AELP / AELP2 — System Learnings & First■Wave Outputs

Generated 2025-09-29 12:57 UTC

1) Core problem and why

We need to simulate real■life response to creatives and placements so the RL agent can learn policies without burning budget. AELP2 turns historical Meta performance and vendor creative evidence into placement■aware baselines and calibrated forecasts, then uses an offline bandit simulator to learn allocation policies and output a slate and launch playbook.

2) Architecture — AELP↔AELP2 and RL simulator

```
AELP (foundational experiments, legacy pipelines) v AELP2 (production pipelines +
planner + RL sim) | | Meta Insights Vendor/Ad Library v \ /
+----------------------------+ | Ingestion + Normalization |
+----------------------------+ | v +----------------------------+ | Features +
Ranker Scoring | | (new■ad p_win, lcb) | +----------------------------+ | v
+----------------------------+ +-------------------+ | Baselines (by placement) |
---> | Monte■Carlo | +----------------------------+ | CAC/Volume | |
+-------------------+ v | +----------------------------+ v | Offline RL Simulator
| -------> Portfolio policies +----------------------------+ | | Planner UI/API v
| +----------------------------+ v | Launch Playbooks (Meta) | --------> Live
tests & feedback +----------------------------+
```

```
RL Simulator (offline) -------------------- Inputs: - Per■blueprint priors: p_win,
novelty - Baseline draws: CPM, CTR, CVR by placement (triangular p10/p50/p90) -
Budget constraints & caps Loop: 1) Sample CTR/CVