, P2c first-run wizard from PR #88). The 'one static screen' framing predates the wizard.
  FIX : Reword to 'no router; App.tsx renders a first-run SetupWizard + the ingest panel' so the SPA surface description matches the post-#88 code.

? docs/ANNOTATE_AND_TRAIN.md @docs/ANNOTATE_AND_TRAIN.md:6,84
  CLAIM: Relation note: peer of LOCAL_APPLIANCE_ARCHITECTURE.md (the latter lands with PR #13).
  EVID : git ls-files confirms docs/LOCAL_APPLIANCE_ARCHITECTURE.md already exists in the repo; it has long since landed, so the forward-looking 'lands with PR #13' / 'cross-link once PR #13 merges' phrasing is outdated.
  FIX : Drop the 'lands with PR #13' future-tense framing ? the doc exists now; add the cross-link.

? docs/DRY_RUN_CUJ.md @docs/DRY_RUN_CUJ.md:11
  CLAIM: The SPA is 'Fully en/kn/hi (the switcher / ?lang=)' and the multilingual workflow is en/kn/hi.
  EVID : web/src/i18n/index.ts:25 ? `export const SUPPORTED_LOCALES = ["en", "kn", "hi", "es", "da", "id"]`; es/da/id were added to the SPA in PR #87 (cf59fbf 'add es/da/id to the marketing site (parity with the SPA's 6 languages)'). The SPA now supports 6 languages, not 3.
  FIX : Update 'en/kn/hi' to 'en/kn/hi/es/da/id (6 languages)' where it describes the current switcher; the Hindi/Kannada run results in §7 stay accurate as historical record.

? docs/MULTILINGUAL_EXPANSION.md @docs/MULTILINGUAL_EXPANSION.md:108
  CLAIM: §10 References cites internal 'docs/UI_PROPOSAL.md'.
  EVID : git ls-files | grep -i UI_PROPOSAL returns nothing (exit 1) ? no such file exists anywhere in the repo.
  FIX : Remove the docs/UI_PROPOSAL.md reference or repoint it to the correct doc (the engine repo has UI docs; this repo does not contain UI_PROPOSAL.md).

? docs/MULTILINGUAL_EXPANSION.md @docs/MULTILINGUAL_EXPANSION.md:50
  CLAIM: SUPPORTED_LOCALES entries are at i18n/index.ts:18.
  EVID : web/src/i18n/index.ts: SUPPORTED_LOCALES is declared at line 25 (= ['en','kn','hi','es','da','id']), not line 18 (line 18 is the hiCommon import). Line-number drift only.
  FIX : Change the citation from i18n/index.ts:18 to i18n/index.ts:25.

? docs/MULTILINGUAL_EXPANSION.md @docs/MULTILINGUAL_EXPANSION.md:49
  CLAIM: The <html lang> sync is at i18n/index.ts:59-65.
  EVID : web/src/i18n/index.ts:59-65 is the detection-order block (order/lookupLocalStorage/etc.); the syncDocumentLang() function for <html lang> begins at line 70.
  FIX : Repoint the <html lang> citation to ~i18n/index.ts:70 (syncDocumentLang).

? docs/PER_RALLY_SELECTION.md @docs/PER_RALLY_SELECTION.md:6, :164
  CLAIM: References §0.6 / §10 cite 'REQUEST_ACCESS_FORM.md' (repo-root style) as a peer doc.
  EVID : git ls-files shows the file is at docs/REQUEST_ACCESS_FORM.md, not at repo root. The bare 'REQUEST_ACCESS_FORM.md' reference is path-imprecise (it now lives under docs/).
  FIX : Reference it as docs/REQUEST_ACCESS_FORM.md.

? docs/PORTABILITY.md @docs/PORTABILITY.md:25
  CLAIM: `node --test src/` runs BOTH paths.
  EVID : package.json "test": "node --test" (no src/ arg) is what CI runs (.github/workflows/test.yml:25 `npm test`). The split targets are test:cf / test:vm. `node --test src/` is not the command CI uses; the bare `node --test` discovers and runs all *.test.js.
  FIX : Change to: `npm test` (node --test) runs BOTH paths, and CI runs it on every PR.

? docs/REQUEST_ACCESS_FORM.md @docs/REQUEST_ACCESS_FORM.md:55
  CLAIM: The hidden honeypot field is named `hp` (anti-spam; if filled -> silently drop).
  EVID : site/public/index.html:602 `<input class="hp" type="text" name="company" ...>` ? `hp` is the CSS class, the field NAME is `company`. The Worker honeypot check keys on `company`: site/src/index.js:21 `return !!(p && typeof p.company === "string" && p.company.trim() !== "")`, and the client drops on `f.company` (public/index.html:722).
  FIX : In the §3 table, change the honeypot row's field from `hp` to `company` (note `hp` is the CSS class).

? docs/REQUEST_ACCESS_FORM.md @docs/REQUEST_ACCESS_FORM.md:85
  CLAIM: P3: per-submission email notify is LIVE (per-submission to `hello@`).
  EVID : Email-notify is implemented (site/src/notify.js cfEmailNotify + wrangler.toml `[[send_email]]`), but the destination is avi.dullu@gmail.com, not hello@: wrangler.toml:24 `destination_address = "avi.dullu@gmail.com"`, :28 `NOTIFY_TO = "avi.dullu@gmail.com"`. (The form's contact mailto fallback does use hello@, but submission notifications go to avi.dullu@gmail.com.)
  FIX : Change 'per-submission to `hello@`' to 'per-submission to avi.dullu@gmail.com' to match wrangler.toml NOTIFY_TO.

? docs/REQUEST_ACCESS_FORM.md @docs/REQUEST_ACCESS_FORM.md:66
  CLAIM: Anti-spam: add Cloudflare Turnstile in P2 if spam appears.
  EVID : Internal inconsistency: §4 line 66 says Turnstile lands 'in P2', but the §6 tracker lists Turnstile under P3 (line 85). P2 (the Worker, PR #3) shipped without Turnstile, so §4's 'P2' is the wrong phase label.
  FIX : Change 'in P2' to 'in P3' in §4 to match the tracker.

? site/README.md @site/README.md:7
  CLAIM: public/ ? static assets served as-is (index.html, capture.html, og.png, robots.txt).
  EVID : git ls-files site/public/ also lists own-model.html, flyer.html/flyer-kn.html/flyer-hi.html, i18n.js, img/ (studio/annotate/flyer-qr), video/tool-cuj.mp4, and locales/{en,kn,hi,es,da,id}/common.json ? a 6-language i18n + flyers + SPA-video surface absent from the README's 4-file list.
  FIX : Broaden the public/ description to mention the localized pages + locales/ (6 langs), the flyers, own-model.html, img/, and video/, not just the original four files.

? site/README.md @site/README.md:16
  CLAIM: npm test # node --test src/ (no install needed)
  EVID : site/package.json scripts.test = 'node --test' (no src/ path argument). Running it executes 86 tests across src/ and other dirs (? tests 86, pass 86).
  FIX : Change the annotation to '# node --test (86 tests, no install needed)'.

? site/README.md @site/README.md:8
  CLAIM: src/index.js ? the Worker: serves public/ and handles POST /api/request-access -> D1.
  EVID : site/src/index.js:84 returns env.ASSETS.fetch(request) for non-form paths ? on Cloudflare the static assets are served by the [assets] binding (wrangler.toml:12-14 ASSETS), not by Worker code. The Worker only owns the /api/request-access path; it delegates 'everything else: static assets'.
  FIX : Reword to 'handles POST /api/request-access -> D1 and delegates all other paths to the Cloudflare static-assets binding (ASSETS)'.

? web/README.md @web/README.md:25-26
  CLAIM: `src/api/client.ts` ? typed engine client (`listRallies` / `compileReel` / `highlightUrl` / `getConfig`)
  EVID : All four named exports exist (web/src/api/client.ts:101,111,128,133), so the claim is not wrong ? but the client now also exports processVideo (:85), getJob (:96), getCapabilities (:138), minShotCount (:143) and the chunked-upload family uploadInit/uploadPart/uploadComplete/uploadAbort/uploadFile (:152-245). The four-function summary undersells the current surface.
  FIX : Append "+ processVideo / getJob / getCapabilities / chunked uploadFile" (or note the list is illustrative) so the client description isn't read as exhaustive.

654. assistant (2026-06-22T09:45:49.801Z)

Thorough result: **44 stale findings across 13 docs** (2 came back `fresh`, 0 archive candidates). The pattern is clear ? the docs froze at their authoring PR while the SPA, engine

API, and localization shipped past them (e.g. `README.md` still says "pre-alpha scaffold", `web/README.md` calls `App.tsx` a "shell", several docs mark engine endpoints `/api/health` · `/api/capabilities` · `/api/videos` as "net-new/Todo" when they're live on the engine's master).

I'll fix these as a **parallel fan-out** (one fixer per stale doc, each re-verifying against the live repo ? including the sibling engine repo for cross-repo claims ? before editing), then review the whole diff myself before committing. First: branch off `origin/master`, and fetch the engine repo so cross-repo verification is against fresh `origin/master`.

655. user (2026-06-22T09:45:55.372Z)

```
=== khelsutra: clean tree? branch off origin/master ===
[no tracked changes above = clean]
branch 'docs/freshness-sweep' set up to track 'origin/master'.
Switched to a new branch 'docs/freshness-sweep'
branch: docs/freshness-sweep | HEAD: 50b2bd7 chore(repo): ignore local runtime artifacts +
correct stale NEXT_STEPS claims (#94)
=== engine: fetch origin (for cross-repo verification) ===
engine origin/master: 5c2eec8 chore(security): P0 (safe scope) ? scrub personal paths from code
+ add leak-scan CI guard (#314)
```

656. assistant (2026-06-22T09:46:00.622Z)

On `docs/freshness-sweep`, engine fetched (its `origin/master` is at #314). Now the fixer fan-out ? 12 stale docs (I'll handle `NEXT_STEPS.md`'s one finding myself with non-brittle phrasing). Each fixer re-verifies before editing, edits only its own doc, and reports code-changes-needed rather than touching code.

657. user (2026-06-22T09:46:23.274Z)

```
Workflow launched in background. Task ID: w8ehcyzd2
Summary: Apply the verified doc-staleness fixes, one surgical fixer per doc, re-verifying
against the live repo
Transcript dir: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\subagents\workflows\wf_3d035317-127
Script file: C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-fix-
wf_3d035317-127.js
(Edit this file with Write/Edit and re-invoke Workflow with {scriptPath:
"C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-fix-
wf_3d035317-127.js"} to iterate without resending the script.)
Run ID: wf_3d035317-127
To resume after editing the script: Workflow({scriptPath: "C:\Users\avidu\.claude\projects\C--
Users-avidu-Projects-badminton-highlight-
indexer\8d4866bf-8edd-4da9-909e-022ddee12675\workflows\scripts\khelsutra-doc-fix-
wf_3d035317-127.js", resumeFromRunId: "wf_3d035317-127"}) ? completed agents return cached
results.

You will be notified when it completes. Use /workflows to watch live progress.
```

658. assistant (2026-06-22T09:46:38.784Z)

While the fixers run, let me handle `NEXT_STEPS.md`'s one finding myself ? but instead of bumping `#93 ? #94` (which would just lag again the moment this sweep merges as the next PR), I'll make the count non-brittle.

659. user (2026-06-22T09:46:41.384Z)

```
The file C:\Users\avidu\Projects\khelsutra\NEXT_STEPS.md has been updated successfully. (file
state is current in your context ? no need to Read it back)
```

660. assistant (2026-06-22T09:46:44.508Z)

NEXT_STEPS handled. Now waiting on the 12 fixer agents ? I'll review the complete diff, apply any code-comment fixes they flag, run the site suite, then commit and PR.

661. user (2026-06-22T09:51:11.922Z)

<task-notification>
<task-id>w8ehcyzd2</task-id>
<tool-use-id>toolu_01Qd3S5FjpYY4MjRbZrfjzDh</tool-use-id>
<output-file>C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\w8ehcyzd2.output</output-file>
<status>completed</status>
<summary>Dynamic workflow "Apply the verified doc-staleness fixes, one surgical fixer per doc, re-verifying against the live repo" completed</summary>
<result>{"reports":[{"doc":"Fixed 3 of 3 stale README.md findings (2 surgical edits; the 3rd was resolved transitively per the audit's own suggested-fix guidance, so no separate edit). All claims re-verified against the live khelsutra repo and cross-checked against deploy docs before editing. No 'unverifiable' perDoc items required action. Edited file: C:\\\\Users\\\\avidu\\\\Projects\\\\khelsutra\\\\README.md. The other README.md findings the audit marked 'correct' (engine sibling 'rename pending', canonical planning docs in engine repo, deploy/ kit, site/, web/ tech stack, License) were left untouched as accurate. Docs-only change; no code edits needed.","editsApplied":[{"location":"README.md:9 (Status line)","summary":"Flipped the stale 'pre-alpha scaffold. SPA lands in web/ per execution-plan milestone M1.' to the accurate current status: 'alpha ? the web/ SPA is built and multilingual (6 languages: en/kn/hi/es/da/id), with the per-rally selection MVP and first-run setup wizard shipped; khelsutra.guru / app.khelsutra.guru are live (Cloudflare Pages + cloudflared tunnel + Access). See NEXT_STEPS.md.' Verified against the live repo: web/ has 31 commits and 51 source files incl. App.tsx, SetupWizard.tsx, IngestPanel.tsx; i18n SUPPORTED_LOCALES = [en,kn,hi,es,da,id]; deploy/droplet/README.md:100 '? Live-verified on the droplet (2026-06-22)' with app.khelsutra.guru up via Pages+tunnel+Access."},{"location":"README.md:16-18 (Structure list)","summary":"Moved 'control-plane/ ? (beta) Cloud Run FastAPI control plane' out of the present-tense Structure bullet list into a separate 'Planned (beta, not yet present): control-plane/ ? Cloud Run FastAPI control plane.' note, so a reader does not expect the path to exist. Verified: control-plane/ directory is absent (git ls-files returns nothing; top-level tracked dirs are deploy/docs/locales/site/web only)."}],"skipped":[{"finding":"README.md:4 ? tagline 'the web app (alpha SPA) now, the hosted control plane at beta' (medium, stale).","reason":"The audit's own suggested fix was to KEEP 'web app (alpha SPA) now' and instead align the surrounding intro/status so it no longer describes the SPA as not-yet-landed. That alignment is fully achieved by the README.md:9 status-line edit (and the control-plane is now honestly tagged 'Planned (beta)'). The tagline text itself is directionally correct and 'control plane at beta' is forward-looking, consistent with the new Planned note ? so no separate text edit to line 4 is warranted. Editing it further would be scope creep with no remaining contradiction to fix."}],"codeChangesNeeded":[],"notes":"Edited file: C:\\\\Users\\\\avidu\\\\Projects\\\\khelsutra\\\\README.md (only file touched). No git write commands run; read-only verification only. Cross-repo engine claims in this doc were the 'correct'-verdict ones (sibling repo name, canonical planning docs) and needed no change, so no engine-symbol re-anchoring was required for the applied edits."},{"doc":"deploy/DEPLOY_ALPHA.md","editsApplied":[{"location":"DEPLOY_ALPHA.md:3 (header Status)","summary":"Flipped Status from `READY` (\"owner executes the cloud steps\") to `LIVE`, citing the live-verification on the droplet 2026-06-22 in deploy/droplet/README.md:100-106 (DNS->Cloudflare->Access->tunnel->engine chain confirmed; Pages/Tunnel/Access up). Noted the only remaining human steps: a rotated Gemini key + a browser OTP login to exercise the upload->reel CUJ ? honest, not an over-claim of a fully-exercised CUJ."},{"location":"DEPLOY_ALPHA.md:66 (§2.2 export block)","summary":"Widened RALLY_ALLOWED_HOSTS from \"api.khelsutra.guru\" to \"api.khelsutra.guru,localhost,127.0.0.1\" to match the deployed kit engine.env.example:17. Added a rationale: setting this var REPLACES the engine's loopback-only default (verified in engine origin/master backend/main.py::_allowed_hosts_from_env ? `return hosts or [\"localhost\",\"127.0.0.1\",\"testserver\"]`), so without the loopback entries the on-box /api/health checks would 400."},{"location":"DEPLOY_ALPHA.md:70 (§2.2 ? callout)","summary":"Aligned the inline RALLY_ALLOWED_HOSTS value in the warning callout to the same \"api.khelsutra.guru,localhost,127.0.0.1\" so the prose no longer contradicts the code block above it; same finding (host list too narrow), kept internally consistent."}],"skipped":[],"codeChangesNeeded":[],"notes":"Both audited staleFindings for deploy/DEPLOY_ALPHA.md were re-verified against the live repo and applied surgically.\\n\\nFinding 1 (header stale -> LIVE, severity medium): CONFIRMED.

deploy/droplet/README.md:100-106 records \\\"Live-verified on the droplet (2026-06-22)\\\" with the full DNS->Cloudflare->Access->tunnel->engine chain confirmed and Pages/Tunnel/Access up. The `READY` header lagged. Flipped to `LIVE` but kept it HONEST ? the upload->reel CUJ itself still needs a rotated Gemini key + a browser OTP login (per droplet README:105-106), so I stated that rather than claiming a fully-exercised CUJ.\\n\\nFinding 2 (host list too narrow, severity low): CONFIRMED. The deployed kit deploy/droplet/engine.env.example:17 uses `RALLY_ALLOWED_HOSTS=api.khelsutra.guru,localhost,127.0.0.1` with the explicit rationale \\\"loopback added for on-box health checks\\\", whereas DEPLOY_ALPHA.md §2.2 only set `api.khelsutra.guru`. Verified the WHY against engine origin/master backend/main.py::_allowed_hosts_from_env (line ~89): when RALLY_ALLOWED_HOSTS is set it REPLACES (does not append) the loopback-only default `[\\\"localhost\\\",\\\"127.0.0.1\\\",\\\"testserver\\\"]`, so the narrow value drops loopback and would 400 on-box /api/health. Fixed in both the §2.2 export block (line 66) and the §2.2 ? callout (line 70) for internal consistency.\\n\\nUnverifiable items in the perDoc.findings (none flagged as actionable for this doc) were not touched. No code changes required; nothing skipped. I did NOT run any git write commands ? git was read-only (engine origin/master verification of the Host-allowlist semantics)."},{"doc":"docs/ANNOTATE_AND_TRAIN.md","editsApplied":[{"location":"Header line 3 (Status)","summary":"Flipped Status DRAFT -> IN PROGRESS (P0-P4 shipped #15/#22/#23/#25; only owner-gated P5-P9 remain) and bumped Last updated 2026-06-20 -> 2026-06-22. Verified gate/PR history on both repos' origin/master."},{"location":"Line 5 (Tracking anchors)","summary":"Corrected the false 'index in docs/README.md' claim to 'NOT yet indexed in docs/README.md (still to add)'. Verified: git grep for annotate/own-model in docs/README.md on khelsutra origin/master returns nothing."},{"location":"§7 P0 row (line 71)","summary":"Changed status ? -> ? and reworded to 'doc landed #15; docs/README.md index entry still to add', since the doc exists (#15) but is not actually indexed. Kept #15 link."},{"location":"§7 P5 row (line 76)","summary":"Split P5 into P5a (annotator-UI screenshots ?, cite #25 ? site/public/img/annotate-enable.png + annotate-tool.png are committed and referenced in own-model.html:111-112) and P5b (real footage screenshots ?, owner-gated). Verified the two PNGs are tracked and #25 (1906c82) added them."},{"location":"§7 P1-P4 rows (lines 72-75)","summary":"Replaced the '(this PR)' PR-column placeholders with [#22] (own-model page P1-P4, b1110e1) and added [#23] (org-link follow-up) on the P4 row. Verified PR titles via git log on origin/master."},{"location":"Relation note (line 6) + Nice-to-knows (line 84)","summary":"Dropped the future-tense 'lands with PR #13' / 'once PR #13 merges' framing for LOCAL_APPLIANCE_ARCHITECTURE.md, which already exists in the repo (git ls-files confirms)."},{"location":"§4.3 Reuse map ? SPA bullet (line 53)","summary":"Reworded 'web/src/App.tsx is one static screen' to 'App.tsx renders a first-run SetupWizard + the ingest panel' (post-#88). Kept the 'no router' sub-claim (still true; no react-router dep). Verified App.tsx:11/242 imports/renders SetupWizard on origin/master."},{"location":"TL;DR (line
... (truncated 41294 chars, full result in C:\Users\avidu\AppData\Local\Temp\claude\C--Users-avidu-Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tasks\w8ehcyzd2.output)</result>
<usage><agent_count>12</agent_count><subagent_tokens>998215</subagent_tokens><tool_uses>205</tool_uses><duration_ms>288538</duration_ms></usage>
</task-notification>

662. assistant (2026-06-22T09:51:17.594Z)

All 12 fixers reported. Let me pull the structured summary (edits + any code-changes-needed), then ? most importantly ? review the actual `git diff` myself before trusting any of it.

663. user (2026-06-22T09:51:22.760Z)

LOG: 12/12 fixers reported; 63 edits applied; 2 code-changes-needed

=== Fixed 3 of 3 stale README.md findings (2 surgical edits; the ===
  edits: 2 | skipped: 1 | codeNeeds: 0
   SKIP: README.md:4 ? tagline 'the web app (alpha SPA) now, the hosted control plane at beta' (medium, stale). -> The audit's own suggested fix was to KEEP 'web app (alpha SPA) now' and instead align the surrounding intro/status so it no longer describes the SPA as not-yet-landed. That alignment is fully achieved by the README.md:9 status-line edit (and the control-plane is now honestly tagged 'Planned (beta)'). The tagline text itself is directionally correct and 'control plane at beta' is forward-looking, consistent with the new Planned note ? so no separate text edit to line 4 is warranted. Editing it further would be scope creep with no

remaining contradiction to fix.

=== deploy/DEPLOY_ALPHA.md ===
  edits: 3 | skipped: 0 | codeNeeds: 0

=== docs/ANNOTATE_AND_TRAIN.md ===
  edits: 13 | skipped: 2 | codeNeeds: 1
   SKIP: DoD bullet (line 92) 'doc indexed in docs/README.md' and §7 DoD claim -> Left as-is: line 92 is an UNCHECKED '- [ ]' DoD requirement, so it honestly states indexing is still to do ? it is not a false completion claim. The false completion claims were at lines 5 and 71, both corrected. No change needed here.
   SKIP: perDoc 'correct'-verdict citations the audit did not flag (e.g. gt_loader.py:14-90 at line 47; stub_runner synth :124-149 range; rally_annotator.lua line refs) -> These were verdict=correct in the audit and re-verified as accurate (load_gt_intervals at :44 within the cited 14-90 span; synth function spans :124-149; row['start_time'] at :139). Not stale; left untouched to avoid scope creep.

=== docs/DRY_RUN_CUJ.md ===
  edits: 2 | skipped: 2 | codeNeeds: 0
   SKIP: Engine line-number drift in §1/§8 (main.py:274/283/159, ref.py:60 cited; engine origin/master now ~main.py:581/590/458) ? DRY_RUN_CUJ.md:15,68-72,76 -> Audit verdict was 'correct' (not 'stale'), severity low, with an explicitly OPTIONAL suggestedFix ('Optionally refresh the cited line numbers'). It is not in this doc's staleFindings worklist; the routing logic the lines describe is intact. Refreshing brittle engine line anchors here would be scope creep beyond the assigned stale findings. (ref.py:60 'gs://?gcs' still verifies on engine origin/master, so it is not even drifted.)
   SKIP: GCS-artifact / run-result specifics: 12 rallies, video_id 05e0e577, 50.9MB/76.5s reels, screencast paths gs://khelsutra-rally-corpus/promo/*.mp4, staged bucket video ? DRY_RUN_CUJ.md:54-60,82-85 -> Audit verdict 'unverifiable' ? these are GCS bucket objects and a one-off local run's outputs, not checkable from the repo, with no counter-evidence. Per instructions, do not act on unverifiable items unless now positively verifiable; I have no GCS access.

=== docs/LOCAL_APPLIANCE_ARCHITECTURE.md (KhelSutra repo) ? fixe ===
  edits: 13 | skipped: 5 | codeNeeds: 0
   SKIP: perDoc unverifiable: T11 retention sweep tool (engine #299) + auto-scheduling (engine #301) marked ?. -> References engine-repo GitHub PR/issue numbers not falsifiable from the local checkouts; the audit itself marked this unverifiable (severity low) with no suggestedFix. Left as-is.
   SKIP: perDoc unverifiable: §3 droplet filesystem/port facts (/srv/khelsutra-site, node :8787, ~/rally, MinIO 9000/9001, GCS emulator 9023). -> Live-droplet (168.144.126.176) facts not checkable from the repo; audit marked unverifiable with no suggestedFix. No edit.
   SKIP: perDoc unverifiable: D10/§12 claim that this doc 'supersedes DESKTOP_APP_PLAN.md'. -> DESKTOP_APP_PLAN.md is an engine-repo doc not present in the khelsutra tree; supersession status can't be confirmed from this repo. Audit marked unverifiable, no suggestedFix. No edit.
   SKIP: §3 line 47 /api/health anchor '[verified main.py:127 ...]' ? on origin/master /api/health is at main.py:307, so the :127 line number has drifted. -> NOT in the audit's worklist for this doc (the audit flagged capabilities/auth/native/deploy/header, not this health line-number drift). Avoided scope creep; the substantive claim ('/api/health now EXISTS') remains correct. The accurate :307 anchor is already added in §3's new 'Shipped since' bullet.
   SKIP: §12 References block carries several pre-drift engine line anchors (e.g. main.py:42 for CORS). -> Not individually called out in this doc's staleFindings; editing the whole reference list would be a wholesale anchor-refresh beyond the audited fixes. Left for a dedicated anchor-refresh pass.

=== docs/MULTILINGUAL_EXPANSION.md ===
  edits: 6 | skipped: 2 | codeNeeds: 0
   SKIP: §8 Open question Q6 (line 95): 'Does site/ localize for these too for v1, or SPA-only until FRIENDLY B3 lands?' ? contains the now-satisfied 'until FRIENDLY B3 lands' framing. -> Not in the audit worklist. The EX3/SPA-only stale finding scoped its suggestedFix explicitly to §1/§4/§6 prose and the EX3 tracker row; line 95 is an owner-facing Open Question (a decision prompt), not a factual staleness claim, and was not flagged. Editing it would be scope creep beyond the findings. The §9 DoD line 104 ('after FRIENDLY B3') is likewise an unchecked DoD criterion whose dependency is harmlessly satisfied; left as-is.
   SKIP: Changelog '168 web tests green' (line 114) ? live suite now reports 223 passed. ->

Audit verdict is 'correct' (point-in-time changelog note dated 2026-06-21), not stale; suggestedFix was explicitly optional ('no fix needed'). Left untouched to preserve the historical record per the doc's voice.

=== docs/PER_RALLY_SELECTION.md ===
  edits: 7 | skipped: 2 | codeNeeds: 0
   SKIP: Finding 5 ? line-number drift in anchors OUTSIDE §3/§10 (the §2 decision table D4/D5/D7 at lines 29-32, §4.2 floor anchor line 71, §4.2 compile/compiler anchors line 72, §5 risk-table anchors lines 111-112, §9 DoD line 157, and the changelog lines 170-173). -> Finding 5's enumerated docLocation is specifically `docs/PER_RALLY_SELECTION.md:38-46, :163 (§10 anchors)` ? i.e. §3 and §10 only. The decision table, data-flow floor anchor, risk-table rows, DoD, and changelog were not in the worklist; the contracts they describe still hold (only line numbers drifted) and rewriting them would be scope creep. The Changelog in particular is historical record (documents what was true at each ship) and must not be retro-edited.
   SKIP: Finding 1 suggestedFix's literal 'Flip P6 to ?'. -> Applying the full ? would be dishonest: while both engine halves shipped, the SPA-side offline banner is NOT wired and the SPA does not yet consume download_url (verified: no /api/health or download_url usage in khelsutra web/src). Used the audit's offered alternative ('or split') and marked P6 ? with an explicit engine-done / SPA-pending note, per the honesty gate on status markers.

=== docs/PORTABILITY.md ===
  edits: 2 | skipped: 2 | codeNeeds: 1
   SKIP: PORTABILITY.md:11 ? optional update of `handleRequest(request, store)` to `handleRequest(request, store, notify)` to reflect the added notify injection. -> Audit verdict is 'correct' (not stale) ? listed only as an optional refinement, not a stale finding on my worklist. Per instructions I act only on stale findings and do not scope-creep. The 2-arg form it describes still accurately conveys the store-by-injection point.
   SKIP: PORTABILITY.md:57-61 ? 'Verified live (2026-06-20)' DO droplet provision/teardown/restore run figures (12 KB DB, rows:2). -> Audit verdict is 'unverifiable' ? a historical external live-run claim not falsifiable from the repo, and consistent with the shipped scripts (provision/teardown with export+RESTORE_DB exist and the booted-server round-trip is proven in CI). No positive evidence of staleness, so left unchanged.

=== docs/README.md ===
  edits: 2 | skipped: 2 | codeNeeds: 0
   SKIP: The audit's perDoc/staleFindings entries with doc == 'README.md' (no docs/ prefix) ? 'Status: pre-alpha scaffold' (README.md:9), 'web app (alpha SPA) now ... hosted control plane at beta' (README.md:4), 'control-plane/ ? (beta) Cloud Run FastAPI control plane' (README.md:17). -> Out of scope: these describe the repo-ROOT README.md (a structure list with 'pre-alpha scaffold', 'Structure', 'control-plane/', 'License' sections), NOT my assigned file docs/README.md (a docs index with none of that content). My task is to edit ONLY docs/README.md, so I left these for whoever is assigned the root README.
   SKIP: docs/README.md:16-18 sub-suggestion to also call the doc 'peer of the archived FRIENDLY_AND_MULTILINGUAL.md'. -> Not skipped ? applied. (Listed here only to note the FRIENDLY_AND_MULTILINGUAL.md doc is genuinely archived: the index entry at line 9 marks it 'DONE ? ARCHIVED (2026-06-22)', so 'archived' is honest.)

=== docs/REQUEST_ACCESS_FORM.md ===
  edits: 4 | skipped: 2 | codeNeeds: 0
   SKIP: perDoc.findings (verdict=correct): optional refresh of the §3 human-facing label prose to the post-PR#75 plain-language copy ('GPU machine for local processing?' is the pre-#75 wording). -> Not a staleness finding (verdict was 'correct' ? every field name and enum value still matches). The suggestedFix is explicitly 'Optional' and would require transcribing live UI copy verbatim; doing so risks scope creep and re-introducing inaccuracy. Left as-is per the no-scope-creep rule.
   SKIP: perDoc.findings (verdict=unverifiable): 'Deployed LIVE on Cloudflare' claim in the §6 P2 row. -> Marked unverifiable in the audit with no suggestedFix; deployment is owner-run runtime state with no in-repo artifact to confirm or refute. Nothing to edit.

=== site/README.md ===
  edits: 6 | skipped: 0 | codeNeeds: 0

=== web/README.md ===
  edits: 3 | skipped: 1 | codeNeeds: 0
   SKIP: perDoc findings with verdict=correct: client.ts correlation id (web/README.md:26),

types.ts motion-only fabrication note (:27-28), log.ts observability seed (:29), dev proxy 5173?8000 (:11), `python -m backend.main --server` (:13), Checks typecheck/test/build (:17-22), and the See-docs links (:4-6). -> All verdict=correct and re-verified true against the live repo (vite.config.ts proxy, package.json scripts, engine backend/main.py --server, types.ts:20-26, both linked docs exist). No edit warranted ? not in the staleFindings worklist.

######## CODE CHANGES NEEDED ########
  [docs/ANNOTATE_AND_TRAIN.md] C:/Users/avidu/Projects/khelsutra/docs/README.md: Add an ANNOTATE_AND_TRAIN.md entry to the 'Tracked work' list so the doc is actually indexed (the doc's P0 deliverable / DoD requires it). Currently git grep for annotate/own-model in docs/README.md on origin/master returns nothing. (The doc repeatedly claimed it is indexed in docs/README.md, but it is not. I corrected the doc's claims to honest 'not yet indexed' (P0 ?), but the real fix to satisfy P0/DoD is to add the index entry in docs/README.md ? a different file outside my edit scope.)
  [docs/PORTABILITY.md] site/server.js:1 (header comment): Change the header comment "Always-on VM host for the KhelSutra site ? dependency-free (Node 22+ only)." to say Node 24+ (e.g. "dependency-free (Node 24+ only ? node:sqlite is unflagged at >=24)."), (The audit's suggestedFix for the Node-version staleness covers BOTH docs/PORTABILITY.md:64 AND site/server.js:1, which both still say 'Node 22+'. The repo authoritatively requires >=24 (provision-vm.sh NODE_MAJOR>=24, CI node-version 24; node:sqlite was flagged before ~Node 22.5). The doc half is fixed; the server.js comment is a code file outside my assigned doc, so it is recorded here rather than edited.)

TOTAL edits: 63 skipped: 21

664. assistant (2026-06-22T09:51:38.634Z)

Solid work ? 63 surgical edits, with disciplined skips (preserved historical changelogs, refused to act on unverifiable GCS claims, marked P6 `?` engine-done/SPA-pending rather than a dishonest `?`). Two code/cross-file needs flagged. But agent self-reports aren't proof ? let me **review the actual diff** before trusting any of it.

665. user (2026-06-22T09:51:40.557Z)

```
=== diff stat ===
 NEXT_STEPS.md                        |  3 ++-
 README.md                            |  5 +++--
 deploy/DEPLOY_ALPHA.md               |  6 +++---
 docs/ANNOTATE_AND_TRAIN.md           | 40 +++++++++++++++++++----------------
 docs/DRY_RUN_CUJ.md                  |  4 ++--
 docs/LOCAL_APPLIANCE_ARCHITECTURE.md | 28 ++++++++++++++-----------
 docs/MULTILINGUAL_EXPANSION.md       | 12 +++++------
 docs/PER_RALLY_SELECTION.md          | 42 ++++++++++++++++++++------------------
 docs/PORTABILITY.md                  |  6 +++---
 docs/README.md                       | 10 ++++++---
 docs/REQUEST_ACCESS_FORM.md          |  8 +++----
 site/README.md                       | 23 +++++++++++++++++----
 web/README.md                        | 14 ++++++++++--
 13 files changed, 118 insertions(+), 83 deletions(-)

=== sanity: any accidental code-file edits? (should be docs/*.md, *.md only) ===
NEXT_STEPS.md
README.md
deploy/DEPLOY_ALPHA.md
docs/ANNOTATE_AND_TRAIN.md
docs/DRY_RUN_CUJ.md
docs/LOCAL_APPLIANCE_ARCHITECTURE.md
docs/MULTILINGUAL_EXPANSION.md
docs/PER_RALLY_SELECTION.md
docs/PORTABILITY.md
docs/README.md
docs/REQUEST_ACCESS_FORM.md
site/README.md
web/README.md
```

666. assistant (2026-06-22T09:51:45.676Z)

Clean ? only `.md` files touched, no stray code edits. Now the full diff review.

667. user (2026-06-22T09:51:47.425Z)

```
diff --git a/NEXT_STEPS.md b/NEXT_STEPS.md
index b53bb2e..bf6ea5c 100644
--- a/NEXT_STEPS.md
+++ b/NEXT_STEPS.md
@@ -12,3 +12,4 @@ live). **Git auto-deploy is connected** (Workers Builds, root `site/`): merge
to
 dependency-free **Node + SQLite** host (`site/server.js`), **live + browsable on the DO droplet
?
-http://168.144.126.176:8787** (systemd `khelsutra-site`). PRs through #93 merged (master).
Docs:
+http://168.144.126.176:8787** (systemd `khelsutra-site`). Shipped across ~90 merged PRs (the
form,
+6-language localization, the per-rally MVP, droplet portability). Docs:
 `docs/REQUEST_ACCESS_FORM.md`, `docs/PORTABILITY.md`.
diff --git a/README.md b/README.md
index d1a591d..0f6180f 100644
--- a/README.md
+++ b/README.md
@@ -8,3 +8,3 @@ the web app (alpha SPA) now, the hosted control plane at beta.
   `UI_PROPOSAL.md` · `HOSTING_PLAN.md` · `UI_ALPHA_EXECUTION_PLAN.md`.
-- **Status:** pre-alpha scaffold. SPA lands in `web/` per execution-plan milestone M1.
+- **Status:** alpha ? the `web/` SPA is built and multilingual (6 languages:
en/kn/hi/es/da/id), with the per-rally selection MVP and first-run setup wizard shipped;
`khelsutra.guru` / `app.khelsutra.guru` are live (Cloudflare Pages + cloudflared tunnel +
Access). See [`NEXT_STEPS.md`](NEXT_STEPS.md).

@@ -16,3 +16,4 @@ the web app (alpha SPA) now, the hosted control plane at beta.
 - `docs/` ? product docs as they migrate from the engine repo
-- `control-plane/` ? (beta) Cloud Run FastAPI control plane
+
+**Planned (beta, not yet present):** `control-plane/` ? Cloud Run FastAPI control plane.

diff --git a/deploy/DEPLOY_ALPHA.md b/deploy/DEPLOY_ALPHA.md
index 311e025..e9d7cc2 100644
--- a/deploy/DEPLOY_ALPHA.md
+++ b/deploy/DEPLOY_ALPHA.md
@@ -2,3 +2,3 @@

-> **Status:** `READY` (owner executes the cloud steps) · **Topology:** split ? SPA on
Cloudflare Pages + engine on an `api.*` cloudflared tunnel, Cloudflare Access over both (the
engine's `HOSTING_PLAN.md` §2 alpha plan). · **Cost:** ~$0/mo standing (Pages + Tunnel + Access
are free ?50 users) + Gemini per-run (~$0.05/match).
+> **Status:** `LIVE` (deploy live-verified on the droplet 2026-06-22 ?
`deploy/droplet/README.md`; only a rotated Gemini key + a browser OTP login remain to exercise
the upload?reel CUJ) · **Topology:** split ? SPA on Cloudflare Pages + engine on an `api.*`
cloudflared tunnel, Cloudflare Access over both (the engine's `HOSTING_PLAN.md` §2 alpha plan).
· **Cost:** ~$0/mo standing (Pages + Tunnel + Access are free ?50 users) + Gemini per-run
(~$0.05/match).

@@ -65,3 +65,3 @@ export GEMINI_API_KEY="<your-rotated-key>"
 export RALLY_CORS_ALLOW_ORIGINS="https://app.khelsutra.guru"   # lock CORS to the SPA origin
-export RALLY_ALLOWED_HOSTS="api.khelsutra.guru"                # REQUIRED ? see ? below;
without it EVERY tunneled request 400s
+export RALLY_ALLOWED_HOSTS="api.khelsutra.guru,localhost,127.0.0.1"  # REQUIRED ? see ? below;
without it EVERY tunneled request 400s (loopback added so on-box /api/health checks still pass ?
setting this var REPLACES the loopback-only default)
 export RALLY_BIND_HOST="127.0.0.1"                             # keep the loopback bind
(cloudflared dials 127.0.0.1:8000)
@@ -69,3 +69,3 @@ export RALLY_BIND_HOST="127.0.0.1"                             # keep the
```

loopba

-> **? Two env vars you MUST set, or the tunnel is dead-on-arrival.** With `edge_auth_enabled: true`, the engine (a) flips its bind to `0.0.0.0` and (b) keeps the DNS-rebinding **Host-header guard** at its loopback-only default ? so it boots "healthy" but then **`400`s every request `cloudflared` forwards with `Host: api.khelsutra.guru`** (including the §6 `/api/health` check). The boot only *warns* about this (`EXPOSED_HOST_ALLOWLIST_LOOPBACK`), it does not stop. Set **`RALLY_ALLOWED_HOSTS="api.khelsutra.guru"`** (the load-bearing fix ? the Host allowlist is independent of the bind) and **`RALLY_BIND_HOST="127.0.0.1"`** (so the bind stays loopback as §0/§6 claim; `cloudflared` dials loopback anyway). *Verified against engine `origin/master`: with the allowlist unset, `GET /api/health` with `Host: api.khelsutra.guru` ? `400`; with it set ? `200`.*
+> **? Two env vars you MUST set, or the tunnel is dead-on-arrival.** With `edge_auth_enabled: true`, the engine (a) flips its bind to `0.0.0.0` and (b) keeps the DNS-rebinding **Host-header guard** at its loopback-only default ? so it boots "healthy" but then **`400`s every request `cloudflared` forwards with `Host: api.khelsutra.guru`** (including the §6 `/api/health` check). The boot only *warns* about this (`EXPOSED_HOST_ALLOWLIST_LOOPBACK`), it does not stop. Set **`RALLY_ALLOWED_HOSTS="api.khelsutra.guru,localhost,127.0.0.1"`** (the load-bearing fix ? the Host allowlist is independent of the bind; the loopback entries keep the on-box `/api/health` check passing, since setting this var REPLACES the loopback-only default) and **`RALLY_BIND_HOST="127.0.0.1"`** (so the bind stays loopback as §0/§6 claim; `cloudflared` dials loopback anyway). *Verified against engine `origin/master`: with the allowlist unset, `GET /api/health` with `Host: api.khelsutra.guru` ? `400`; with it set ? `200`.*

diff --git a/docs/DRY_RUN_CUJ.md b/docs/DRY_RUN_CUJ.md
index a360566..a85d03c 100644
--- a/docs/DRY_RUN_CUJ.md
+++ b/docs/DRY_RUN_CUJ.md
@@ -2,3 +2,3 @@

-> **Status:** `DRAFT` ? owner-gated · **Owner:** Avi Dullu · **Created:** 2026-06-21 · **Goal for this session.**
+> **Status:** `DONE ? DELIVERED (2026-06-21)` · **Owner:** Avi Dullu · **Created:** 2026-06-21. DoD §6 boxes 1?3 ? (Path A + Path B/B-P2 live e2e); the only open box (droplet, §6 box 4) is **owner-parked** (GPU subscription). Archive once that box closes.
 > **Umbrella:** a milestone of `docs/LOCAL_APPLIANCE_ARCHITECTURE.md` (the appliance) + `docs/PER_RALLY_SELECTION.md` (the shipped SPA back-half).
@@ -10,3 +10,3 @@ Operate the **web UI** ? point at a **video on the GCS bucket** ? the engine
 ## 1. What already works vs the gap `[verified]`
-**? Back half (shipped ? PER_RALLY_SELECTION M0):** the SPA loads `?video=<id>` ? `POST /api/query` (rallies, default all-selected) ? deterministic query/select ? `POST /api/compile` ? `<video>` preview + **`<a download>` saves the MP4 locally**. Reel download-to-local **works today** (`GET /api/highlights/{video_id}/{filename}` streams it; `/api/compile` also returns `download_url`). Fully **en/kn/hi** (the switcher / `?lang=`).
+**? Back half (shipped ? PER_RALLY_SELECTION M0):** the SPA loads `?video=<id>` ? `POST /api/query` (rallies, default all-selected) ? deterministic query/select ? `POST /api/compile` ? `<video>` preview + **`<a download>` saves the MP4 locally**. Reel download-to-local **works today** (`GET /api/highlights/{video_id}/{filename}` streams it; `/api/compile` also returns `download_url`). Fully **en/kn/hi/es/da/id (6 languages)** (the switcher / `?lang=`).

diff --git a/docs/MULTILINGUAL_EXPANSION.md b/docs/MULTILINGUAL_EXPANSION.md
index 5e209ca..7724338 100644
--- a/docs/MULTILINGUAL_EXPANSION.md
+++ b/docs/MULTILINGUAL_EXPANSION.md
@@ -15,3 +15,3 @@ Add **Thai, Indonesian, Danish, Spanish, Chinese (Simplified), Japanese, Korean*
 Why now. en/kn/hi shipped and proved the foundation is locale-pluggable (a Kannada-default browser auto-renders Kannada; the switcher walks kn?hi?en). The owner wants to broaden reach beyond the India beachhead ? **SE Asia (th, id), Europe (da, es), and East Asia (zh, ja, ko)**.
-**Concrete outcome ("good"):** a user whose browser is in any of the 7 languages lands on the Studio (and, once `site/` is localized via FRIENDLY B3, the marketing pages too), and ? without help ? understands the product, runs the tool, and gets a reel, **entirely in their language**, with the fonts rendering crisply offline.
+**Concrete outcome ("good"):** a user whose browser is in any of the 7 languages lands on the Studio (and the marketing pages too ? FRIENDLY B3 landed [#70], with es/da/id site parity

shipped [#87]; th/zh/ja/ko site parity still pending), and ? without help ? understands the product, runs the tool, and gets a reel, **entirely in their language**, with the fonts rendering crisply offline.
 Read to get current: this doc §3?§4; [`FRIENDLY_AND_MULTILINGUAL.md`](archives/FRIENDLY_AND_MULTILINGUAL.md) (the foundation this rides on, esp. §4 design + §6.5 negotiation tests); `web/src/i18n/index.ts` (the detector/ICU/fallback config) + `web/src/main.tsx` (the `@fontsource` font-load pattern) + `web/src/i18n/hygiene.test.ts` (key-parity + no-bare-strings + offline guards).
@@ -48,4 +48,4 @@ Per-language diligence (Unicode CLDR plural-rules chart; SIL OFL; `@fontsource`;
 Reuse map (what's reused vs new):
-- *Reused unchanged:* the negotiation contract (persisted pref ? `?lang=` ? navigator ? `en`), the ICU catalog format, the `@fontsource` `unicode-range`-scoped self-host pattern (`main.tsx:7-8`), the key-parity + no-bare-strings + offline CI (`hygiene.test.ts`), the draft?review pipeline, the `<html lang>` sync (`i18n/index.ts:59-65`).
-- *New:* 7 `common.json` catalogs; `SUPPORTED_LOCALES` entries (`i18n/index.ts:18`); **4 new Noto font imports**; an **es 3-form plural** allowance in the catalog + tooling; a **CJK lazy-load + offline-subset** strategy (Phase C); Thai line-break CSS safety net.
+- *Reused unchanged:* the negotiation contract (persisted pref ? `?lang=` ? navigator ? `en`), the ICU catalog format, the `@fontsource` `unicode-range`-scoped self-host pattern (`main.tsx:7-8`), the key-parity + no-bare-strings + offline CI (`hygiene.test.ts`), the draft?review pipeline, the `<html lang>` sync (`i18n/index.ts:70` `syncDocumentLang`).
+- *New:* 7 `common.json` catalogs; `SUPPORTED_LOCALES` entries (`i18n/index.ts:25`); **4 new Noto font imports**; an **es 3-form plural** allowance in the catalog + tooling; a **CJK lazy-load + offline-subset** strategy (Phase C); Thai line-break CSS safety net.
 - **? Catalog-home (resolved by FRIENDLY B3):** catalogs now live at repo-root **`/locales/{en,kn,hi,es,da,id}/common.json`** ? the shared top-level dir from the I18N_PLAN "D-CATALOG-HOME" diagram. FRIENDLY **B3** hoisted them out of `web/src/i18n/locales/` (the SPA re-points its imports there; the `site/` shim consumes a `copy:locales` mirror). These es/da/id catalogs moved with the hoist, as anticipated.
@@ -67,3 +67,3 @@ i18n is **presentation-only** ? no new personal-data collection, no biometrics
 - The **CJK offline bundle is large** (tens of MB) unless a subset build is adopted ? a real size trade-off, not free.
-- `site/` marketing localization for these 7 **depends on FRIENDLY B3** landing first (the shared shim); until then this is SPA-only.
+- `site/` marketing localization rides **FRIENDLY B3** (the shared shim) ? **B3 landed ([#70])** and **es/da/id site parity shipped ([#87])**, so the Phase-A trio is no longer SPA-only; only the **th/zh/ja/ko** site parity remains.

@@ -86,3 +86,3 @@ Legend: ? Todo · ? In progress · ? Done · ? Blocked/gated. **One smal
 | EX2 | **Owner-machine font eyeball** per new script (th, zh, ja, ko) ? cloud CI can't render fonts | ? | EB*/EC* | **Owner (owner machine)** | ? | ? |
-| EX3 | **`site/` marketing parity** for the 7 (rides FRIENDLY **B3** shared `t()` shim) | ? | FRIENDLY B3 | Blocked on B3 | ? | ? |
+| EX3 | **`site/` marketing parity** for the 7 (rides FRIENDLY **B3** shared `t()` shim) ? **es/da/id site parity shipped**; th/zh/ja/ko remain | ? | FRIENDLY B3 (landed) | No (after the catalogs ship) | ? | B3 [#70] + es/da/id site parity [#87] |

@@ -107,3 +107,3 @@ Legend: ? Todo · ? In progress · ? Done · ? Blocked/gated. **One smal
 ## 10. References
-**Internal `[verified]`:** `web/src/i18n/{index.ts,hygiene.test.ts}`, `web/src/main.tsx` (font-load pattern), repo-root `/locales/{en,kn,hi,es,da,id}/common.json` (hoisted by FRIENDLY B3 ? D-CATALOG-HOME); [`FRIENDLY_AND_MULTILINGUAL.md`](archives/FRIENDLY_AND_MULTILINGUAL.md), engine `docs/I18N_PLAN.md`, `docs/UI_PROPOSAL.md`.
+**Internal `[verified]`:** `web/src/i18n/{index.ts,hygiene.test.ts}`, `web/src/main.tsx` (font-load pattern), repo-root `/locales/{en,kn,hi,es,da,id}/common.json` (hoisted by FRIENDLY B3 ? D-CATALOG-HOME); [`FRIENDLY_AND_MULTILINGUAL.md`](archives/FRIENDLY_AND_MULTILINGUAL.md), engine `docs/{I18N_PLAN.md,UI_PROPOSAL.md}`.
 External `[researched]`: Unicode CLDR v48 plural-rules chart (https://www.unicode.org/cldr/charts/48/supplemental/language_plural_rules.html) + plural-rules spec; W3C BCP-47 / Chinese script subtags (`zh-Hans`/`zh-Hant`); SIL Open Font License 1.1 + Noto CJK LICENSE (github.com/notofonts/noto-cjk); `@fontsource-variable/noto-sans-{thai,sc,jp,kr}`; W3C Thai line-breaking (`thai-gap`); react-i18next + i18next-icu.
diff --git a/docs/PORTABILITY.md b/docs/PORTABILITY.md
index 3800776..6370687 100644

```
--- a/docs/PORTABILITY.md
+++ b/docs/PORTABILITY.md
@@ -24,4 +24,4 @@

-**Enforced, not asserted.** `node --test src/` runs BOTH paths and CI
(`.github/workflows/test.yml`,
-Node 24) runs them on every PR, so the repo stays compatible with both hosts:
+**Enforced, not asserted.** `npm test` (`node --test`) runs BOTH paths and CI
(`.github/workflows/test.yml`,
+Node 24) runs it on every PR, so the repo stays compatible with both hosts:
 - `src/index.test.js` ? the Cloudflare Worker + D1 logic.
@@ -63,3 +63,3 @@ reproducible across providers.
 ## VM deploy runbook ("if Cloudflare acts out")
-Node 22+ only ? **no `npm install`** (node:sqlite is built in).
+Node 24+ only (node:sqlite is unflagged at >=24) ? **no `npm install`** (node:sqlite is built
in).
 ```
diff --git a/docs/README.md b/docs/README.md
index f206dcf..cd01671 100644
--- a/docs/README.md
+++ b/docs/README.md
@@ -15,5 +15,6 @@
   eyeball (human gates), and **B7** (engine backend l10n, deferred).
-- **`MULTILINGUAL_EXPANSION.md`** ? proposal to add **7 more locales** (th/id/da/es/zh-
Hans/ja/ko)
- with en/kn/hi-level diligence; peer of `FRIENDLY_AND_MULTILINGUAL.md`, reusing its
foundation.
- `DRAFT` ? **owner-gated (approval to start)**, [#58].
+- **`MULTILINGUAL_EXPANSION.md`** ? `IN PROGRESS`: **Phase A (es/da/id) shipped** ([#67]);
**Thai
+ (B) / CJK (C) remain owner-gated**. Adds **7 locales** (th/id/da/es/zh-Hans/ja/ko) beyond the
+ en/kn/hi foundation with en/kn/hi-level diligence; peer of the archived
+ `FRIENDLY_AND_MULTILINGUAL.md`, reusing its foundation.
 - **`LOCAL_APPLIANCE_ARCHITECTURE.md`** ? the single-box, UI-powered rally indexer + reel
@@ -25,2 +26,5 @@
   `/api/query` + `/api/compile` contract. `IN PROGRESS` (decisions locked).
+- **`ANNOTATE_AND_TRAIN.md`** ? annotate-your-footage ? train-your-own-model USP pages (the MIT
+ `rally-annotator` VLC extension + the engine's `backend/worker train` CLI). `DRAFT` ?
+ owner-gated on scope; peer of `LOCAL_APPLIANCE_ARCHITECTURE.md`.
 - **`REQUEST_ACCESS_FORM.md`** ? segmented lead-capture form + Worker/D1 backend (shipped).
diff --git a/docs/REQUEST_ACCESS_FORM.md b/docs/REQUEST_ACCESS_FORM.md
index f0470b9..6cb5e6d 100644
--- a/docs/REQUEST_ACCESS_FORM.md
+++ b/docs/REQUEST_ACCESS_FORM.md
@@ -54,6 +54,6 @@ someone raises a hand, **without shipping lead data to a third party**
(privacy-
 | Consent | checkbox | "Okay to email me about the KhelSutra alpha" |
-| `hp` | hidden honeypot | anti-spam; if filled ? silently drop |
+| `company` | hidden honeypot | anti-spam; if filled ? silently drop (`hp` is the CSS class) |

 ## 4. Design / architecture
-- **Frontend** (`site/index.html`): the header + hero "Request access" buttons scroll to the
+- **Frontend** (`site/public/index.html`): the header + hero "Request access" buttons scroll to
the
   `#join` form; `#join` becomes the form. On submit, JS `POST`s JSON to `/api/request-access`.
@@ -65,3 +65,3 @@ someone raises a hand, **without shipping lead data to a third party**
(privacy-
   has_gpu, storage (csv), message, consent, country, user_agent`.
-- **Anti-spam:** honeypot first; add **Cloudflare Turnstile** in P2 if spam appears.
+- **Anti-spam:** honeypot first; add **Cloudflare Turnstile** in P3 if spam appears.
 - **Notify (P2):** optional email ping to `hello@` per submission (MailChannels / Cloudflare
Email).
@@ -84,3 +84,3 @@ Legend: ? Todo · ? In progress · ? Done. One PR per row off `origin/mast
 | P2 | Worker + D1 + `wrangler.toml` + `schema.sql` + **unit tests (10 green)** | ? | ? |
LIVE on Cloudflare (Worker `hidden-king-khelsutra` + D1 `khelsutra-access`; persists +
```
```

```
emails per submission) |
-| P3 | Anti-spam (Turnstile) + per-submission email notify | No | ? | **email-notify LIVE**
(per-submission to `hello@`); Turnstile still ? |
+| P3 | Anti-spam (Turnstile) + per-submission email notify | No | ? | **email-notify LIVE**
(per-submission to `avi.dullu@gmail.com`); Turnstile still ? |
 | P4 | Tiny admin/export view (auth'd) to read responses | owner | ? | ? |
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

668. assistant (2026-06-22T09:52:20.044Z)

Batch 1 is excellent ? honest status flips (note `DEPLOY_ALPHA` keeps the caveat "only a rotated Gemini key + browser OTP login remain", not an over-claim), correct path/anchor fixes, and `docs/README.md` already got the `ANNOTATE_AND_TRAIN` index entry (so code-need #1 is moot). Now the engine-heavy docs ? and I'll independently spot-check the load-bearing engine claims they assert, since those are the riskiest.

669. user (2026-06-22T09:52:21.696Z)

```
<persisted-output>
Output too large (43.6KB). Full output saved to: C:\Users\avidu\.claude\projects\C--Users-avidu-
Projects-badminton-highlight-indexer\8d4866bf-8edd-4da9-909e-022ddee12675\tool-
results\bvsuihy4v.txt

Preview (first 2KB):
diff --git a/docs/ANNOTATE_AND_TRAIN.md b/docs/ANNOTATE_AND_TRAIN.md
index 09630a4..7db31bb 100644
--- a/docs/ANNOTATE_AND_TRAIN.md
+++ b/docs/ANNOTATE_AND_TRAIN.md
@@ -2,6 +2,6 @@

-> **Status:** `DRAFT` ? owner-gated on scope · **Owner:** Avi · **Created:** 2026-06-20 ·
**Last updated:** 2026-06-20
+> **Status:** `IN PROGRESS` ? marketing surface (P0?P4) shipped (#15, #22, #23, #25); only
owner-gated P5?P9 remain · **Owner:** Avi · **Created:** 2026-06-20 · **Last updated:**
2026-06-22
 > **Lifecycle:** `DRAFT ? IN PROGRESS ? DONE ? archived` (move to `docs/archives/` when DONE)
-> **Tracking anchors:** §7 progress tracker is the source of truth; index in `docs/README.md`;
pointer in `NEXT_STEPS.md` + memory `current-state` once work starts.
-> **Relation:** peer of `khelsutra/docs/PER_RALLY_SELECTION.md` and
`khelsutra/docs/LOCAL_APPLIANCE_ARCHITECTURE.md` (the latter lands with PR #13); mirrors the
engine repo's `PROJECT_DOC_TEMPLATE` discipline. Documents an EXTERNAL public tool (`rally-
annotator`, MIT) and the engine's `backend/worker train` CLI; it does not own those repos.
+> **Tracking anchors:** §7 progress tracker is the source of truth; NOT yet indexed in
`docs/README.md` (still to add); pointer in `NEXT_STEPS.md` + memory `current-state` once work
starts.
+> **Relation:** peer of `khelsutra/docs/PER_RALLY_SELECTION.md` and
`khelsutra/docs/LOCAL_APPLIANCE_ARCHITECTURE.md` (both already landed); mirrors the engine
repo's `PROJECT_DOC_TEMPLATE` discipline. Documents an EXTERNAL public tool (`rally-annotator`,
MIT) and the engine's `backend/worker train` CLI; it does not own those repos.
 > **Honesty note:** every engine-fact claim is tagged `[verified]` (read against code this
session), `[design]` (proposed here), or `[researched]` (needs sign-off). Not legal advice.
@@ -11,3 +11,3 @@
 ## 0. TL;DR
...
</persisted-output>
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

670. assistant (2026-06-22T09:52:31.046Z)

The diff is large ? let me read it in full while independently verifying the load-bearing engine claims these docs now assert (don't trust agent reports on cross-repo facts).

671. user (2026-06-22T09:52:36.442Z)

```
=== engine origin/master: API endpoints the docs now claim 'exist' ===
origin/master:backend/main.py:307:@app.get("/api/health")
origin/master:backend/main.py:334:@app.get("/api/version")
origin/master:backend/main.py:340:@app.get("/api/capabilities")
origin/master:backend/main.py:374:@app.get("/api/videos")
--- compile returns download_url? ---
origin/master:backend/main.py:413:                    "download_url":
f"/api/reels/{quote(e.name)}",
origin/master:backend/main.py:768:        download_url =
f"/api/highlights/{quote(req.video_id)}/{quote(out_name)}"
origin/master:backend/main.py:769:        return {"status": "success", "filepath": out_path,
"download_url": download_url}
--- list_videos in db? ---
origin/master:backend/infrastructure/database.py:545:    def list_videos(self, limit:
Optional[int] = None) -> List[Dict[str, Any]]:
=== native_wasb_runner on master (not branch-only)? ===
backend/pipeline/detectors/native_wasb_runner.py
tests/test_native_wasb_runner.py
=== Host-allowlist REPLACE semantics + CORS-from-env + TrustedHost + edge-auth ===
origin/master:backend/main.py:33:from starlette.middleware.trustedhost import
TrustedHostMiddleware
origin/master:backend/main.py:66:def _cors_origins_from_env(env: Optional[Dict[str, str]] =
None) -> List[str]:
origin/master:backend/main.py:89:def _allowed_hosts_from_env(env: Optional[Dict[str, str]] =
None) -> List[str]:
origin/master:backend/main.py:100:    # IPv4 loopback (the default bind) + the TestClient host.
NB: Starlette's TrustedHostMiddleware
origin/master:backend/main.py:110:    TrustedHostMiddleware,
```

672. assistant (2026-06-22T09:52:44.960Z)

```
All engine claims independently confirmed on engine `origin/master`: `/api/health` (307),
`/api/capabilities` (340), `/api/videos` (374), `download_url` from compile (768-769),
`db.list_videos` (545), `native_wasb_runner.py` on master, plus
`_cors_origins_from_env`/`_allowed_hosts_from_env`/`TrustedHostMiddleware`. Let me confirm the
one subtle claim (REPLACE-not-append host semantics) and read the full engine-heavy diff.
```

673. user (2026-06-22T09:52:47.070Z)

```
=== _allowed_hosts_from_env: REPLACE or append the loopback default? ===
def _allowed_hosts_from_env(env: Optional[Dict[str, str]] = None) -> List[str]:
    """P0c: Host-header allowlist ? a DNS-rebinding guard for the localhost appliance. A
malicious
    web page can point a domain it controls at 127.0.0.1 and script requests to the local
server;
    the browser sends ``Host: attacker.com``, which this rejects (the server only answers
loopback
    Host headers by default). ``RALLY_ALLOWED_HOSTS`` (comma-separated) overrides; the default
keeps
    Starlette's TestClient host ``testserver`` so the in-process suite isn't broken. Read from
the
    ENV (not config.json) so importing ``backend.main`` does zero filesystem I/O; a
routable/edge
    deploy sets its public host (or ``*`` when fronted by an auth gate that validates the
host)."""
    src = os.environ if env is None else env
    raw = (src.get("RALLY_ALLOWED_HOSTS", "") or "").strip()
    hosts = [h.strip() for h in raw.split(",") if h.strip()]
    # IPv4 loopback (the default bind) + the TestClient host. NB: Starlette's
TrustedHostMiddleware
    # matches on host.split(":")[0], which mangles a bracketed IPv6 Host ("[::1]:port"), so IPv6
    # loopback is NOT supported by this guard ? access the appliance via 127.0.0.1/localhost.
    return hosts or ["localhost", "127.0.0.1", "testserver"]
```

DNS-rebinding guard. Bound to localhost by default, so only loopback Host headers a (a remote page can't read responses cross-origin AND its Host header is rejected here off so an unexpected ``www.`` host is rejected rather than 307-redirected.

674. assistant (2026-06-22T09:52:54.817Z)

```
Confirmed ? `return hosts or [...]` means `RALLY_ALLOWED_HOSTS` **replaces** the loopback
default, so the `DEPLOY_ALPHA` fix (adding `localhost,127.0.0.1`) prevents a real on-box
`/api/health` 400. That's a genuinely valuable runbook correctness fix, not just freshness. Now
reading the full engine-heavy diff to confirm honesty/formatting.
```

675. user (2026-06-22T09:52:57.407Z)

```
1       diff --git a/docs/ANNOTATE_AND_TRAIN.md b/docs/ANNOTATE_AND_TRAIN.md
2       index 09630a4..7db31bb 100644
3       --- a/docs/ANNOTATE_AND_TRAIN.md
4       +++ b/docs/ANNOTATE_AND_TRAIN.md
5       @@ -2,6 +2,6 @@
6
7       -> **Status:** `DRAFT` ? owner-gated on scope · **Owner:** Avi · **Created:** 2026-06-20
· **Last updated:** 2026-06-20
8       +> **Status:** `IN PROGRESS` ? marketing surface (P0?P4) shipped (#15, #22, #23, #25);
only owner-gated P5?P9 remain · **Owner:** Avi · **Created:** 2026-06-20 · **Last updated:**
2026-06-22
9        > **Lifecycle:** `DRAFT ? IN PROGRESS ? DONE ? archived` (move to `docs/archives/` when
DONE)
10      -> **Tracking anchors:** §7 progress tracker is the source of truth; index in
`docs/README.md`; pointer in `NEXT_STEPS.md` + memory `current-state` once work starts.
11      -> **Relation:** peer of `khelsutra/docs/PER_RALLY_SELECTION.md` and
`khelsutra/docs/LOCAL_APPLIANCE_ARCHITECTURE.md` (the latter lands with PR #13); mirrors the
engine repo's `PROJECT_DOC_TEMPLATE` discipline. Documents an EXTERNAL public tool (`rally-
annotator`, MIT) and the engine's `backend/worker train` CLI; it does not own those repos.
12      +> **Tracking anchors:** §7 progress tracker is the source of truth; NOT yet indexed in
`docs/README.md` (still to add); pointer in `NEXT_STEPS.md` + memory `current-state` once work
starts.
13      +> **Relation:** peer of `khelsutra/docs/PER_RALLY_SELECTION.md` and
`khelsutra/docs/LOCAL_APPLIANCE_ARCHITECTURE.md` (both already landed); mirrors the engine
repo's `PROJECT_DOC_TEMPLATE` discipline. Documents an EXTERNAL public tool (`rally-annotator`,
MIT) and the engine's `backend/worker train` CLI; it does not own those repos.
14       > **Honesty note:** every engine-fact claim is tagged `[verified]` (read against code
this session), `[design]` (proposed here), or `[researched]` (needs sign-off). Not legal advice.
15      @@ -11,3 +11,3 @@
16       ## 0. TL;DR
17      -Ship ONE new static **marketing page** (`site/public/own-model.html`, slug `/own-
model`) with three anchored sections ? **Annotate your footage · Train your own model · Setup &
use** ? telling the bring-your-own-model story and linking the real, public, MIT **rally-
annotator** VLC plugin `[verified]`. The single most important caveat: only the **ANNOTATE half
is real today** ? the rally-annotator ships (v1.6.4, MIT) and its CSV is directly ingestible by
the engine with **no transform** `[verified]` ? while the **TRAIN half is an open, in-build
CLI** whose GPU-free path trains a **pipeline-validation mock only**, and whose **real model
needs a GPU + an in-flight native WASB runner** behind an **owner-gated ship decision**
`[verified gate_verdict harness.py:98]`. The Studio SPA gets a **design-tracked, owner-gated**
correction/contribute screen, **never a "Train now" button**. Defensible claim = **ownership /
licensing / privacy / consent / minors-excluded**, NOT accuracy-superiority (there is no
benchmark vs a paid LLM).
18      +Ship ONE new static **marketing page** (`site/public/own-model.html`, slug `/own-
model`) with three anchored sections ? **Annotate your footage · Train your own model · Setup &
use** ? telling the bring-your-own-model story and linking the real, public, MIT **rally-
annotator** VLC plugin `[verified]`. The single most important caveat: only the **ANNOTATE half
is real today** ? the rally-annotator ships (v1.6.4, MIT) and its CSV is directly ingestible by
the engine with **no transform** `[verified]` ? while the **TRAIN half is an open, in-build
CLI** whose GPU-free path trains a **pipeline-validation mock only**, and whose **real model
needs a GPU + an in-flight native WASB runner** behind an **owner-gated ship decision**
```

`[verified gate_verdict harness.py:142]`. The Studio SPA gets a **design-tracked, owner-gated** correction/contribute screen, **never a "Train now" button**. Defensible claim = **ownership / licensing / privacy / consent / minors-excluded**, NOT accuracy-superiority (there is no benchmark vs a paid LLM).

@@ -24,3 +24,3 @@ KhelSutra's differentiator vs paid LLMs is that **you can own a model trained on

| D5 | Frame USP as ownership/licensing/privacy/consent, NOT accuracy | `[verified]` no user-model-vs-LLM benchmark exists | No accuracy-superiority copy anywhere |

-| D6 | Paid LLMs never a training source; MIT/Apache/BSD only; minors excluded; opt-in | `[verified CLAUDE.md]` + `harness.py:178,189-192` attestation | Reuse the site's existing minors + opt-in wording (`index.html:342-349`) |

+| D6 | Paid LLMs never a training source; MIT/Apache/BSD only; minors excluded; opt-in | `[verified CLAUDE.md]` + `harness.py:222,233-237` attestation | Reuse the site's existing minors + opt-in wording (`index.html:342-349`) |

| D7 | Annotator CSV needs NO transform for the engine label path | `[verified]` §4.2 schema bridge | Document the MD5+stem join as the only real prereq |

@@ -30,5 +30,5 @@ KhelSutra's differentiator vs paid LLMs is that **you can own a model trained on

-**The train loop (partly real).** The engine's `backend/worker train` (cmd_train, `__main__.py:231`) pulls `labels/` (+ `videos/` for the detector source), builds per-video trajectory CSVs, assembles a manifest, **shells out** to `python -m training.gen0.harness` (the one-way train/serve wall, `training/README.md:9-11`), and pushes a versioned bundle to `weights/<version>/`. Two trajectory sources, identical downstream: **synth** (DEFAULT, GPU-free, a CPU mock that ENCODES the labels ? plumbing-meaningful only, `stub_runner.py:124-149,131-133`) and **detector** (real WASB, M3.5, **GPU-gated, IN-FLIGHT** ? the Linux-native runner does not exist yet, `02:15`, `04:98-113`). The harness fits a ~1M-param LogisticRegression on CPU and emits the artifact bundle (`weights.json` + reviewable `manifest.json`); **serving consumes only the bundle, never a code import** (ADR-008). The **ship-gate (`gate_verdict`, `harness.py:98`) currently refuses a commercial ship** regardless of F1: WASB checkpoint provenance C3 unresolved (commercial ship refused) + F1@tIoU not yet calibrated (#172).

+**The train loop (partly real).** The engine's `backend/worker train` (cmd_train, `__main__.py:332`) pulls `labels/` (+ `videos/` for the detector source), builds per-video trajectory CSVs, assembles a manifest, **shells out** to `python -m training.gen0.harness` (the one-way train/serve wall, `training/README.md:9-11`), and pushes a versioned bundle to `weights/<version>/`. Two trajectory sources, identical downstream: **synth** (DEFAULT, GPU-free, a CPU mock that ENCODES the labels ? plumbing-meaningful only, `stub_runner.py:124-149,131-133`) and **detector** (real WASB, M3.5, **GPU-gated, IN-FLIGHT** ? the Linux-native runner does not exist yet, `02:15`, `04:98-113`). The harness fits a ~1M-param LogisticRegression on CPU and emits the artifact bundle (`weights.json` + reviewable `manifest.json`); **serving consumes only the bundle, never a code import** (ADR-008). The **ship-gate (`gate_verdict`, `harness.py:142`) currently refuses a commercial ship** regardless of F1: WASB checkpoint provenance C3 unresolved (commercial ship refused) + F1@tIoU not yet calibrated (#172).

-**Guardrails (verbatim, non-negotiable).** Paid LLMs (Gemini/OpenAI/Anthropic) = inference/serving/fallback ONLY, **never a training data source** (bundle attests `no_paid_llm_training_labels:True`, `harness.py:192`); label source must be human (annotator golden CSVs, `harness.py:178`). Commercial-clean deps only ? **MIT/Apache-2.0/BSD**. Owned base = Qwen3-VL (Apache) primary; Gemma-4 AMBER/bench-only; Gemma 3 rejected. **Minors excluded** from the training pool; per-user training data needs **explicit opt-in** `[verified CLAUDE.md]`.

+**Guardrails (verbatim, non-negotiable).** Paid LLMs (Gemini/OpenAI/Anthropic) = inference/serving/fallback ONLY, **never a training data source** (bundle attests `no_paid_llm_training_labels:True`, `harness.py:236`); label source must be human (annotator golden CSVs, `harness.py:222`). Commercial-clean deps only ? **MIT/Apache-2.0/BSD**. Owned base = Qwen3-VL (Apache) primary; Gemma-4 AMBER/bench-only; Gemma 3 rejected. **Minors excluded** from the training pool; per-user training data needs **explicit opt-in** `[verified CLAUDE.md]`.

@@ -48,7 +48,7 @@ Annotator writes header `rally_number,start_time,end_time,ending_reason,sport,sh

-The worker's own inference CSV uses `start,end,shot_count,ending_reason` (`__main__.py:62-74`) ? order/short-name differ, but order never matters and `gt_loader` accepts

both aliases; the annotator's `*_time` form is the **safer** of the two. **The ONE real prerequisite is an IDENTITY/JOIN constraint, not a schema gap:** training keys video?label by **file MD5** and stem-matches `labels/<name>.rallies.csv` ? `videos/<name>.<ext>` (`__main__.py:271-280`; `02:48-60`). Stems must agree and the staged video bytes must be **byte-identical** to what was annotated ? any re-encode changes the MD5 and breaks the join (collector manifest currently omits `md5`, M2 prereq, `02:62-68`).
+The worker's own inference CSV uses `start,end,shot_count,ending_reason` (`__main__.py:72-85`) ? order/short-name differ, but order never matters and `gt_loader` accepts both aliases; the annotator's `*_time` form is the **safer** of the two. **The ONE real prerequisite is an IDENTITY/JOIN constraint, not a schema gap:** training keys video?label by **file MD5** and stem-matches `labels/<name>.rallies.csv` ? `videos/<name>.<ext>` (`__main__.py:383-384`; `02:48-60`). Stems must agree and the staged video bytes must be **byte-identical** to what was annotated ? any re-encode changes the MD5 and breaks the join (collector manifest currently omits `md5`, M2 prereq, `02:62-68`).

4.3 Reuse map

-- **Marketing** `[verified]`: clone `site/public/capture.html` skeleton + brand tokens (`index.html:25-32`); reuse `.honest`/`.nots` components (`:152-158`, used `:351-354`, `:400-415`); add a `/own-model` alias in `site/server.js:49-51` (Worker `env.ASSETS.fetch` needs no change, `site/src/index.js:84`); cross-link nav `:218-223`, the "gets sharper" section `:327-356`, footer `:528`; screenshots under `site/public/img/` (precedent `site/og.png` + `make_og.py`).
-- **SPA** `[design, gated]`: today `web/src/App.tsx` is one static screen, **no router** (`web/package.json`). A correction screen needs react-router-dom (or a hash switch) + the typed client `web/src/api/client.ts`/`types.ts` + `web/src/lib/log.ts`; it realizes P10 in `PER_RALLY_SELECTION.md:122` and depends on a **new** engine endpoint `PATCH /api/segments/{id}` (dead columns `candidate_segments.human_label/notes/confirmed_segment_id`).
+- **Marketing** `[verified]`: clone `site/public/capture.html` skeleton + brand tokens (`index.html:25-32`); reuse `.honest`/`.nots` components (`:152-158`, used `:351-354`, `:400-415`); add a `/own-model` alias in `site/server.js:99` (`LOCALIZED_PAGES` set at `:41`; Worker `env.ASSETS.fetch` needs no change, `site/src/index.js:84`); cross-link nav `:218-223`, the "gets sharper" section `:327-356`, footer `:528`; screenshots under `site/public/img/` (precedent `site/public/og.png` + `site/make_og.py`).
+- **SPA** `[design, gated]`: today `web/src/App.tsx` renders a first-run `SetupWizard` + the ingest panel, **no router** (`web/package.json`). A correction screen needs react-router-dom (or a hash switch) + the typed client `web/src/api/client.ts`/`types.ts` + `web/src/lib/log.ts`; it realizes P10 in `PER_RALLY_SELECTION.md:122` and depends on a **new** engine endpoint `PATCH /api/segments/{id}` (dead columns `candidate_segments.human_label/notes/confirmed_segment_id`).
- **Annotator content** `[verified]`: embed CSV example + install commands + ending-reason table **verbatim** from the repo; link, don't re-author.
@@ -59,3 +59,3 @@ The worker's own inference CSV uses `start,end,shot_count,ending_reason` (`__mai
- **Minors excluded from any training pool** `[verified CLAUDE.md + live site "Minors are protected"]` ? keep prominent (audience = academies with children).
-- **No paid-LLM laundering**: the model never trains on a paid LLM's output (terms/copyright), attested in the bundle (`harness.py:192`).
+- **No paid-LLM laundering**: the model never trains on a paid LLM's output (terms/copyright), attested in the bundle (`harness.py:236`).
- **Residual**: base-model (Qwen3-VL) pretraining provenance is unauditable `[researched]` ? owner risk-acceptance required before VLM "train your own model" copy goes beyond roadmap.
@@ -70,8 +70,9 @@ Legend: ? Todo · ? In progress · ? Done · ? Gated. One small PR per r
|----|-------------|-----------|--------|--------|----|
-| P0 | This tracked doc in `khelsutra/docs/`, indexed in `docs/README.md` | ? | No | ? |
-| P1 | `site/public/own-model.html` (clone capture.html; 3 sections; honesty fences; placeholder shots) | P0 | No | ? | (this PR) |
-| P2 | `/own-model` alias in `site/server.js` (Worker needs no change) | P1 | No | ? | (this PR) |
-| P3 | Cross-links: nav + footer (the "gets sharper" section link is a nice-to-have, deferred) | P1 | No | ? | (this PR) |
-| P4 | Embed verbatim repo text (CSV/install/ending-reasons) + GitHub link | P1 | No | ? | (this PR) |
-| P5 | Capture + commit real screenshots (consented, no minors, faces redacted);
+| P0 | This tracked doc in `khelsutra/docs/`, indexed in `docs/README.md` | ? | No | ? | [#15](https://github.com/avidullu/khelsutra/pull/15) |
-| P1 | `site/public/own-model.html` (clone capture.html; 3 sections; honesty fences; placeholder shots) | P0 | No | ? | (this PR) |
-| P2 | `/own-model` alias in `site/server.js` (Worker needs no change) | P1 | No | ? | (this PR) |
-| P3 | Cross-links: nav + footer (the "gets sharper" section link is a nice-to-have, deferred) | P1 | No | ? | (this PR) |
-| P4 | Embed verbatim repo text (CSV/install/ending-reasons) + GitHub link | P1 | No | ? | (this PR) |
-| P5 | Capture + commit real screenshots (consented, no minors, faces redacted);

replace placeholders | P1 | ? Owner (footage) | ? | ? |
+| P0 | This tracked doc in `khelsutra/docs/` (doc landed #15; `docs/README.md` index entry still to add) | ? | No | ? | [#15](https://github.com/avidullu/khelsutra/pull/15) |
+| P1 | `site/public/own-model.html` (clone capture.html; 3 sections; honesty fences; placeholder shots) | P0 | No | ? | [#22](https://github.com/avidullu/khelsutra/pull/22) |
+| P2 | `/own-model` alias in `site/server.js` (Worker needs no change) | P1 | No | ? | [#22](https://github.com/avidullu/khelsutra/pull/22) |
+| P3 | Cross-links: nav + footer (the "gets sharper" section link is a nice-to-have, deferred) | P1 | No | ? | [#22](https://github.com/avidullu/khelsutra/pull/22) |
+| P4 | Embed verbatim repo text (CSV/install/ending-reasons) + GitHub link; org-link follow-up | P1 | No | ? | [#22](https://github.com/avidullu/khelsutra/pull/22), [#23](https://github.com/avidullu/khelsutra/pull/23) |
+| P5a | Annotator-UI screenshots (VLC View-menu + annotator dialog) committed, placeholders replaced | P1 | No | ? | [#25](https://github.com/avidullu/khelsutra/pull/25) |
+| P5b | Capture + commit real *footage* screenshots (consented, no minors, faces redacted) | P1 | ? Owner (footage) | ? | ? |
| P6 | Owner: slug + 1-vs-3-page + train-positioning honesty | ? | ? Owner | ? | ? |
@@ -83,3 +84,3 @@ Legend: ? Todo · ? In progress · ? Done · ? Gated. One small PR per r
Blocking gates: (1) Greenlight how prominently to position "train your own model" (CLI + roadmap). (2) Approve the SPA correction screen + the new `PATCH /api/segments/{id}` endpoint. (3) Owner risk-acceptance on Qwen3-VL provenance before VLM copy. (4) Provide consented no-minors footage for the 5 screenshots.
-**Nice-to-knows:** slug (`/own-model` vs `/byo-model` vs `/annotate`); one page vs three; react-router-dom vs hash switch; cross-link `docs/LOCAL_APPLIANCE_ARCHITECTURE.md` once PR #13 merges.
+**Nice-to-knows:** slug (`/own-model` vs `/byo-model` vs `/annotate`); one page vs three; react-router-dom vs hash switch; cross-link `docs/LOCAL_APPLIANCE_ARCHITECTURE.md` (already landed).

@@ -94,3 +95,3 @@ Legend: ? Todo · ? In progress · ? Done · ? Gated. One small PR per r
Annotator `[verified]`: `rally-annotator/README.md`, `vlc/rally_annotator.lua` (HEADER :128; now_seconds :201-207; marks :608-628; save_all :310-335; resolve_out_path :253-267), `docs/CSV_FORMAT.md`, `docs/ENDING_REASONS.md`, `LICENSE` (MIT), `.github/workflows/ci.yml`.
-**Engine `[verified]`:** `backend/worker/__main__.py` (cmd_train :231; _write_intervals_csv :62-74; join :271-280; harness shell-out :321-328), `backend/pipeline/detectors/stub_runner.py` (synth :124-149, :131-133, :139), `training/gen0/harness.py` (gate_verdict :98; label_source :178; attestation :189-192), `backend/eval/gt_loader.py` (:14-90), `docs/SETUP_NEW_MACHINE.md`, `docs/DECOUPLED_COMPUTE/02-COMPUTE-MAP-AND-CONTRACT.md`, `docs/DECOUPLED_COMPUTE/04-BUILD-AND-TEST-PLAN.md`, `docs/OWNED_MODEL_IMPLEMENTATION_PLAN.md`, `CLAUDE.md`.
+**Engine `[verified]`:** `backend/worker/__main__.py` (cmd_train :332; _write_intervals_csv :72-85; join :383-384; harness shell-out :424-432), `backend/pipeline/detectors/stub_runner.py` (synth :124-149, :137, :139), `training/gen0/harness.py` (gate_verdict :142; label_source :222; attestation :233-237), `backend/eval/gt_loader.py` (load_gt_intervals :44; _START_NAMES :30), `docs/SETUP_NEW_MACHINE.md`, `docs/DECOUPLED_COMPUTE/02-COMPUTE-MAP-AND-CONTRACT.md`, `docs/DECOUPLED_COMPUTE/04-BUILD-AND-TEST-PLAN.md`, `docs/OWNED_MODEL_IMPLEMENTATION_PLAN.md`, `CLAUDE.md`.
KhelSutra `[verified]`: `site/public/index.html`, `site/public/capture.html`, `site/server.js`, `site/src/index.js`, `web/src/App.tsx`, `web/package.json`, `docs/README.md`, `docs/PER_RALLY_SELECTION.md`.
@@ -98,2 +99,3 @@ Legend: ? Todo · ? In progress · ? Done · ? Gated. One small PR per r
Changelog
+- 2026-06-22 ? Freshness pass: flipped Status DRAFT?IN PROGRESS (P0?P4 shipped #15/#22/#23/#25); cited real PRs in §7 (was "(this PR)"); split P5 into P5a (annotator-UI shots ? #25) / P5b (footage shots ? owner); corrected the "indexed in `docs/README.md`" claims (P0 ? ?, still to add); dropped the "lands with PR #13" future-tense (LOCAL_APPLIANCE_ARCHITECTURE.md already landed); reworded the SPA "one static screen" line for the post-#88 SetupWizard; refreshed drifted engine line citations (cmd_train :332, gate_verdict :142, label_source :222, attestation :233-237, server.js alias :99, OG asset paths).
- 2026-06-20 ? Initial DRAFT consolidating annotator + train-loop + UI-surface research into one tracked doc (multi-agent design pass; schema-bridge + no-images + train-guardrails re-verified against source before writing).
diff --git a/docs/LOCAL_APPLIANCE_ARCHITECTURE.md b/docs/LOCAL_APPLIANCE_ARCHITECTURE.md

```
index eefbfdc..ed7cbf6 100644
--- a/docs/LOCAL_APPLIANCE_ARCHITECTURE.md
+++ b/docs/LOCAL_APPLIANCE_ARCHITECTURE.md
@@ -2,3 +2,3 @@

-> **Status:** `DRAFT` ? owner-gated (auth choice, retention, cost ceiling, GPU host, P5 bridge) · **Owner:** Avi · **Created:** 2026-06-20 · **Last updated:** 2026-06-21
+> **Status:** `IN PROGRESS` ? owner-gated (auth choice, retention, cost ceiling, GPU host, P5 bridge) · **Owner:** Avi · **Created:** 2026-06-20 · **Last updated:** 2026-06-22
 > **Lifecycle:** `DRAFT ? IN PROGRESS ? DONE ? archived` (move to `docs/archives/` when DONE)
@@ -13,3 +13,3 @@

-The **single most important caveat:** the engine has **zero app-layer auth and CORS `*`** `[verified backend/main.py:42]`, so safety is structural **only** because the engine stays bound to `127.0.0.1:8000` `[verified backend/main.py:638]`, a firewall blocks inbound `8000`, and a reverse proxy/tunnel is the sole authenticated ingress ? **the gate, not the app, is the tenancy boundary.**
+The **single most important caveat:** the engine's CORS default is `*` but is now **env-configurable** with `allow_credentials=False` `[verified backend/main.py::_cors_origins_from_env, origin/master]`, and it now ships a **Host-header guard** (`TrustedHostMiddleware`) `[verified backend/main.py, origin/master]` + an **optional edge-auth integration** (verifies the `Cf-Access-Jwt-Assertion` JWT) `[verified backend/api/auth.py, origin/master]` + an **exposure-safety boot posture** (`enforce_startup_checks`) `[verified backend/config/startup.py, origin/master]` ? so "zero app-layer auth" is **no longer literally true**. Even so, safety is still structural **only** because the engine stays bound to `127.0.0.1:8000` `[verified backend/main.py:638]`, a firewall blocks inbound `8000`, and a reverse proxy/tunnel is the sole authenticated ingress ? **the gate, not the app, is the tenancy boundary.**

@@ -44,3 +44,3 @@ The decoupled engine just shipped a `StorageBackend` ABC + a `python -m backend.
 - **Storage:** `storage.backend` default = `"local"` `[verified models.py:298]`; `LocalStorage.fetch_to_local` is **identity** `[verified local.py:33]`, `put` no-ops on `src==dst` ? so on a single box the worker's pull?pipeline?push **collapses to read-local/write-local**. URI scheme `parse_uri` `[verified ref.py]`: bare/`local://` ? local; `s3://` (AWS/R2/**DO Spaces**/MinIO, by endpoint) ; `gcs://`/`gs://` ; `gdrive://` ? **follow-up stub (emulated only)**. Credentials never in config ? boto3 default chain / `~/.aws` `[verified s3.py]`.
-- **Compute seam:** `indexing.detector_impl` ? {`auto`=WSL, `native`=Linux-native, `stub`=CPU mock}; default `"auto"` `[verified models.py:129]`, resolved in `trajectory_hybrid._resolve_runner` `[verified :76-85]`. **Native WASB is in-flight on branch `feat/native-wasb-runner`:** `native_wasb_runner.py` **exists** `[verified]`; **`fusion_hybrid._build_native_runner` returns `NativeWasbRunner`** `[verified fusion_hybrid.py:86]` while **`trajectory_hybrid._build_native_runner` still raises** `[verified :57-63]`. Acceptance gate = **owner byte-parity** (WSL CSV == native CSV) on a real GPU box ? not exercisable on the droplet or in the offline suite. Treat native as **not production-verified**.
+- **Compute seam:** `indexing.detector_impl` ? {`auto`=WSL, `native`=Linux-native, `stub`=CPU mock}; default `"auto"` `[verified models.py:129]`, resolved in `trajectory_hybrid._resolve_runner` `[verified :76-85]`. **Native WASB `native_wasb_runner.py` is now on `origin/master`** (no longer branch-only; `tests/test_native_wasb_runner.py` is on master too) `[verified git ls-tree origin/master]`; **`fusion_hybrid._build_native_runner` returns `NativeWasbRunner`** `[verified fusion_hybrid.py:86, origin/master]` while **`trajectory_hybrid._build_native_runner` still raises `NotImplementedError`** `[verified trajectory_hybrid.py:61, origin/master]`. Acceptance gate = **owner byte-parity** (WSL CSV == native CSV) on a real GPU box ? not exercisable on the droplet or in the offline suite. Treat native as **not production-verified**.
 - **Worker CLI:** `python -m backend.worker {infer|train}` pulls a `--video-uri` ? pipeline ? pushes `outputs/<video_id>/{intervals.csv,report.json}` `[verified worker/__main__.py:88,149]`. The validated GPU-free/key-free combo: `detector_impl=stub` + `--set indexing.skip_ai_handoff=true` + `storage.s3`.

@@ -48,3 +48,4 @@ The decoupled engine just shipped a `StorageBackend` ABC + a `python -m backend.
 - **Droplet reality:** `168.144.126.176`, **CPU-only, NO GPU**, region `blr1`; co-hosts
```

the marketing site (`/srv/khelsutra-site`, node `:8787`) **and** the engine (`~/rally`); ran the full train+infer e2e against DO Spaces with the **GPU mocked** `[SETUP_NEW_MACHINE.md, DECOUPLED_COMPUTE/HANDOFF.md]`. Shared testbed: MinIO `9000/9001` + GCS emulator `9023` also run there.

-- **Absent (net-new):** `GET /api/capabilities`, `GET /api/videos` (only single-row `get_video(id)`), `GET /api/reels/{video_id}`, any source/proxy streaming endpoint, any **Dockerfile**, any per-video **cost ledger** `[verified ? grep returns none]`.

+- **Shipped since this doc's first draft:** `GET /api/capabilities` now **exists on `origin/master`** `[verified backend/main.py:340, origin/master]` (alongside `/api/health` :307 + `/api/version` :334).

+- **Absent (net-new):** `GET /api/videos` (only single-row `get_video(id)`), `GET /api/reels/{video_id}`, any source/proxy streaming endpoint, any **Dockerfile**, any per-video **cost ledger** `[verified ? grep returns none]`.

@@ -95,3 +96,3 @@ All sources normalize to **"a video the pipeline can read"** ? a local path (M

| **CPU droplet ? motion** | `segmenter=motion` | nothing | REAL, low-quality | only no-key/no-GPU real segmenter; never surface `server/receiver/events` (fabricated) |

-| **GPU host ? native WASB** (in-flight) | `wasb`+`detector_impl=native` ? `NativeWasbRunner` (in-proc torch) | a **GPU host ? DO GPU droplet / GCE GPU VM / Cloud Run+GPU**; `WASB_WEIGHTS` | **REAL**, on-device | partially wired (fusion builds it, trajectory refuses); **owner byte-parity gate**; **Cloud Run/GCE need a containerized worker ? no Dockerfile exists, and WASB's torch-1.11 conda env makes the image non-trivial** |

+| **GPU host ? native WASB** (`native_wasb_runner.py` on master; owner-parity-gated) | `wasb`+`detector_impl=native` ? `NativeWasbRunner` (in-proc torch) | a **GPU host ? DO GPU droplet / GCE GPU VM / Cloud Run+GPU**; `WASB_WEIGHTS` | **REAL**, on-device | partially wired (fusion builds it, trajectory refuses); **owner byte-parity gate**; **Cloud Run/GCE need a containerized worker ? no Dockerfile exists, and WASB's torch-1.11 conda env makes the image non-trivial** |

| **Owner Win+WSL ? auto** | `wasb`+`detector_impl=auto` (WSL2+conda) | Windows+WSL2+GPU | REAL (today) | Windows-only transport; the working real-GPU path for owner verification |

@@ -102,3 +103,3 @@ All sources normalize to **"a video the pipeline can read"** ? a local path (M

4.4 Auth / exposure (non-negotiable invariant)

-No engine auth, by design; the gate is external. **Invariant (all three must hold):** engine bound to `127.0.0.1` + firewall closes `8000` + edge gate up. Go-live checklist `[HOSTING_PLAN.md:102-108]`: Access/basicauth ON; **rotate the chat-shared `GEMINI_API_KEY`**; set `security.allowed_video_root`; rate-limit; per-user quota.

+Auth lives at the gate, by design *(the engine now also ships an **optional** in-app edge-auth check that verifies the `Cf-Access-Jwt-Assertion` JWT ? `backend/api/auth.py` on `origin/master` ? but the external gate remains the boundary here)*. **Invariant (all three must hold):** engine bound to `127.0.0.1` + firewall closes `8000` + edge gate up. Go-live checklist `[HOSTING_PLAN.md:102-108]`: Access/basicauth ON; **rotate the chat-shared `GEMINI_API_KEY`**; set `security.allowed_video_root`; rate-limit; per-user quota.

@@ -112,3 +113,3 @@ No engine auth, by design; the gate is external. **Invariant (all three must hol

4.6 Reuse map

-**Reused (no re-plumb):** the full `/api` loop; the in-process worker; the chunked upload endpoints (un-called); LocalStorage default; the detector seam incl. the landing `NativeWasbRunner`; the **RallyPicker** design + anchors; the `web/` scaffold; the `HOSTING_PLAN` gate; the `SETUP_NEW_MACHINE` runbook. **Net-new:** Caddyfile + systemd units; `GET /api/capabilities`; `GET /api/videos`; `GET /api/reels/{video_id}`; `GET /api/health`; the SPA console screens; ingest-from-bucket wiring; a retention sweep; (reserved) the bucket-worker bridge.

+**Reused (no re-plumb):** the full `/api` loop; the in-process worker; the chunked upload endpoints (un-called); LocalStorage default; the detector seam incl. the landing `NativeWasbRunner`; the **RallyPicker** design + anchors; the `web/` scaffold; the `HOSTING_PLAN` gate; the `SETUP_NEW_MACHINE` runbook. **Net-new:** Caddyfile + systemd units; `GET /api/videos`; `GET /api/reels/{video_id}`; the SPA console screens; ingest-from-bucket wiring; a retention sweep; (reserved) the bucket-worker bridge. *(`GET /api/capabilities` and `GET /api/health` have since shipped on engine `origin/master` ? see §3.)*

@@ -127,3 +128,3 @@ Launch-readiness directive (owner 2026-06-20): **logs and monitoring

taken serio
|---|---|---|
-| No app auth + CORS `*` `[main.py:42]` | open if any of {localhost-bind, firewall, gate} slips | enforce all three; **LIVE posture test** that raw `:8000` is refused from outside |
+| CORS default `*` (now env-lockable, `allow_credentials=False`) + auth lives at the edge `[main.py::_cors_origins_from_env + api/auth.py, origin/master]` | open if any of {localhost-bind, firewall, gate} slips | enforce all three; **LIVE posture test** that raw `:8000` is refused from outside |
| Test ports on shared droplet | MinIO `9000/9001` + GCS emu `9023` | firewall closes them |
@@ -132,3 +133,3 @@ Launch-readiness directive (owner 2026-06-20): **logs and monitoring taken serio
| Native WASB unverified | partially wired; owner-parity-gated | treat as not-production-verified until owner byte-parity on a real GPU box |
-| New engine endpoints | `capabilities`/`videos`/`reels`/`health` don't exist | coordinate engine PRs with active sessions; **branch from `origin/master`, re-verify HEAD, explicit `git add`** |
+| New engine endpoints | `capabilities`/`health` **now exist on `origin/master`**; `videos`/`reels` still don't | coordinate engine PRs with active sessions; **branch from `origin/master`, re-verify HEAD, explicit `git add`** |
| Drive ingest overstated | `gdrive://` is a stub | don't present Drive as live; require Drive?bucket pre-step until real |
@@ -163,3 +164,3 @@ Legend: ? Todo · ? In progress · ? Done · ? Blocked/gated. **One smal
| T6 | S3 embed **RallyPicker** + compile + download | T4 | no | ? | ? |
-| T7 | *(engine)* `GET /api/capabilities` + `GET /api/health` + banner | T1 | coord | ? | ? |
+| T7 | *(engine)* `GET /api/capabilities` + `GET /api/health` + banner ? **engine half shipped on `origin/master`** (`/api/capabilities` main.py:340, `/api/health` :307); only the SPA banner remains | T1 | coord | ? | engine (on master) |
| T8 | *(engine)* `GET /api/videos` + `GET /api/reels/{video_id}` (S4 reload-survival) | ? | coord | ? | ? |
@@ -174,3 +175,3 @@ Legend: ? Todo · ? In progress · ? Done · ? Blocked/gated. **One smal

-**? Alpha deploy path chosen = SPLIT, not the single-origin shell (owner 2026-06-21).** The alpha ships the **split** topology (SPA on Cloudflare Pages + engine on an `api.*` cloudflared tunnel + CF Access over both ? the engine `HOSTING_PLAN.md` §2), **not** the single-origin Caddy shell of **T2** (kept as the documented alternative ? §0/D1). The **deploy kit + runbook** ([`deploy/DEPLOY_ALPHA.md`](../deploy/DEPLOY_ALPHA.md) + `deploy/cloudflared.config.example.yml`) is shipped (? templates+runbook); the actual stand-up (tunnel + Pages + CF Access + key rotation) is the **owner** step. The **T3** raw-`:8000`-refused posture test is documented as the deploy kit's §6 verification. Enabled by: engine exposure-safety boot posture (engine #294) + SPA `VITE_API_BASE` ([#64]).
+**? Alpha deploy path chosen = SPLIT, not the single-origin shell (owner 2026-06-21).** The alpha ships the **split** topology (SPA on Cloudflare Pages + engine on an `api.*` cloudflared tunnel + CF Access over both ? the engine `HOSTING_PLAN.md` §2), **not** the single-origin Caddy shell of **T2** (kept as the documented alternative ? §0/D1). The **deploy kit + runbook** ([`deploy/DEPLOY_ALPHA.md`](../deploy/DEPLOY_ALPHA.md) + `deploy/cloudflared.config.example.yml`) is shipped (? templates+runbook). **The alpha stand-up is now LIVE (2026-06-22):** the engine booted on the droplet (`168.144.126.176`), `app.khelsutra.guru` is up, and the full **DNS ? Cloudflare ? Access ? tunnel ? engine** chain plus the **T3** raw-`:8000`-refused posture test (boot-posture fail-fast, Host-guard `400?200`, path-jail, T4 upload?jail, edge-auth `401`-gate) were all **exercised live** `[deploy/droplet/README.md:100-106]`. Only the **Gemini-key rotation** + a **browser OTP login** remain. Enabled by: engine exposure-safety boot posture (engine #294) + SPA `VITE_API_BASE` ([#64]).

@@ -181,4 +182,4 @@ Legend: ? Todo · ? In progress · ? Done · ? Blocked/gated. **One smal
11. Definition of done
-- [ ] CUJ-1 works end to end on the droplet (Gemini **and** stub) behind the gate, via the ingest menu.
-- [ ] Raw `:8000`/`:900x` unreachable from outside (**tested**, not assumed).

143   +- [~] CUJ-1 works end to end on the droplet (Gemini **and** stub) behind the gate, via the ingest menu. *(2026-06-22: chain LIVE end-to-end ? DNS?Access?tunnel?engine, T4 upload?jail exercised; the full upload?reel pass still needs the Gemini key + a browser OTP login `[deploy/droplet/README.md:100-106]`.)*
144   +- [x] Raw `:8000`/`:900x` unreachable from outside (**tested** live 2026-06-22, not assumed) `[deploy/droplet/README.md:100-106]`.
145   - [ ] Local upload wired; bucket-URI ingest at least designed/tracked; Drive scoped honestly (not faked).
146   @@ -195,2 +196,3 @@ Legend: ? Todo · ? In progress · ? Done · ? Blocked/gated. **One smal
147   ### Changelog
148   +- 2026-06-22 ? Freshness pass: `DRAFT`?`IN PROGRESS`; re-anchored engine facts against `origin/master` ? `/api/capabilities` + `/api/health` now exist (T7 engine half ?); `native_wasb_runner.py` is on master (no longer branch-only); CORS `*` is now env-lockable + the engine ships a Host-guard / optional Cf-Access edge-auth / boot-posture (so "zero app-layer auth" is no longer literal); alpha deploy went LIVE 2026-06-22 (`app.khelsutra.guru`), ticking the raw-`:8000`-refused DoD box.
149   - 2026-06-20 ? Initial `DRAFT`. Consolidated three single-box-appliance designs (FastAPI-native / unified-origin / decoupled-profile) into the reverse-proxied unified-origin spine + compute-as-a-setting framing, via a multi-agent design pass; every engine-fact anchor re-verified against source before writing (catching the native-WASB partial-wiring + `ai_max_workers` default nuances). Folded in the owner's ingest-menu (Drive/bucket/local), multi-provider GPU, and first-class Observability directives.
150   diff --git a/docs/PER_RALLY_SELECTION.md b/docs/PER_RALLY_SELECTION.md
151   index 25e6024..1afae68 100644
152   --- a/docs/PER_RALLY_SELECTION.md
153   +++ b/docs/PER_RALLY_SELECTION.md
154   @@ -5,3 +5,3 @@
155   > **Tracking anchors:** §7 progress tracker is the source of truth; indexed in `docs/README.md`; pointer in `NEXT_STEPS.md` + memory `current-state` once work starts.
156   -> **Relation:** second major tracked work item in the `khelsutra` repo (peer of `REQUEST_ACCESS_FORM.md`); mirrors the engine repo's `PROJECT_DOC_TEMPLATE` discipline. Primarily realizes alpha screen **A5 Highlights** (engine repo `UI_PROPOSAL.md:47`, exec-plan **M4** `UI_ALPHA_EXECUTION_PLAN.md:39`); its correction-loop extension realizes **A4 Rally Review** (`UI_PROPOSAL.md:46`, exec-plan **M3** `:38`). Consumes the engine's `/api/query` + `/api/compile` contract.
157   +> **Relation:** second major tracked work item in the `khelsutra` repo (peer of `docs/REQUEST_ACCESS_FORM.md`); mirrors the engine repo's `PROJECT_DOC_TEMPLATE` discipline. Primarily realizes alpha screen **A5 Highlights** (engine repo `UI_PROPOSAL.md:47`, exec-plan **M4** `UI_ALPHA_EXECUTION_PLAN.md:39`); its correction-loop extension realizes **A4 Rally Review** (`UI_PROPOSAL.md:46`, exec-plan **M3** `:38`). Consumes the engine's `/api/query` + `/api/compile` contract.
158   > **Honesty note:** every engine-fact claim is tagged `[verified]` (read against engine code this session), `[design]` (proposed here), or `[researched]` (needs sign-off). Not legal/financial advice.
159   @@ -11,3 +11,3 @@
160   ## 0. TL;DR
161   -Build the **pick-rallies ? reel** surface as a **hybrid**: a selectable rally **list/grid** is the workhorse; a **deterministic query bar** ("rallies over 10s, longest first") pre-selects into the **same** selection set; a sticky footer compiles the selection into one downloadable MP4. The entire MVP runs on **existing engine endpoints** ? `POST /api/query` ? `POST /api/compile` ? `GET /api/highlights` `[verified backend/main.py:331-340, 342-396, 307-328]` ? with **zero new engine intelligence and zero new endpoints**. The single most important caveat: queries and chips are scoped to **length / count / order / shot-count only**; shot-type, player, and score are **roadmap-only** (no detector, no column) and must not be faked. The biggest UX-vs-effort tension: **no per-rally preview is possible today** ? only the *final compiled reel* can stream ? so MVP selects on text metadata, and visual preview/timeline is a deliberate later milestone.
162   +Build the **pick-rallies ? reel** surface as a **hybrid**: a selectable rally **list/grid** is the workhorse; a **deterministic query bar** ("rallies over 10s, longest first") pre-selects into the **same** selection set; a sticky footer compiles the selection into one downloadable MP4. The entire MVP runs on **existing engine endpoints** ? `POST /api/query` ? `POST /api/compile` ? `GET /api/highlights` `[verified backend/main.py::{query_play_segments, compile_highlights, download_highlight}]` ? with **zero new engine intelligence and zero new endpoints**. The single most important caveat: queries and chips are scoped to **length / count

/ order / shot-count only**; shot-type, player, and score are **roadmap-only** (no detector, no column) and must not be faked. The biggest UX-vs-effort tension: **no per-rally preview is possible today** ? only the *final compiled reel* can stream ? so MVP selects on text metadata, and visual preview/timeline is a deliberate later milestone.

@@ -40,8 +40,8 @@ The whole design rests on the engine's real, already-shipped contract. Read agai

-- **The rally store is `play_segments`** ? columns `id, video_id, start_time, end_time, duration, server, receiver, metadata` `[verified database.py:124-135]`. `id` is the integer the UI selects and passes to compile; `start_time/end_time/duration` are REAL seconds (duration computed at insert).

-- **Listing rallies = `POST /api/query {video_id}` ? `{play_segments:[...]}`** ? the *only* rally-list path today; it JSON-parses `metadata` and joins `events` `[verified main.py:331-340, database.py:427-475]`.

-- **Building a reel = `POST /api/compile {video_id, segment_ids:[int], output_filename}`** ? loads `videos.filepath`, filters all segments to the requested `segment_ids`, applies the `min_shot_count` floor, extracts `time_pairs=[(start,end)]`, and calls `VideoCompiler.compile_highlights(raw_path, segments, name)` `[verified main.py:342-396, compiler.py:303]`, which **merges overlaps** then per-clip re-encodes ? concat ? one MP4 `[verified compiler.py:311-317]`. Selection order/duplicates are therefore safe.

-- **Downloading = `GET /api/highlights/{video_id}/{filename}`** ? streams the compiled MP4. Compile returns a **server-local filepath, NOT a URL** `[verified main.py:394]`, so the SPA constructs this download URL itself. The compiled reel is the **only streamable video that exists today** `[verified main.py:307-328]`.

-- **The floor knob is readable** ? `GET /api/config` exposes `stitching.min_shot_count` `[verified main.py:121-123]`, so the UI can warn before a selected rally is silently dropped.

-- **Job entry** ? `GET /api/jobs/{job_id}` returns `result.video_id` `[verified main.py:274-293]`, which is how the MVP reaches a video without a (nonexistent) list-videos endpoint.

+- **The rally store is `play_segments`** ? columns `id, video_id, start_time, end_time, duration, server, receiver, metadata` `[verified database.py CREATE TABLE play_segments]`. `id` is the integer the UI selects and passes to compile; `start_time/end_time/duration` are REAL seconds (duration computed at insert).

+- **Listing rallies = `POST /api/query {video_id}` ? `{play_segments:[...]}`** ? the *only* rally-list path today; it JSON-parses `metadata` and joins `events` `[verified main.py::query_play_segments, database.py::query_play_segments]`.

+- **Building a reel = `POST /api/compile {video_id, segment_ids:[int], output_filename}`** ? loads `videos.filepath`, filters all segments to the requested `segment_ids`, applies the `min_shot_count` floor, extracts `time_pairs=[(start,end)]`, and calls `VideoCompiler.compile_highlights(raw_path, segments, name)` `[verified main.py::compile_highlights, compiler.py::compile_highlights]`, which **merges overlaps** then per-clip re-encodes ? concat ? one MP4 `[verified compiler.py::_merge_overlapping]`. Selection order/duplicates are therefore safe.

+- **Downloading = `GET /api/highlights/{video_id}/{filename}`** ? streams the compiled MP4. Compile now returns **both** a server-local `filepath` **and** a ready-to-use `download_url` (`/api/highlights/...`) `[verified main.py::compile_highlights]`, so the SPA may use `download_url` directly; the hand-built-URL fallback is now optional, not required. The compiled reel is the **only streamable video that exists today** `[verified main.py::download_highlight]`.

+- **The floor knob is readable** ? `GET /api/config` exposes `stitching.min_shot_count` `[verified main.py::get_config]`, so the UI can warn before a selected rally is silently dropped.

+- **Job entry** ? `GET /api/jobs/{job_id}` returns `result.video_id` `[verified main.py::get_job]`, which is how the MVP reaches a video; a list-videos endpoint (`GET /api/videos`) has since shipped (§4.3).

@@ -51,3 +51,3 @@ The whole design rests on the engine's real, already-shipped contract. Read agai

- **Roadmap ? no column, no detector:** shot-type per shot, player identity, match/point score, game/set grouping. None exist anywhere in the schema; do not consume them `[verified database.py whole-schema 122-148]`.

-- **Correction-loop columns exist but are dead:** `candidate_segments.human_label/notes/confirmed_segment_id` are defined but no endpoint writes them `[verified database.py:166-168]` ? that's the M4 corpus extension (engine PR-4/B6,

```
`UI_ALPHA_EXECUTION_PLAN.md`).
182     +- **Correction-loop columns exist but are dead:**
`candidate_segments.human_label/notes/confirmed_segment_id` are defined but no endpoint writes
them `[verified database.py CREATE TABLE candidate_segments]` ? that's the M4 corpus extension
(engine PR-4/B6, `UI_ALPHA_EXECUTION_PLAN.md`).
183
184     @@ -72,3 +72,3 @@ Gemini segmenter (process time) writes `play_segments {id,
start/end_time, durat
185      ? **Build** ? `POST /api/compile {video_id, segment_ids: [...selectedIds],
output_filename}` ? stitcher de-dupes via `_merge_overlapping` `[compiler.py:311-317]`, cuts
`time_pairs` from `videos.filepath`, re-encodes per clip ? concat ? one MP4 `[main.py:386-393,
compiler.py:303]`
186     -? compile returns a **server-local filepath, not a URL** `[main.py:394]` ? SPA builds
`GET /api/highlights/{video_id}/{output_filename}` to preview in `<video>` + download
`[main.py:307-328]`.
187     +? compile returns **both** a server-local `filepath` **and** a ready-to-use
`download_url` (`/api/highlights/...`) `[main.py::compile_highlights]` ? the SPA may use
`download_url` directly (or build `GET /api/highlights/{video_id}/{output_filename}` itself) to
preview in `<video>` + download `[main.py::download_highlight]`.
188
189     @@ -82,6 +82,6 @@ Gemini segmenter (process time) writes `play_segments {id,
start/end_time, durat
190      | `GET /api/jobs/{job_id}` | Carry `result.video_id` into the picker | **exists** |
191     -| `POST /api/compile` ? add `download_url` | Return `/api/highlights/...` so SPA stops
hand-building it (additive, low-risk; today returns only a filepath `main.py:394`) | **modify**
|
192     -| `GET /api/health` | Engine-offline banner ping the SPA is specified to use
(`HOSTING_PLAN.md:42`) | **new** |
193     +| `POST /api/compile` ? add `download_url` | Returns `/api/highlights/...` so the SPA
needn't hand-build it (compile now returns both `filepath` **and** `download_url` `[verified
main.py::compile_highlights]`) | **exists** |
194     +| `GET /api/health` | Engine-offline banner ping the SPA is specified to use
(`HOSTING_PLAN.md:42`) ? shipped `[verified main.py::health]` | **exists** |
195      | `GET /api/videos/{video_id}/rallies` | RESTful/cacheable alias returning the same
`{play_segments}` shape | **new** |
196     -| `GET /api/videos` | List a user's videos for a Match picker (only `get_video(id)`
exists, `database.py:418`) ? needs DB enumerate + identity scoping | **new** |
197     +| `GET /api/videos` | List videos for a Match picker ? the enumerate endpoint shipped
`[verified main.py::list_videos, database.py::list_videos]`; only identity/owner-scoping remains
gated | **exists** |
198      | `GET /api/videos/{video_id}/source` (Range) *or* `/clip?start=&end=` | Stream
source/per-span video so a rally previews before selection ? neither exists | **new** |
199     @@ -115,3 +115,3 @@ No silent drops: unsupported concepts (`query.unavailable_request` +
`dropped_un
200      | No auth / user scoping anywhere (CORS `*`, `videos.user_id` never set/filtered) |
High for multi-tester | Fine for single-box local alpha; a ship-blocker for any shared host ?
gate `GET /api/videos` behind PR-3 scoping |
201
```

676. assistant (2026-06-22T09:53:48.263Z)

Careful, durable work ? the `PER_RALLY` fixer converted brittle line anchors to **symbol references** (`main.py::compile_highlights`), and the honesty discipline held (P6 ? `?` engine-done/SPA-pending, not a false `?`). Before I trust the symbol references and the remaining engine claims (auth.py, startup.py, harness line anchors), let me verify those symbols resolve on engine `origin/master`, and read the `site/README` + `web/README` diffs.

677. user (2026-06-22T09:53:54.492Z)

```
=== PER_RALLY symbol refs resolve? (main.py + database.py + compiler.py) ===
origin/master:backend/main.py:298:def get_config():
origin/master:backend/main.py:308:def health():
origin/master:backend/main.py:375:def list_videos(limit: Optional[int] = None):
origin/master:backend/main.py:638:def get_job(job_id: str):
origin/master:backend/main.py:671:def download_highlight(video_id: str, filename: str):
```

```
origin/master:backend/main.py:695:def query_play_segments(req: QueryRequest):
origin/master:backend/main.py:706:def compile_highlights(req: CompileRequest):
origin/master:backend/infrastructure/database.py:545:    def list_videos(self, limit:
Optional[int] = None) -> List[Dict[str, Any]]:
origin/master:backend/infrastructure/database.py:564:    def query_play_segments(self, video_id:
str, filters: Dict[str, Any]) -> List[Dict[str, Any]]:
origin/master:backend/pipeline/stitcher/compiler.py:83:    def _merge_overlapping(segments:
List[Tuple[float, float]]) -> List[Tuple[float, float]]:
origin/master:backend/pipeline/stitcher/compiler.py:305:    def compile_highlights(self,
raw_video_path: str, segments: List[Tuple[float, float]], output_filename: str) -> str:
=== LOCAL_APPLIANCE: auth.py + startup.py symbols ===
origin/master:backend/api/auth.py:5:``Cf-Access-Jwt-Assertion`` header is **required** and
**verified** (RS256) against the team JWKS ?
origin/master:backend/api/auth.py:23:CF_ACCESS_HEADER = "Cf-Access-Jwt-Assertion"
origin/master:backend/config/startup.py:63:    """Raised by ``enforce_startup_checks`` when an
active profile fails a FATAL check."""
origin/master:backend/config/startup.py:253:def enforce_startup_checks(
=== ANNOTATE: gate_verdict + cmd_train anchors ===
origin/master:training/gen0/harness.py:142:def gate_verdict(eval_res: dict, floor:
Optional[dict]) -> dict:
origin/master:training/gen0/harness.py:196:    verdict = gate_verdict(eval_res, floor)
origin/master:backend/worker/__main__.py:332:def cmd_train(args: argparse.Namespace) -> int:

=== site/README.md + web/README.md diffs ===
diff --git a/site/README.md b/site/README.md
index a793400..ca8f9fb 100644
--- a/site/README.md
+++ b/site/README.md
@@ -6,5 +6,11 @@ The public site at **[khelsutra.guru](https://khelsutra.guru)** ? a static lan
 ## Layout
-- `public/` ? static assets served as-is (`index.html`, `capture.html`, `og.png`,
`robots.txt`).
-- `src/index.js` ? the Worker: serves `public/` and handles `POST /api/request-access` ? D1.
-- `src/index.test.js` ? unit tests for the form logic.
+- `public/` ? static assets served as-is: `index.html`, `capture.html`, `own-model.html`, the
+  flyers (`flyer.html` / `flyer-kn.html` / `flyer-hi.html`), `og.png`, `robots.txt`, `img/`,
+  `video/`, and `locales/` (6 languages: en/kn/hi/es/da/id) with `i18n.js`.
+- `src/index.js` ? the Worker handler: handles `POST /api/request-access` ? D1 and delegates
all
+  other paths to the Cloudflare static-assets binding (`ASSETS`).
+- `src/store.js` ? storage adapters (D1 for the Worker, `node:sqlite` for the VM host).
+- `src/notify.js` ? notification adapters (Email Routing send binding on Cloudflare; no-op on
the VM).
+- `src/*.test.js` / `src/layout.browser.mjs` ? the test suite (form logic, VM host, assets,
+  deploy-config, i18n/locales-drift, DOM/UX, browser layout).
 - `wrangler.toml` ? Worker + static-assets + D1 binding config.
@@ -12,2 +18,11 @@ The public site at **[khelsutra.guru](https://khelsutra.guru)** ? a static
lan
 - `make_og.py` ? regenerates `public/og.png` (Pillow + Windows fonts): `python make_og.py`.
+- `server.js` ? an always-on VM host (dependency-free, `node:sqlite`) that runs the **same**
+  `handleRequest` from `src/index.js`; it's the portability fallback "if Cloudflare acts out"
+  and backs the alpha droplet. See `docs/PORTABILITY.md`.
+
+## Two runtimes, one handler
+The request-access logic in `src/index.js` is runtime-agnostic: it runs both as the
**Cloudflare
+Worker** (this README's focus; D1 + Email Routing) and on an **always-on VM host**
(`server.js`;
+`node:sqlite`, no email yet). All storage/notify coupling lives in `src/store.js` /
`src/notify.js`.
+See `docs/PORTABILITY.md`.

@@ -15,3 +30,3 @@ The public site at **[khelsutra.guru](https://khelsutra.guru)** ? a static lan
 ```
-npm test   # node --test src/ (no install needed)
```
```

```
+npm test       # node --test (86 tests, no install needed)
 npm run dev    # npx wrangler dev ? local Worker + D1
diff --git a/web/README.md b/web/README.md
index 6ac4d9e..99c42f6 100644
--- a/web/README.md
+++ b/web/README.md
@@ -24,3 +24,4 @@ npm run build      # vite build
 ## Layout
-- `src/api/client.ts` ? typed engine client (`listRallies` / `compileReel` / `highlightUrl` /
`getConfig`),
+- `src/api/client.ts` ? typed engine client (`listRallies` / `compileReel` / `highlightUrl` /
`getConfig`,
+  plus `processVideo` / `getJob` / `getCapabilities` / chunked `uploadFile`),
   grounded in the verified engine contract; every call carries a correlation id.
@@ -29,2 +30,11 @@ npm run build      # vite build
 - `src/lib/log.ts` ? structured client logging + correlation ids (observability seed, appliance
§5).
-- `src/App.tsx` ? the console shell (screens land in M2+).
+- `src/components/` ? the screen components: per-rally selection (`RallyGrid` / `QueryBar` /
+  `TapFilters` / `SelectionFooter`), in-UI ingest (`IngestPanel`), the P2c first-run wizard
+  (`SetupWizard`), `LanguageSwitcher`, `CompileResult`, `StatusChip`.
+- `src/hooks/` ? `useRallySelection`, `useIngestJob` (process ? poll job), `useUpload` (chunked
upload).
+- `src/i18n/` ? the 6-language catalog (en/kn/hi/es/da/id; `index.ts` `SUPPORTED_LOCALES`).
+- `src/lib/` ? pure helpers: `query` (parse the rally query bar), `selection`, `floor`,
`format`,
+  `errors` / `errorText` (alongside `log.ts` above).
+- `src/App.tsx` ? the operator console: per-rally selection, in-UI ingest (`IngestPanel`), the
P2c
+  first-run wizard (`SetupWizard`) and reel compile, in 6 languages. Drives the engine via
+  `listRallies` / `compileReel`.
```

678. assistant (2026-06-22T09:54:34.195Z)

Everything verifies cleanly against engine `origin/master`:
- **PER_RALLY symbol refs** all resolve: `main.py::{get_config:298, health:308, list_videos:375, get_job:638, download_highlight:671, query_play_segments:695, compile_highlights:706}`, `database.py::{list_videos, query_play_segments}`, `compiler.py::{_merge_overlapping, compile_highlights}`.
- **auth.py:23** `CF_ACCESS_HEADER = "Cf-Access-Jwt-Assertion"` and **startup.py:253** `enforce_startup_checks` ? confirm the LOCAL_APPLIANCE auth rewrite.
- **harness.py:142** `gate_verdict` and **__main__.py:332** `cmd_train` ? confirm the ANNOTATE re-anchoring is accurate.
- `site/README` and `web/README` now describe the real shipped surface.

No errors, no over-claims (status flips keep honest caveats, P6 stays `?`, native WASB stays "not production-verified"). One code-comment fix remains (the `server.js` Node-version, flagged by the PORTABILITY fixer). Let me check for any other stale "Node 22" and apply it.

679. user (2026-06-22T09:54:35.873Z)

C:\Users\avidu\Projects\khelsutra\site\server.js:1:// Always-on VM host for the KhelSutra site ? dependency-free (Node 22+ only).

680. assistant (2026-06-22T09:54:39.516Z)

Only `site/server.js:1`. Applying the consistency fix (the repo authoritatively requires Node 24 ? CI `node-version: 24`, `provision-vm.sh` `NODE_MAJOR>=24`).

681. user (2026-06-22T09:54:40.647Z)

The file C:\Users\avidu\Projects\khelsutra\site\server.js has been updated successfully. (file state is current in your context ? no need to Read it back)

682. assistant (2026-06-22T09:54:47.112Z)

```
The change is a code comment + markdown docs ? not browser-observable, so preview verification
doesn't apply here. The right gate is the test suite (which also boots `server.js`, exercising
the file I just edited). Running it, then committing.
```

683. user (2026-06-22T09:54:50.364Z)

```
? tests 86
? suites 0
? pass 86
? fail 0
? cancelled 0
? skipped 0
? todo 0
? duration_ms 1759.5717
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

684. assistant (2026-06-22T09:54:57.774Z)

```
86/86 green, `server.js` boots fine. Re-verifying HEAD, staging explicit paths, and committing.
```

685. user (2026-06-22T09:55:02.095Z)

```
=== HEAD/branch re-verify ===
docs/freshness-sweep
[empty above = no commits yet, clean base]
=== stage explicit paths ===
=== staged (expect 14 files, nothing else) ===
M  NEXT_STEPS.md
M  README.md
M  deploy/DEPLOY_ALPHA.md
M  docs/ANNOTATE_AND_TRAIN.md
M  docs/DRY_RUN_CUJ.md
M  docs/LOCAL_APPLIANCE_ARCHITECTURE.md
M  docs/MULTILINGUAL_EXPANSION.md
M  docs/PER_RALLY_SELECTION.md
M  docs/PORTABILITY.md
M  docs/README.md
M  docs/REQUEST_ACCESS_FORM.md
M  site/README.md
M  site/server.js
M  web/README.md
```

686. assistant (2026-06-22T09:55:06.517Z)

```
14 files staged, nothing stray. Committing, pushing, and opening the PR.
```

687. user (2026-06-22T09:55:23.553Z)

```
[docs/freshness-sweep 4ae9145] docs: freshness sweep ? reconcile 13 docs with the live repo (+
engine origin/master)
 14 files changed, 119 insertions(+), 84 deletions(-)
=== only my commit vs origin/master ===
4ae9145 docs: freshness sweep ? reconcile 13 docs with the live repo (+ engine origin/master)
=== push ===
remote:
branch 'docs/freshness-sweep' set up to track 'origin/docs/freshness-sweep'.
To https://github.com/avidullu/khelsutra.git
 * [new branch]      docs/freshness-sweep -> docs/freshness-sweep
```

688. user (2026-06-22T09:55:45.769Z)

```
https://github.com/avidullu/khelsutra/pull/95
```

```
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

689. assistant (2026-06-22T09:55:49.208Z)

```
PR **[#95](https://github.com/avidullu/khelsutra/pull/95)** opened. Watching CI to green, then
merging.
```

690. user (2026-06-22T09:56:53.357Z)

```
web     pass    40s
https://github.com/avidullu/khelsutra/actions/runs/27944489353/job/82685667762
layout  pending 0
https://github.com/avidullu/khelsutra/actions/runs/27944489353/job/82685667806
Workers Builds: hidden-king-khelsutra   pass    0
https://dash.cloudflare.com/213c12cf60de9041dd8a3521aa206472/workers/services/view/hidden-king-
khelsutra/production/builds/d767d758-45f6-4c62-ad2b-37084f777c55
layout  pass    48s
https://github.com/avidullu/khelsutra/actions/runs/27944489353/job/82685667806
site    pass    17s
https://github.com/avidullu/khelsutra/actions/runs/27944489353/job/82685667773
web     pass    40s
https://github.com/avidullu/khelsutra/actions/runs/27944489353/job/82685667762
Workers Builds: hidden-king-khelsutra   pass    0
https://dash.cloudflare.com/213c12cf60de9041dd8a3521aa206472/workers/services/view/hidden-king-
khelsutra/production/builds/d767d758-45f6-4c62-ad2b-37084f777c55
layout  pass    48s
https://github.com/avidullu/khelsutra/actions/runs/27944489353/job/82685667806
site    pass    17s
https://github.com/avidullu/khelsutra/actions/runs/27944489353/job/82685667773
web     pass    40s
https://github.com/avidullu/khelsutra/actions/runs/27944489353/job/82685667762
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

691. assistant (2026-06-22T09:56:55.814Z)

```
All four green (`site`, `web`, `layout`, Cloudflare preview). Merging and confirming clean
slate.
```

692. user (2026-06-22T09:57:06.822Z)

```
 site/server.js                       |  2 +-
 web/README.md                        | 14 ++++++++++--
 14 files changed, 119 insertions(+), 84 deletions(-)
HEAD: 7dc7de5 docs: freshness sweep ? reconcile 13 docs with the live repo (+ engine
origin/master) (#95)
open PRs: 0
branches:
  * master
Shell cwd was reset to C:\Users\avidu\Projects\badminton-highlight-indexer
```

693. assistant (2026-06-22T09:57:19.613Z)

```
Merged ? **PR [#95](https://github.com/avidullu/khelsutra/pull/95)** is in, `master` at
`7dc7de5`, **0 open PRs, only `master` locally**. Clean slate. Let me add a brief checkpoint to
the handoff (session-wrap convention), then summarize.
```

694. user (2026-06-22T09:57:21.378Z)

```
1       ---
2       name: current-state
3       description: "Live micro-checkpoint state for Claude ? updated every checkpoint (not
only on request). What just landed, what's next, open decisions."
4       metadata:
5         node_type: memory
```

```
6         type: project
7         originSessionId: b13673cc-f087-43b1-9527-35dc0b67b301
8       ---
9
10      # Current state ? updated continuously (see [[working-agreement-checkpoints]])
11
12      > **Scope of this file = LIVE only:** the current handoff + *today's* checkpoints + open
PRs/decisions. Keep it read-first-short (?1 page). Older checkpoints are archived in [[current-
state-archive]] ? when today's checkpoints age out (next clean-slate day), move them there,
don't let them pile up here.
13
14      **Checkpoint:** 2026-06-22 (GOLDEN-FIXTURES **P0 ? #314 MERGED**, master `5c2eec8`; a
SEPARATE thread from packaging/UX). Owner: "all owned golden videos are owned+minor-free; pick
(a); make P0 now." On scouting, P0 was bigger/riskier than the audit's "4 files": the video
names are FUNCTIONAL (eval_baselines keys ? training floor + the gitignored
`output/human_lovo_manifest.json` + data files; `MULTICOURT_CLIPS`; hot in
QUALITY_ITERATIONS/the quality track) ? surfaced it ? owner picked **SAFE SCOPE**. Shipped:
scrubbed personal `C:\Users\avidu` paths from 3 code docstrings + **`tools/leak_scan.py`**
(structural PII guard: 32-hex MD5 + personal paths over
backend/training/scripts/tools/eval_baselines; tests/ excluded; optional gitignored
`.leakscan_private` for local name scanning) + CI `leak-scan` job + 10 tests. Clean scan; suite
**1365?/7s**, cov **85.27%**, all CI green. **DEFERRED (safe-scope):** the functional video-name
rename + git-history rewrite (need manifest/data/training-floor coordination + quality-track
pause; do before open-sourcing). ?? **NEXT = P3 real Tier-0 fixtures** ? owner APPROVED (all
owned cleared + **(a)** commit de-attributed per-frame CSV); needs the golden **trajectory CSVs
+ label CSVs** (`Annotation Setup/Collect/Trajectories/` + `Golden Labelled/` / OneDrive
`golden-data/`) ? I locate them or owner points me, then `--freeze-real` per video. Golden-
fixtures track = #186/#189/#191/#192/#193/#194 + #314, all merged+audited. [[golden-set-
vendoring]]
15
16      **Checkpoint:** 2026-06-22 (? **GCS = corpus source-of-truth ? `docs/DATA_IN_GCS.md`
shipped; F: label backup done; 3 owner decisions open**). Owner goal: run ALL train+infer from
the cloud with ALL data in `gs://khelsutra-rally-corpus` ? remove the local box as a blocker;
document where data lives so other sessions pick up from the bucket. [[gcs-data-source-of-
truth]]
17      - ? **Authoritative map `docs/DATA_IN_GCS.md`** (indexed in docs/README.md; branch
`work/golden-corpus-hygiene`, commit `8884c9f`): bucket layout · current contents + **drift** ·
how `backend.worker infer|train` consume `gs://` · the gap · phased migration plan + DoD. The
**cloud engine already exists** (archived DECOUPLED_COMPUTE: `worker infer --video-uri` /
`worker train --corpus-uri`, URI backends in `backend/storage/ref.py`, `deploy/gcp/`) ? the doc
records WHERE bytes are, it does NOT redesign.
18      - ? **Gap measured:** corpus *metadata* is **~25 MB & entirely local** ? 20 trajectory
CSVs (17.5 MB at `C:\?\Collect\Trajectories\`), `output/human_lovo_manifest.json` (C:-absolute
paths), most labels. Bucket has only 3 videos, **no `trajectories/`**, inconsistent `_proxy`
label naming. Only the VIDEOS are large ? migrate **Drive?bucket from a cloud box**, not via the
local box.
19      - ? **F: backup done:** the 7 single-copy new-clip labels hash-verified into
`F:\[Khelsutra] GoPro Hero Black 12\KhelSutra\Training\Golden Labelled\`. ? F: root was RENAMED
`GoPro Hero Black 12`?`[Khelsutra] GoPro Hero Black 12` between turns ? live proof of the path-
fragility GCS fixes.
20      - ? **OPEN owner decisions (DATA_IN_GCS §7) before Phase-1 upload:** (1) `labels/`
naming (recycled GoPro names ? key by `video_id`/MD5?); (2) cache `trajectories/` in the bucket
for GPU-free eval?; (3) a bucket-relative eval manifest? On answers I upload
trajectories+labels+manifest (~25 MB, seconds).
21      - ? Branch `work/golden-corpus-hygiene` off origin/master (`689c738`), 1 commit, **NOT
pushed/PR'd yet** (awaiting owner). Serving WIP `6fdd1d4` + `scratch/PACKAGING_PLAN.serving-
wip.md` preserved untouched. Prior #236 served-Gate-A re-grounding MERGED 2026-06-19; original-6
labels turned out **recoverable from F:** (not lost) ? optional follow-up: re-measure litmus on
the full 15-clip corpus. [[compute-stack-gcp-only]] [[gotcha-shared-checkout-branch-collisions]]
22
23      **Checkpoint:** 2026-06-22 (**PACKAGING P2b + P2c ? BOTH MERGED; v1 LIGHT engine-
complete + wizard'd**). Same session as the P3 checkpoint below; owner: "complete P2b-install" ?
"start P2c wizard SPA now" ? "merge #310 now". **? P2b-install ? engine
[#310](https://github.com/Khelsutra/badminton-highlight-indexer/pull/310) MERGED** (master
`f8b2c3b`): `write_light_default_config` seeds a truly-local + **`motion`** (gate **O8** =
```

resolved) torch-free default `config.json` on the first `indexer --app` (minimal, profile-driven, NEVER overwrites; placed before `load_app_config`, scoped to `--app`); `scripts/install.py` (stdlib, cross-platform ? `uv tool install` else pip + a per-OS launcher shortcut that **pins `RALLY_APP_HOME`** ? load-bearing since `load_dotenv` runs at import + writes an uninstall manifest) + `scripts/uninstall.py` (manifest-driven; KEEPS user data unless `--purge`) + README. **12-agent adversarial review (`wf_2be29ecf`) folded:** MAJOR PATH-robustness (`uv tool update-shell` + resolve+embed the abs `indexer` path + warn) + uninstall **containment guard** (`_within(path, app_home)` refuses rmtree outside app-home ? a corrupted/foreign manifest can't nuke an arbitrary dir) + `--dry-run` fidelity + tests for the destructive path. Suite **1352?**, ruff+mypy clean, cov **85.17%**. ? **CI was billing-blocked for hours** ? the engine repo bills under the **`Khelsutra` ORG** (separate from the personal account); owner fixed the ORG billing ? CI green ? merged. **? P2c first-run wizard ? khelsutra [#88](https://github.com/avidullu/khelsutra/pull/88) MERGED** (master `7ae9cee`): `web/src/components/SetupWizard.tsx` ? 3 steps (AI-key *guided-setup* since no key-write endpoint yet / compute from `GET /api/capabilities` / videos) rendered before `IngestPanel`, shows once per browser (localStorage), 6 langs (en/kn/hi/es/da/id, brand stays Latin "KhelSutra" per #77), `getCapabilities()`+`EngineCapabilities`. **223 web tests** (+ the i18n parity/brand/no-bare guards) + **browser-verified e2e against a live local engine** (all 3 steps render real caps: ffmpeg+GPU detected; offline degrades gracefully). **13-agent review (`wf_191a9b86`) folded:** MAJOR a11y (the "Get a free Gemini key" link was below the 44px tap floor ? `min-h-11`) + loading-window finish-button (gated to `offline` only) + `role=dialog`?`region` + kn brand spacing + the no-bare-strings guard now covers SetupWizard. ? a sibling reorg-merged **#87** (es/da/id catalogs) mid-build ? rebased (auto-merged clean) + re-synced `site/public/locales` via `npm run copy:locales` (a `locales-drift.test` enforces the mirror). Both built in **isolated worktrees off `origin/master`** (since removed); sibling worktrees untouched. **Also this session (owner asks): deployed the real Gemini key (from `G:\My Drive\gemini.api_key.txt`) to the droplet `/etc/khelsutra/engine.env` + validated live (55 models, $0) + ran the FULL suite on the droplet (1313 passed); injecting the key auto-started `khelsutra-engine` (enabled unit) on loopback ? owner: "keep it running" (NOT publicly exposed ? `api./app.khelsutra.guru` have no public DNS route yet).** ? **LOOP CLOSED ? engine [#312](https://github.com/Khelsutra/badminton-highlight-indexer/pull/312) MERGED** (master `293e5c0`): re-bundled `backend/ui/` from khelsutra@`7ae9cee` via `scripts/bundle_ui.sh` so **`indexer --app` now SERVES the wizard** (`/api/version` spa_commit=7ae9cee; `/` serves `index-GSXqL6iY.js`; live-boot proven). Pure vendored-asset refresh; 53 UI/skew contract tests + full suite 1355? + cov 85.25%. **? v1 LIGHT is now feature-complete ? only OWNER steps remain: a clean-Windows/macOS-box acceptance run + ffmpeg-acquisition UX** (winget/brew/apt is still a manual dep). North-star (native-WASB-in-default-download P3b/P4a + training P4b + annotation P4c) stays future + counsel(O3)/owner-GPU-gated. ?? **THIS SESSION TALLY: 3 PRs merged (#310 P2b-install, khelsutra #88 P2c wizard, #312 bundle) + P3 GPU validation ($0) + droplet key/suite. 0 open PRs, 0 VMs, all worktrees/branches cleaned.** North-star (native-WASB-in-default-download P3b/P4a + training P4b + annotation P4c) stays future + counsel(O3)/owner-GPU-gated. [[packaging-app-plan]] [[next-session-packaging-app]] [[gotcha-shared-checkout-branch-collisions]] [[working-agreement-merging]]

24

25 **Checkpoint:** 2026-06-22 (**PACKAGING P3 ? GPU validation of the LIGHT-tier WHEEL on a real L4, end-to-end, $0**). Owner greenlit the spend (AskUserQuestion). Provisioned an **L4** `rally-gpu` (`g2-standard-4`, **asia-south1-a** Mumbai) ? shipped the `origin/master` archive ? **built the wheel ON the box + installed it NON-editable** (`badminton_highlight_indexer-0.1.0-py3-none-any.whl`, base LIGHT, no extras) ? **`import backend.main` works with `torch installed = False`** (the P2e torch-free promise PROVEN on a GPU box; every backend `import torch` is function-local) ? built the **SEPARATE** WASB conda env (py3.8/torch1.11) reached via `$WASB_PYTHON` (the native path is a pure `subprocess` shell-out ? `native_wasb_runner.py`). Then booted the **packaged launcher `indexer --app` from a NON-repo cwd** (app-home isolated state) ? `/api/health` ok · `/api/version` reports the bundled SPA (khelsutra@46e0244, content_hash ac2a3d6) · `/api/capabilities` = `ffmpeg:true,ffprobe:true / nvidia_smi_present:true / wasb_weights_present:true / gemini_key_present:false` · **bundled SPA served at `/`** ? drove the REAL API `POST /api/process(wasb_hybrid) ? poll /api/jobs ? /api/query ? /api/compile ? /api/reels` with the tuned 22-rally recipe (`detector_impl=native` + `skip_ai_handoff=true` + `min_shot_count=0`) ? **`RALLIES_DETECTED=22`** (reproduced the golden result via the PACKAGED stack, not `scripts/`) ? **`p3_highlights.mp4` 73.7 MB** watermarked reel ("Dynamic highlight reel saved", all 22 segments). **Proof preserved to `gs://khelsutra-rally-corpus/highlights/p3_{packaging_highlights.mp4,run.log,app.log}`** (survives teardown). **VM DELETED + audited `0 instances / 0 disks ? $0`** (~14 min ? **$0.17**, far under the $100/?10k cap). Ops-Agent auto-installed + verified (NVML receiver healthy, GPU graphs live; boot-time duplicate-write export errors are transient, settled). ? The ONLY hiccup was cosmetic:

my driver's `urlopen(timeout=60)` tripped on the synchronous 22-clip `/api/compile` (~65s) ? **client-side only; the server completed the stitch** (sync endpoint runs to completion regardless of client disconnect). The packaging insight WORTH KEEPING: **a plain torch-free LIGHT wheel + a separately-provisioned WASB conda env = full GPU-quality reel** ? the dual-env architecture (PACKAGING_PLAN §4.4) is real on hardware. ? **P3 DONE** ? next packaging items = **P2b-install** then **P2c wizard** (per [[next-session-packaging-app]] / [[packaging-app-plan]] §7). Screencast = owner records from the preserved reel/logs (or a local LIGHT-tier `indexer --app` browser capture, $0). Built in a detached worktree off `origin/master` (since removed); the serving checkout was untouched. [[gcp-vm-always-delete]] [[gcp-ops-agent-enable]] [[packaging-app-plan]]
26
27 **Checkpoint:** 2026-06-22 (**ALPHA-DEPLOY runbook DE-RISKED ? verified vs engine `origin/master` + repro'd; 2 khelsutra PRs**). Owner: "make T2 deploy real" + "make the PRs, run tests on droplet/GPU within the bill cap." **T2 itself is owner-gated INFRA** (CF Pages/Access/`cloudflared` tunnel + GEMINI_API_KEY rotation) ? nothing for Claude to *execute*; the high-leverage move = **adversarially verify `deploy/DEPLOY_ALPHA.md` against the REAL deploy code** (engine `origin/master`, NOT this stale serving checkout) via an 11-agent verification workflow + a deploy-readiness skeptic, then **repro the failures hermetically** ($0 ? droplet/GPU NOT needed, so the bill stayed ~$0). **Found 2 deploy-BREAKERS the original kit shipped:** (1) **`RALLY_ALLOWED_HOSTS` missing** ? `edge_auth` ON ? engine binds `0.0.0.0` + Host-guard stays loopback-only default ? **400s EVERY tunneled request** (boots "healthy" then dead; only a WARN). **Independently CORRECTED the skeptic:** `RALLY_BIND_HOST=127.0.0.1` does NOT fix it (the Host-allowlist is decoupled from the bind ? read `backend/main.py:88-128`); `RALLY_ALLOWED_HOSTS="api.khelsutra.guru"` is the load-bearing fix. **Repro (real Starlette `TrustedHostMiddleware`, exact engine allowlist): `/api/health` Host=`api.khelsutra.guru` ? 400 unset / 200 set.** (2) **gs:// ingest REJECTED** by the required `allowed_video_root` jail (repro'd 400 for `gs://`/`s3://`/`gdrive://`/`..`; pinned by `test_process_hosted_mode_rejects_nonlocal_uri`) ? §6/§7 over-promised it; alpha path = **browser upload**. **2 PRs MERGED (CI green, squash):** **[khelsutra #78](https://github.com/avidullu/khelsutra/pull/78)** (runbook fixes: +2 env vars, gs:// drop, 401-not-500, config.json `--setup` location, ~95MB chunk ? the runbook is now correct; the owner just deploys from it) + **[#79](https://github.com/avidullu/khelsutra/pull/79)** (stale §7 tracker prose). **Master moved since last handoff: A3(#74)/A6(#75)/A7(#76) MERGED + #77 kn/hi console native-review pass** ? A-workstream effectively DONE bar the es/da/id+watermark native review. Verified in an **isolated `origin/master` worktree (since removed)**; serving checkout (`serving/exposure-safety-preflight`) untouched. **? Engine-side thread to route to the SERVING session:** consider escalating the Host-allowlist WARN?ERROR (the one exposure misconfig that doesn't fail-fast yet 100%-breaks); #78 unblocks the deploy regardless. **Follow-up ? owner started the CF setup:** confirmed **NO $25/Pro plan needed** (Pages + Tunnel + Access all free, Access ?50 users; enabling Zero Trust may ask for a card ? pick the **Free** plan). **Team domain LOCKED = `https://khelsutra.cloudflareaccess.com`** (`cf_team_domain`). Owner does the key rotation + firewall (single user). Shipped the **tailored droplet kit ? [khelsutra #81](https://github.com/avidullu/khelsutra/pull/81) MERGED** (`deploy/droplet/`: `config.json`/`engine.env`/`khelsutra-engine.service`/`README` pre-filled for `app.`/`api.khelsutra.guru` + `avi.dullu@gmail.com` + `/srv/khelsutra` jail; declarative templates, JSON-validated, NOT yet run on the box). **? DEPLOYED + LIVE-VERIFIED ON THE DROPLET (owner: "ssh in with the previous-session key + take control"):** SSH works via **`~/.ssh/khelsutra_droplet`** ? `root@claude-do-rally-test` (Ubuntu 24.04, 2-core; ALSO runs `khelsutra-site.service` on :8787 ? left untouched; :8000 free; no git creds on box). Pushed engine `origin/master` (git-archive?tar over SSH, no box creds), `/opt/khelsutra/venv` + `pip install -e .[auth]` (**no torch in core deps**; cv2 needs `libgl1`+`libglib2.0-0t64`). Wrote `/srv/khelsutra/config.json` (AUD placeholder), `/etc/khelsutra/engine.env` (GEMINI placeholder + `RALLY_ALLOWED_HOSTS=api.khelsutra.guru,localhost,127.0.0.1` + `RALLY_BIND_HOST=127.0.0.1` + CORS), `khelsutra-engine.service` (enabled, NOT started ? correctly **fail-fasts until AUD+key**). cloudflared 2026.6.1 installed. **LIVE-PROVEN on the box (4/4):** boot-posture fail-fast ? exit 1 `refusing to start`; Host-guard `api.khelsutra.guru` ? **400 unset / 200 with RALLY_ALLOWED_HOSTS** (the #78 fix observed on the deploy target); real exposed config boots **`127.0.0.1:8000`** (RALLY_BIND_HOST overrides the edge_auth 0.0.0.0). **REMAINING = owner CF-account + key ONLY:** (1) rotated GEMINI key ? env; (2) create CF Access app ? paste AUD into config.json ? `systemctl start khelsutra-engine`; (3) CF Pages for `app.`; (4) `cloudflared tunnel login` (browser) + create/route. **OR** give me a scoped **CF API token + account-id + the rotated key** and I'll do Access/Pages/tunnel via API + run the full upload?reel CUJ. ? `StandaloneUser` Windows path = just different config strings later (nothing hardcoded).
28 **?? DEPLOY EXECUTED (2026-06-22, owner gave SSH-via-`~/.ssh/khelsutra_droplet` + a CF API token at `G:\My Drive\cloudflare.token.txt` for acct `213c12cf60de9041dd8a3521aa206472`; "do

exhaustive testing, make it worth the money"). RESOURCE INVENTORY:** Engine on the droplet `/opt/khelsutra/{badminton-highlight-indexer (origin/master),venv}`, config `/srv/khelsutra/config.json` (edge_auth ON, **AUD=`c980da67ed888fff935fcd913ae1eb4da63f9d65805d6e1c5ca76f92f517082b`**, **`indexing.default_segmenter=gemini`** ? Finding A fix), env `/etc/khelsutra/engine.env` (RALLY_* set, **GEMINI key = PLACEHOLDER**), `systemd khelsutra-engine` ACTIVE on 127.0.0.1:8000. **CF Access app** id `2bb7d8d7-e8df-4810-8fc4-40680e3aeb1f` (domains app.+api.khelsutra.guru, CORS+OPTIONS, policy `ec776118?` allow avi.dullu@gmail.com), team `khelsutra.cloudflareaccess.com`. **Tunnel** `khelsutra-studio` id `55142398-732c-4803-8a9c-48766f1f5cac`, `cloudflared` ACTIVE on box, ingress api.?127.0.0.1:8000. **Pages** project `khelsutra` LIVE at https://khelsutra.pages.dev (SPA built w/ VITE_API_BASE=https://api.khelsutra.guru/api), custom domain app.khelsutra.guru = `initializing` (pending DNS). Zone khelsutra.guru=`65ef7770a6f25697843b53a093504184`. **EXHAUSTIVE TESTS (box):** full suite **1313 passed/12 skipped**, **ruff clean**, **mypy clean (133 files)**; live: boot-fail-fast, Host-guard 400?200, jail rejects gs://+/etc/passwd+badext, **T4 upload?jail OK**, jobs run IN-PROCESS (no worker svc needed), gemini segmenter torch-free + runs (default yolo_hybrid needs torch?Finding A). **2 BLOCKERS LEFT:** (1) **token lacks `Zone.DNS:Edit`** ? can't create the 2 DNS records: `api` CNAME?`55142398-732c-4803-8a9c-48766f1f5cac.cfargotunnel.com` (proxied) + `app` CNAME?`khelsutra.pages.dev` (proxied). Owner: add DNS:Edit to the token (I finish) OR add the 2 records manually. (2) **Gemini key** (owner rotating in a few hrs ? `/etc/khelsutra/engine.env`; without it processing returns `done` with 0 rallies + the missing-key error only in the LOG, not the job status = **Finding B**, an engine UX gap to route to serving). ? old key was in `~/.keys/`+`.bash_history` ? now GONE (rotate revokes nothing reusable).

29 **?? ALPHA IS LIVE (owner added the 2 DNS records manually after the CF token lacked Zone.DNS:Edit).** **FULL CHAIN VERIFIED:** `api.`/`app.khelsutra.guru` resolve (CF proxy); unauthenticated ? **302 to `khelsutra.cloudflareaccess.com/.../login`** with `kid` = the engine's AUD (gate correct); an Access-authenticated request ? tunnel ? engine **`/api/health` = 200** (DNS?CF?Access?tunnel?engine proven); engine JWT gate ACTIVE (401 on missing/garbage). SPA live on Pages (app.khelsutra.guru). **Honest caveat:** the full `/api/process` pass for a real EMAIL-OTP user is covered by engine unit tests + the code (`auth.py` user_id=email-else-sub) but NOT live-exercised (OTP login is a browser/human step); my Access **service-token** test authenticated through Access (health 200) but the engine 401'd it at /api/process ? CORRECTLY, because service tokens carry no email/sub claim (auth.py:96-98), which email users DO carry ? not a real-user blocker. **2 STEPS LEFT, both owner:** (1) rotate GEMINI key ? `/etc/khelsutra/engine.env` + `systemctl restart khelsutra-engine`; (2) browser OTP login at app.khelsutra.guru ? upload?reel (the human CUJ). **Finding A FIXED in the deploy + committed kit (indexing.default_segmenter=gemini).** **Finding B** (no-key ? job `done` w/ 0 rallies + error only in log) = engine UX gap to route to serving. **? UPDATE:** owner dropped a **valid working Gemini key** into `/etc/khelsutra/engine.env` (len 53, `AQ.A?` prefix ? a NEWER Google key format, NOT `AIza`; **Gemini auth OK, 55 models**). **Full real-Gemini CUJ run on an isolated :8001 local-mode instance:** `/api/process` sample.mp4 (locale=kn) ? job **done, pipeline `success`** in ~30s ? Gemini analyzed it ? **0 rallies (sample.mp4 is a 12s smoke clip, too short for a rally ? pipeline+key PROVEN; reel/watermark path proven separately via the LW1 real-watermark videos)**. Production :8000 untouched. **So the alpha is FUNCTIONALLY COMPLETE** ? only the **human browser one-time-PIN login** (app.khelsutra.guru ? upload?reel) and a **real-rally video** (not the 12s smoke clip) remain to see an actual reel. Kit aligned ? [khelsutra #85](https://github.com/avidullu/khelsutra/pull/85) MERGED (gemini default + live-verified README). Owner's standing key-rotation TODO unchanged (current key works). ? owner may want to **revoke the broad CF token** (`G:\My Drive\cloudflare.token.txt`) now setup is done.

30 **?? DOC-SCAN ? PRIORITIZED NEXT (6-agent workflow, git-grounded vs origin/master, 2026-06-22):** **Packaging answer:** download+run works TODAY for a dev (`pip install` wheel ? `indexer --app` ? bundled torch-free SPA, LIGHT/Gemini tier ? all on origin/master: entrypoint pyproject:86, `--app` main.py:867, `backend/ui/` served main.py:194-210, P2e #308 torch-free). **NOT one-click for a non-tech user** ? missing **P2b-install** (install script/uv/default config/uninstaller) + **P2c first-run wizard** (Gemini key/compute from `/api/capabilities` ? SPA doesn't consume it yet) + ffmpeg is a manual dep. **Ranked next:** (1, Claude) **Finding B** fail-job-on-missing-key (P0, protects the live alpha; gemini.py:99?[]?job marked done; NOT covered by #294); (2-3, Claude) **P2b-install + P2c wizard** (the "download & run" goal); (4, owner) **native review** kn/hi-site+es/da/id+watermark; (5, owner-$) **M7** persistent Cloud-Run service+DEPLOY_REPORT (+settle the P3 question same GPU session); (6, owner) **prominent-court gate #243** measure/promote (cheapest OPEN quality win ? multicourt 1.58× over-fire is the #1 gap); (7) compute picker (after M7); (8) droplet ops hygiene (retention #301 schedule, revoke cloudflared cert.pem, monitoring); (9, owner) grow golden corpus + ratify tolF1 gate. **Stale NEXT_STEPS to ignore:** Gate-A-wiring (done, #146 reverted?opt-in #154), SPA-ingest-screen

(shipped #56/#66), async-job-slice (done?CDS is live); **M9-consent/M11 need a v8 migration (NOT v7 ? locale took v7 #306)**; re-verify every `[gap]` at file:line ([[lesson-verify-engine-surface-before-scoping]]). Full scan: workflow `wf_6abc9156`.
31 **?? AUTONOMOUS SESSION (owner "complete highest-leverage + clean slate"):** **? Finding B DONE ? engine [#311](https://github.com/Khelsutra/badminton-highlight-indexer/pull/311) MERGED + REDEPLOYED to the live droplet + VERIFIED LIVE.** `GeminiPlaySegmenter._resolve_provider` now returns a **`Failure`** (with a user-facing reason) on the 3 unrunnable cases (missing key / unknown sport / unknown provider) instead of `None`?`[]`; `process_video` propagates it. Reuses the existing contract: `_execute_processing` maps Failure?`result.status="failed"`+message (main.py:521), the SPA already shows a done+`result.status=="failed"` job as a failure (`useIngestJob.ts:88-95`), CLI worker already returns rc 3. Genuine "ran, found 0" unchanged. +3 tests; full suite 1335? / ruff / mypy clean; all 9 CI lanes green. **Live proof on :8001 (no key): job.status=done, result.status=failed, message="GEMINI_API_KEY is not set?".** Built in an isolated origin/master worktree (now removed); serving checkout untouched. **cert.pem flagged item = NON-ISSUE** (token-path tunnel ? no account-wide cert on the box). **P2b NOT auto-done ? deliberately:** it needs an OWNER distribution-channel decision (the repo is PRIVATE ? how does a non-tech user GET the wheel: publish to PyPI/make public vs ship a downloadable `.whl`) AND its DoD is a clean-Windows-VM test I can't run autonomously; the install-script + truly-local default config + uninstall manifest are ready to build once the channel is picked. **? Clean slate: both repos on master/serving-branch, no worktrees, no running test VMs, live alpha healthy + now hardened with Finding B.** [[next-session-ux-localization]] [[gotcha-shared-checkout-branch-collisions]] [[lesson-verify-engine-surface-before-scoping]] [[decision-rigor-norm]]
32
33 **Checkpoint:** 2026-06-21 (**LOCALE ? WATERMARK + PMF ANALYTICS ? owner ask; 3 PRs merged + verified via REAL watermark video for all 6 locales**). Owner: "do we record the locale when we store the video? ? spec+build it, localize the watermarks in every supported language, make basic analytics possible." **Verified-first finding:** NOT recorded (SPA sent no locale; no DB column). Built cross-repo, **decoupled**: watermark via `CompileRequest.locale` (transient), analytics via a `jobs.locale` v7 column. **3 PRs merged (CI green):** engine [#304](https://github.com/Khelsutra/badminton-highlight-indexer/pull/304) **LW1** localized reel watermark (`backend/pipeline/stitcher/watermark_i18n.py::localize_watermark` ? text+font per locale; bundled **OFL Noto Sans Kannada/Devanagari** [D7 carve-out, fetched from google/fonts] for kn/hi, RobotoSlab for en/es/da/id; `/api/compile` localizes; unknown/`en` ? default unchanged + graceful font fallback; full suite 1302?) · khelsutra [#73](https://github.com/avidullu/khelsutra/pull/73) **LW2** SPA sends `i18n.resolvedLanguage` to `/api/compile` + `/api/process` (177 web tests) · engine [#306](https://github.com/Khelsutra/badminton-highlight-indexer/pull/306) **LW3** **v7 migration** `jobs.locale` (NULLABLE) + `Database.count_jobs_by_locale()` + `ProcessRequest.locale` (full suite 1315?). **? OWNER-DIRECTED VERIFICATION (before merging LW3):** generated a REAL reel through `VideoCompiler` for EACH of en/kn/hi/es/da/id and eyeballed the baked-in watermark ? ALL render correctly (kn/hi no tofu, Latin brand intact). **Engine migration head now v7** (serving consent migration ? v8; re-checked NO competing v7 at commit + rebase + merge). Engine PRs built in **isolated worktrees** off `origin/master` (serving checkout untouched). ? watermark + catalog strings are AI-drafted ? **native review pending**. ? Track handoff = [[next-session-ux-localization]]. [[gotcha-shared-checkout-branch-collisions]] [[working-agreement-merging]]
34
35 **Checkpoint:** 2026-06-21 (**T11 retention sweep ? engine PR MERGED + both follow-ups done**). The owner-spun T11 chip (`task_2480b9e2`) is now BUILT in the engine. **Extended `tools/cleanup_caches.py`** (not hand-rolled, per [[cache-hygiene-policy]]) with a pre-alpha **upload retention sweep**: a new `"uploads"` scan location + `classify` (`uploaded source` / `partial upload` for the orphaned `.partial-<id>` files) over `security.upload_dir` (default `output/uploads`); `plan_deletions` gains an **independent `uploads_retention_days` window** (privacy) orthogonal to the cache `--older-than` (disk); `scan_windows` now **skips the nested uploads dir** (was silently kept whole as "other dir" ? no retention); loud structured output (retention-window banner + per-category subtotals + per-item PURGED/FAILED + a `--json` plan for a cron/job wrapper). **Dry-run by default; reels/DB/CSVs stay durable.** **Owner gate (AskUserQuestion): uploads 7d default; reels kept durable.** ? **engine [#299](https://github.com/Khelsutra/badminton-highlight-indexer/pull/299) MERGED** (`cb7394b`, all 11 CI checks green: ruff/mypy/full-suite/golden cross-OS×3/wheel-smoke/license-scan/syntax+import); +6 unit tests; **coverage no-regression VERIFIED** (branch **85.30%** vs `origin/master` baseline **85.06%** measured in an isolated worktree; same **9748** measured stmts, branch missed **1433 ?** baseline **1456** ? the +0.24% is obsreport reporting tests in my local lane, NOT my diff; the diff touches **ONLY tools/+tests/** and the tool imports

zero backend/training). Branch cut from `origin/master` in the shared checkout, explicit-path staging; after merge the checkout was **restored to the serving session's branch** + feature branches deleted. **? BOTH FOLLOW-UPS DONE (owner-directed):** (1) automation (unattended scheduled-job wiring for hosted alpha) filed ? **[engine #301](https://github.com/Khelsutra/badminton-highlight-indexer/issues/301)** (enhancement; the tool + `--json` hook is ready, needs a cron/Cloud-Scheduler wrapper + observability + default-OFF guard + in-flight/active-job safety). (2) **khelsutra §9 T11 doc flipped ??? ? [khelsutra #69](https://github.com/avidullu/khelsutra/pull/69) MERGED** (`bea8a2a`; honest framing: sweep TOOL shipped #299, scheduling tracked #301; also updated the §6 retention risk row + §10 blocking-gate-#2 [window DECIDED] + §11 DoD). Done in a dedicated worktree off `origin/master`; the UX session's `khelsutra-b3-wt` worktree was untouched. [[cache-hygiene-policy]] [[ui-driven-cloud-serving-gap]] [[gotcha-shared-checkout-branch-collisions]] [[doc-status-register]]

Checkpoint: 2026-06-21 (**UX track ? 6 khelsutra PRs merged; workstream A done; SPA now 6 languages**). The UX/localization track (separate session from the serving/alpha-shareable one). Owner: "keep merging green CI PRs + improve coverage." **6 PRs merged ? all green CI, browser-verified, each adversarially reviewed via a workflow:** [#61](https://github.com/avidullu/khelsutra/pull/61) **A4** (?44px tap targets + tooltip-text?visible) · [#62](https://github.com/avidullu/khelsutra/pull/62) **A5** (ONE shared `lib/errorText.ts` + `errors.*` across build+load+ingest) · [#63](https://github.com/avidullu/khelsutra/pull/63) **A8** (reassurance + icon+text `StatusChip` + building card + ?/? + aria-live) · **[#66](https://github.com/avidullu/khelsutra/pull/66) T4 ? in-UI file UPLOAD (NORTH-STAR keystone)** (file picker ? chunked `/api/upload/*` via `api/client.ts`+`hooks/useUpload.ts` ? server path ? `/api/process` ? existing back-half) · **[#67](https://github.com/avidullu/khelsutra/pull/67) MULTILINGUAL Phase A** (es/da/id ? **6 languages**; es 3-form plural; a **placeholder-parity** guard that caught+fixed a latent kn/hi A1-drift `{videoId}`) · **[#68](https://github.com/avidullu/khelsutra/pull/68) A9** (localized tap-only filter controls; English chips now en-only; tap==type pinned). **Workstream A done (A1/A4/A5/A8/A9 ?); web tests 103?173 + 2 new invariant guards (placeholder-parity, tap==type).** **Owner decisions (AskUserQuestion):** T4 first; **B3 catalog-home = hoist to repo-root `/locales/` (LOCKED)**. **Chips spun + resolved:** MULTILINGUAL (`task_e15e9643`) ? did it myself (#67); **T11 (`task_2480b9e2`) ? picked up by the serving session ? engine [#299](https://github.com/Khelsutra/badminton-highlight-indexer/pull/299) OPEN**; **B3 (`task_b3060842`) STILL PENDING ? the top next, now fully unblocked** (all 6 catalogs ready to hoist; architectural ? SPA import repoint + new site build step + Vite out-of-root import config ? best as a focused session, ideally split PR-1 hoist / PR-2 shim). **Other remaining:** A3 (recording-edu kn/hi, depends on B3), A6/A7 (site), native review of AI-drafted kn/hi/es/da/id; **T2 = owner stand-up only** (deploy kit #65 shipped) + hosted-mode `allowed_video_root` ? `upload_dir` caveat. ? Caught+fixed a slip: started T4 edits on `master`, moved to a feature branch BEFORE any commit. ? Track handoff = [[next-session-ux-localization]]. [[working-agreement-merging]] [[gotcha-shared-checkout-branch-collisions]]

Checkpoint: 2026-06-21 (**ALPHA-SHAREABLE ? Claude's half DONE; 3 PRs merged**). Owner's #1 = "highlights from a page I can share, no SSH." **KEY CORRECTION (anti-spiral [[lesson-verify-engine-surface-before-scoping]]):** the icebreaker said do "B-P2 ingest screen + M7 Cloud Run *service*" ? but a recon **workflow** (`wf_5de56b64`, 5 readers across both repos) proved **B-P2 was ALREADY shipped** (khelsutra #56, live e2e `DRY_RUN_CUJ.md` §9) + CORS is already env-driven ? **alpha-shareable is a DEPLOY/AUTH task, not an SPA build**. **Owner decisions (2× AskUserQuestion):** topology **(i) Tunnel-to-a-box**, sub-topology **SPLIT** (SPA on CF Pages + engine on `api.*` cloudflared tunnel + CF Access; single-origin Caddy kept as the alt). **3 PRs MERGED:** ? **engine [#294](https://github.com/Khelsutra/badminton-highlight-indexer/pull/294)** (`0c0e064`) **exposure-safety boot posture** ? `startup.py` fails fast at boot (server-only `include_exposure=True`) when `edge_auth_enabled` is ON but half-configured (`EXPOSED_NO_VIDEO_ROOT`/`EXPOSED_EDGE_AUTH_INCOMPLETE` fatal, `EXPOSED_AUTH_EXTRA_MISSING` warn); **default-OFF**; aligned `auth.resolve_tenant` whitespace; adversarial 3-lens review (`wf_35b735a9`) ? no bypass; +16 tests; 1259 green cross-OS. ? **khelsutra [#64](https://github.com/avidullu/khelsutra/pull/64)** (`8bbc557`) **SPA `VITE_API_BASE`** ? `apiBase()` env-driven (default same-origin `/api`) + `vite-env.d.ts` + 3 tests; build-smoke proved the override bakes in. ? **khelsutra [#65](https://github.com/avidullu/khelsutra/pull/65)** (`4627830`) **deploy kit** ? `deploy/DEPLOY_ALPHA.md` (split runbook + CORS-preflight gotcha + posture test) + `cloudflared.config.example.yml` + tracker pointers. Each isolated/default-OFF/adversarial-reviewed/CI-green/merged-on-green; khelsutra built in a dedicated **worktree** (main tree had sibling A8 WIP, since merged #63; ? `rmdir` the `node_modules` junction BEFORE `git worktree

remove`). **? REMAINING for alpha = OWNER stand-up only** (cloudflared tunnel + CF Pages + CF Access + key rotation per `khelsutra/deploy/DEPLOY_ALPHA.md`). **? NEXT engine code (Claude-doable, batch-authorized): M12** (gdrive-api, clean, no DB migration) then **M11** (flywheel skeleton, privacy-gated on v7 consent). [[ui-driven-cloud-serving-gap]] [[lesson-verify-engine-surface-before-scoping]]

40

41 **?? NEXT SESSION FOCUS (2026-06-21, after the big localization+dry-run session) ? ? SUPERSEDED by the checkpoint above (B-P2 confirmed DONE):** SPA is fully **en+kn+hi (DONE)**; a real **UI?engine?reel?download CUJ ran (en+hi)**; a **critical engine bug fixed** (#281). **? NEXT BUILD (owner-picked): (1) Path B ? point at a bucket video IN THE UI.** ? **B-P1 VERIFIED ALREADY DONE ? NO ENGINE CODE (do NOT re-scope it as a feature):** `POST /api/process` has accepted `gs://`/`gcs://` URIs since **M2/#280** (submit resolves the URI + skips the local-exists-404 for non-local backends `main.py:283`; the worker fetches to scratch `acquire_local_input main.py:159` ? same hash?validate?segment?DB pipeline ? queryable API-DB `video_id`); **TESTED** (`tests/test_api_endpoints.py::test_process_nonlocal_uri_skips_exists_and_fetches` + hosted-mode-reject / unknown-scheme-400 / CLI variants). Confirmed by a **3-skeptic verification workflow (0 refuted, gaps:[])**. **Scope pin:** holds in DEFAULT single-user mode (hosted-mode `allowed_video_root` correctly **400s** a bucket URI ? by design). Stale "local-path-only / net-new" docs **CORRECTED ? khelsutra [#55]** (DRY_RUN_CUJ §1/2/4/6/8 + LOCAL_APPLIANCE + README). **? So Path B = ONLY B-P2 = the SPA ingest/progress screen** (enter a `gs://` URI ? `POST /api/process` ? poll `/api/jobs` ? load `result.video_id`; back-half already works) ? **100% khelsutra `web/` SPA + i18n (en/kn/hi), ZERO engine code.** Ready B-P2 sketch: `DRY_RUN_CUJ.md` §2 (new `web/src/{api/types,api/client,hooks/useIngestJob,components/IngestPanel}.tsx` + `App.tsx` videoId?state handoff + `ingest.*` keys). See [[lesson-verify-engine-surface-before-scoping]]. **(2) Droplet deploy (opt 3):** deploy the FIXED engine to `168.144.126.176` (un-provisioned, 2-core/3.8GB) ? re-run CUJ on Linux. **STILL OPEN (localization DoD):** B3 (`site/` marketing still English-only) + native-speaker review of the AI-drafted kn/hi. **Ramp-up:** `khelsutra/docs/DRY_RUN_CUJ.md` + `FRIENDLY_AND_MULTILINGUAL.md` §7 + the dry-run checkpoint below. **Clean slate (at session start):** both repos on master, 0 open PRs, no worktrees/VMs. khelsutra master `2421557` · engine master `2afd54f` (? origin/master `cee4baa`, #283 M4 squashed).

42

43 **Checkpoint:** 2026-06-21 (ENGINE SERVING M5+ ? OWNER BATCH-AUTHORIZED the independent path). Owner redirected: "first surface the owner-gated decisions so I can take them + unblock max independent PRs." Ran a 6-agent grounded gate-map **workflow** (`wf_4849400f`) over README §2/§7/§8 + code ? a decision menu splitting every M5?M12 gate into **(a) default-OFF code I can merge now** vs **(b) thin owner residual**. **Owner answers (AskUserQuestion):** **(1) BLANKET YES** ? write+merge ALL default-OFF serving code (M5,M6,M8,M9-auth,M9-consent-cols,M11-plumbing,M12), one isolated PR each; **+ SPEND OK up to $100/?10k** for due-diligence + **production-grade testing/verification** (so the real M6/M7 cloud runs are ALSO unblocked within the cap ? must respect it, DELETE resources after, confirm cmd+cost before each spend per [[gcp-vm-always-delete]]). **(2) MERGE each on green CI** (OAuth real-testing deferred to a later rigorous pass). **(3) ALL recommended design picks** "with generous budget in mind": **co-located stitch** (D4) · **libx264 ENCODE + NVDEC DECODE-only** (protect parity) · **hard-reject `<2GB` v1** (config knob) · confirm `COMPUTE_DECOUPLED_SERVING/` name · **M6 = Cloud Run JOBS + bucket-poll completion**. **? PLANNED PR ORDER:** A(doc: §8 resolutions + P3 reconcile + tracker housekeeping) ? B(M5 preset+startup-checks+L1?L4 parity test) ? C(M6 `backend/compute/` registry + Dockerfile + mocked-dispatch) ? D(M8 v6 migration + pricebook[$0 placeholder] + USD/INR aggregation + /cost API) ? E(M9-auth Cf-Access JWT verify + tenant scoping; M9-consent inert NULL cols EXTEND D's v6 migration) ? F(M11 capture?store?shadow-eval?goldenize plumbing, `flywheel.*`=OFF) ? G(M12 `GDriveApiStorage`+`gdrive-api://`, fake-Drive tests) + Ops-Agent codify into `deploy/gcp/provision.sh`. Coordinate: D/E/M9-consent share ONE v6 migration (author in D, extend in E). **Genuinely owner-only residual:** real GPU run (M5 runlog) [or I run on cloud GPU within cap], counsel ToS/DPA (M9-consent text), Google OAuth app (M12 real), real GCP prices+FX (M8 calibrate), CF team-domain/AUD+ingress-lock, repo rename, DO droplet/token retire. ? shared checkout ? sibling merged **#284** (`06ff4b7`, bucket-URI regression guard) mid/just-before this session; RE-FETCH origin/master before each branch. **? PROGRESS (this session):** ? **PR-A [#285] MERGED** (`fefa412`, doc: §8 D16/D17 resolutions + P3 reconcile + tracker housekeeping). ? **M5 [#286] MERGED** (`b5af66e`, all 7 CI green) ? `backend/config/startup.py` per-profile fail/warn-fast checks (cloud-serving+non-cloud-storage = the one fatal coherence error; GPU/creds = warnings) wired into server (`_enforce_startup_or_exit`?exit1) + worker (`cmd_infer`/`cmd_train`?rc2), **default-OFF + probe gated behind active profile (worker stays torch-free at profile='')**; **L4 profile-parity test** (`tests/test_profile_parity.py`: same stub trajectory?byte-identical windows+ranges across truly-local/cpu-mock/no-profile + worker-

vs-in-process); 4-lens adversarial review folded (`wf_34671ac1`: MAJOR torch-free fix + minors). Runlog template pre-staged; the real-GPU truly-local **runlog = owner L5**. M5 ? in §7 tracker. ? **M6 [#287] MERGED** (`8cdf65c`, all 7 CI green) ? `backend/compute/` ComputeBackend registry: `get_compute_backend(config)` returns a backend ONLY for `compute_target=cloud_run` (else `None` ? worker's in-process path, **default-OFF**); `CloudRunCompute` dispatches a Cloud Run **Job** (injectable `CloudRunJobClient`; real `GoogleCloudRunJobClient` lazy google-cloud-run) ? **bucket-polls** a done-marker via `execution_policy.poll_until`+storage (D16); container forced **inline** (`compute_target=local_gpu`+native, never re-dispatch); worker `--done-uri` writes a success marker. `deploy/cloudrun/{Dockerfile,README}` = L4 image (CUDA+ffmpeg+DejaVu fonts+`wasb` conda torch-1.11 env mirroring gpu_setup.sh; weights staged per WASB-C3) + M7 runbook. 4-lens review (`wf_87b22fa4`) folded a **BLOCKER** (cloud `fetch` needs a dst ? reads silently returned `{}` on gcs/s3; tests passed only b/c local) + poll-wait test gap + `PolicyTimeout`?clean rc 4. Suite 1167 green. ? **M8 [#288] MERGED** (`43eccc1`, 7 CI green) ? `backend/billing.py` (pure `usage×pricebook`, USD source-of-truth / INR display) + **v6 migration** (NULLABLE cost cols on `jobs` + versioned `pricebook` seeded `placeholder-v0` at $0; additive-NULL, idempotent) + `record_job_cost`/`get_job_cost`/`get_video_cost` + `GET /api/jobs|videos/{id}/cost` + `BillingConfig` (**default-OFF**) + gated `_maybe_record_job_cost`. 3-lens review (`wf_498b24af`) folded a **MAJOR silent-GPU-under-bill** (an unpriced `gpu_type` is now SURFACED ? WARNING + `unpriced_gpu_seconds` breakdown line, not invisible $0) + the **honest no-op** caveat (pipeline doesn't emit `result['usage']` yet ? cost 0 until a follow-up wires the metering; documented + tested) + margin/rounding reconciliation. Suite 1186 green. **?? 4 SERVING PRs THIS SESSION, ALL MERGED: #285(docs §8 D16/D17 + P3 reconcile) + #286(M5 startup-checks+L4 parity) + #287(M6 Cloud Run dispatch shim+image) + #288(M8 cost ledger).** Each default-OFF, isolated, adversarial-reviewed, full-suite+CI green, tracker+changelog in-PR. **? v6 is now SHIPPED (M8 took it) ? the NEXT migration is v7** (M9-consent cols + per-video P1 `workspace_uri` ? the earlier "extend v6" plan shifts to a fresh v7 `_migrate_to_v7`). **? REMAINING authorized batch (ready to build next, blueprints loaded):** **M9-auth** = add `PyJWT[crypto]`; a FastAPI dep extracts `Cf-Access-Jwt-Assertion` ? fetch+cache the team JWKS ? verify sig/exp/iss/aud ? `user_id`; per-tenant scoping on the v4 `user_id` cols; tighten the placeholder CORS (`main.py:40`); all behind `security.edge_auth_enabled=False`; offline RSA/fake-JWKS tests. **M9-consent** = inert NULL `consent_status/is_adult_attested/retention_class` in a **v7** migration (safe no-train default) ? the ToS/DPA TEXT is owner/counsel. **M11** = `flywheel.capture/shadow_eval/promote_stub` reusing `workspace.py::for_video` + `backend/eval/experiment.compare`(:485) + `promotion.py::promote_if_served_safe`, behind `flywheel.*`=OFF + a faked `consent_attested`. **M12** = `GDriveApiStorage(StorageBackend)` + a `gdrive-api://` scheme against an INJECTED fake Drive (mirror `gcs.py` DI), keep the mount-only `gdrive://` (truly-local) untouched. **Ops-Agent codify** into `deploy/gcp/provision.sh` (logging+monitoring APIs + SA metric/log-writer roles + startup-script ? [[gcp-ops-agent-enable]]). **? owner-only residual:** real GPU run ? M5 truly-local runlog (template pre-staged in `runlogs/`); GCP spend for the real M6/M7 run within the **$100/?10k cap** (build the image per `deploy/cloudrun/README`, run on L4, capture DEPLOY_REPORT, **DELETE after**); counsel ToS/DPA (M9-consent); real GCP prices+FX + the M8 usage-emission follow-up; Google OAuth app (M12 real); CF team-domain/AUD + ingress-lock; repo rename; DO droplet/token retire. **?? REAL M7 CLOUD GPU RUN ? ? DONE + VERIFIED (owner: "complete any and every cloud run with GPUs").** Provisioned an **L4** `rally-m7-gpu` (`g2-standard-4`, **asia-southeast1-a** Singapore ? both Mumbai zones STOCKOUT for L4; reads the Mumbai bucket cross-region) ? `bootstrap --gpu` + `gpu_setup.sh` (wasb conda env, **torch 1.11+cu113 cuda True on L4** via PTX-JIT) + staged the 5.82 MiB weights ? `worker -v infer --video-uri gcs://?/mahadevpura_smoke_180s.mp4 --out-uri gcs://? --segmenter wasb_hybrid` ? **22 rallies** (5395 trajectory points, 2930 visible; Phase 2 WASB 363.7s) ? pushed `gs://khelsutra-rally-corpus/outputs/05e0e577?/{intervals.csv,report.json}` (segment_count 22, status success). **VM DELETED, audit 0 instances/0 disks ? $0** (~18 min ? **$0.21**, well under the cap). **? SOLVED the long-flagged "cold-start gives 0 rallies" mystery ? THE MISSING PIECE = `--set indexing.skip_ai_handoff=true`:** `wasb_hybrid` does a Phase-3 **Gemini handoff** by default; with no `GEMINI_API_KEY` on the box it errors ? 0 segments. `skip_ai_handoff=true` = the **fully-local** path ? emits the 22 WASB candidate windows directly as rallies (no Gemini). So a no-Gemini cloud run MUST set `skip_ai_handoff=true` (this is exactly M5's truly-local preset, which sets it). ? minor: `-v` is a TOP-LEVEL flag ? `worker -v infer ?` (not `infer -v`). DEPLOY_REPORT runlog = TODO (after M9-auth). ? **M9-auth [#290] MERGED** (`aac7633`) ? `backend/api/auth.py` Cf-Access JWT verify (RS256: sig+exp+iss+aud; lazy `PyJWT[crypto]`; injectable JWKS ? fully-offline tests incl. alg=none + HS256-confusion rejection) + `resolve_tenant` ? `create_job(owner_id=?)`; `SecurityConfig.{edge_auth_enabled=False,cf_team_domain,cf_access_aud}` + CORS via `RALLY_CORS_ALLOW_ORIGINS` env. **Default-OFF.** Security review (`wf_9b4c71d0`) found NO bypass. ? **OPS-AGENT codify + DRY-RUN GUIDE [#291] MERGED** (`7835f34`, owner asked):

`deploy/gcp/create_gpu_vm.sh` (THE one-command way to spin a GPU VM ? zone-sweep L4 Mumbai?Singapore + **always bakes the Ops-Agent startup-script**) + `ops-agent-startup.sh` + provision.sh (logging/monitoring APIs + SA metric/log-writer roles ? the empty-graphs root cause) + **`docs/CLOUD_GPU_DRYRUN_GUIDE.md`** (step-by-step local-video?cloud-GPU?reel, validated from the M7 run + the local `VideoCompiler` stitch; the **`skip_ai_handoff=true`** no-Gemini path; `min_shot_count:0` so no rally is filtered). **?? 6 SERVING/INFRA PRs MERGED THIS SESSION: #285,#286(M5),#287(M6),#288(M8),#290(M9-auth),#291(ops-agent+guide) + the REAL M7 L4 run (22 rallies, $0).** **? NEXT (owner "keep going"): M11** flywheel plumbing (`flywheel.capture/shadow_eval/promote_stub` reusing `workspace.py::for_video` + `backend/eval/experiment.compare` + `promotion.py::promote_if_served_safe`, behind `flywheel.*`=OFF + a faked `consent_attested`) ? **M12** `GDriveApiStorage(StorageBackend)`+`gdrive-api://` (injected fake-Drive DI; keep mount-only `gdrive://`). ? M9-consent cols = a **v7** migration (M8 took v6). Ops-Agent codify = ? DONE (#291; **owner CONFIRMED Observability live in the cloud console 2026-06-21**). ? **#292 MERGED** (`2afbadf`) ? dry-run font fix (`bootstrap.sh` installs `fonts-dejavu-core`; a 2nd dry run via `create_gpu_vm.sh` PROVED Ops-Agent auto-installs + found the watermark needs the font) + the **?UI-trigger tracking** (NEXT_STEPS ?P0-0 + [[ui-driven-cloud-serving-gap]]). **?? SESSION WRAPPED ? CLEAN SLATE: 7 PRs merged (#285/286/287/288/290/291/292) + the real M7 L4 run (22 rallies); 0 open PRs; 0 GPU VMs (total spend ~$0.61); empty m11/m12 placeholder branches deleted (they had ZERO commits ? the "13k pending lines" the owner saw = untracked `artifacts/`+`scratch/`+sibling docs, NOT my work).** ? **FRESH-SESSION ICEBREAKER = bottom of [[session-handoff]] §"You are here".** ? Owner's #1 NEXT = make it **UI-DRIVEN + alpha-shareable** (khelsutra `web/` B-P2 ingest screen + M7 Cloud-Run *service* deploy + M9-on + a compute picker), THEN M11(flywheel, privacy-sensitive)/M12(gdrive-api). [[gcp-vm-always-delete]] [[gcp-ops-agent-enable]] [[ui-driven-cloud-serving-gap]] [[working-agreement-merging]] [[feedback-launch-quality-bar]] [[compute-stack-gcp-only]]

Checkpoint: 2026-06-21 (STALENESS-FIX / Path-B verify) ? owner: pick up Path B + "fix all stale docs and stop next sessions spiralling into the staleness." **? Verified-FIRST finding: B-P1 was ALREADY DONE** ? `POST /api/process` accepts `gs://`/`gcs://` since **M2/#280** (resolve?skip-exists-404?worker fetch?same pipeline?queryable API-DB `video_id`) + TESTED; confirmed by inline file:line trace AND a **3-skeptic verification workflow (0/3 refuted, gaps:[])**. So the doc-scoped "build the gcs ingest endpoint" was a phantom; Path B = **SPA-only (B-P2)**. **DONE this session:** (1) **khelsutra [#55] MERGED** (master `d0461fd`) ? corrected `DRY_RUN_CUJ` §1/2/4/6/§8 + `LOCAL_APPLIANCE` (`/api/health` EXISTS now M1/#268; bucket-URI ingest NOT a net-new endpoint; T12 engine-half ? done) + `README`; added the anti-spiral convention *"re-verify engine-surface `[gap]`/`net-new` claims at a cited file:line ? the engine ships PRs between sessions."* (2) **Memory steering corrected** (this NEXT-SESSION-FOCUS block + the dry-run checkpoint were stale "local-path-only") + new lesson [[lesson-verify-engine-surface-before-scoping]] + MEMORY.md index. (3) **engine [#284] MERGED** (master `06ff4b7`, all CI green incl. full suite 2m32s) ? regression GUARD `test_process_bucket_uri_yields_queryable_video_id` (parametrized gs://+gcs://) pinning `video_id==MD5(fetched bytes)` (not the URI) + `POST /api/query` queryability (the SPA contract; previously only LOCAL had it); built+removed in an isolated worktree off `origin/master` (engine checkout left on its concurrent-session branch). **? NEXT: B-P2** = the SPA ingest/progress screen (enter gs:// URI ? /api/process ? poll /api/jobs ? load video_id; back-half already works) ? **100% `web/` + i18n (en/kn/hi), ZERO engine code**; ready sketch in `DRY_RUN_CUJ.md` §2 + the workflow synthesis (`web/src/{api/types,api/client,hooks/useIngestJob,components/IngestPanel}.tsx` + `App.tsx` videoId?state + `ingest.*` keys). Then opt-3 droplet (deploy fixed engine + GCS creds + #281 fix ? re-run CUJ on Linux). **? Repos SYNCED TO HEAD at session end:** khelsutra master `d0461fd`; **engine** working tree `06ff4b7` ? NOTE the engine main checkout couldn't switch to `master` (sibling worktree `tier1-windowing` holds it), so it's the merged-but-remote-gone `feat/serving-m4-device-context` branch **reset --hard to origin/master + upstream repointed to origin/master** (safe: tracked tree clean + the 2 divergent commits were the M4 #283 squash, byte-identical to master; old tip tagged `_pre_sync_m4_44c9f01` + reflog). So `git pull --ff-only` works there now; the branch NAME is cosmetic (points at master head).

695. assistant (2026-06-22T09:57:43.389Z)

The handoff is dense and current. Adding one concise checkpoint for this session's brand-repo work, in the file's reverse-chron style.

696. user (2026-06-22T09:57:52.916Z)

The file C:\Users\avidu\.claude\projects\C--Users-avidu-Projects-badminton-highlight-indexer\memory\current-state.md has been updated successfully. (file state is current in your context ? no need to Read it back)

697. assistant (2026-06-22T09:58:10.150Z)

All done ? clean checkpoint reached. Here's the session wrap:

? This session (khelsutra brand repo, synced from `#93` ? now `master 7dc7de5`)

1. Synced to remote head ? was already up to date at `#93`; lots had landed since my last context (org consolidation, 6-language l10n, per-rally MVP, alpha deploy).

2. Stale branches pruned ? verified all **6** local branches mapped to merged PRs (#80/#58/#86/#88/#87/#77), deleted them, pruned remote-tracking refs. Only `master` remains.

3. Doc-freshness sweep ? instead of eyeballing, I ran it as two fan-outs:
- **15-agent audit** (one verifier per doc, every claim checked at `file:line`, cross-repo against the engine's `origin/master`) ? **44 stale findings**. The pattern was consistent: docs froze at their authoring PR while the SPA, engine API, localization, and alpha deploy shipped past them.
- **12-agent fixer fan-out** ? **63 surgical edits across 13 docs**.
- I **independently re-verified every load-bearing engine claim** myself before merging ? `/api/{health,capabilities,videos}`, compile's `download_url`, `native_wasb_runner.py` on master, the `auth.py`/`startup.py` symbols, and the subtle `RALLY_ALLOWED_HOSTS`-replaces-loopback semantics (that one was a genuine **runbook-correctness fix**, not just freshness). Brittle engine line-anchors became symbol refs (`main.py::compile_highlights`) so they won't rot again.

Honesty held throughout: `README` "pre-alpha" ? alpha; `web/README` "shell" ? real console; status flips kept their caveats (P6 stays `?`, native-WASB "not production-verified", footage screenshots owner-gated, historical changelogs preserved). The 2 docs that audited `fresh` were left untouched.

Two PRs, both merged on green CI (site + web + layout): [#94](https://github.com/avidullu/khelsutra/pull/94) (hygiene) · [#95](https://github.com/avidullu/khelsutra/pull/95) (doc sweep). One bug caught by my own verification: a `.gitignore` trailing-comment that silently matched nothing ? fixed before it shipped.

? State

- **0 open PRs, only `master` locally, working tree clean.** Handoff checkpoint added to `current-state.md`.
- The `Queued` items that remain are genuinely **owner-account** actions (Cloudflare HTTPS/HSTS + Analytics token, DMARC, the `khelsutra` repo?org transfer, clean D1 test rows, revoke the droplet SSH key) ? none are code I can complete.

That's a solid clean slate to pick up from. Appreciate the partnership today! ?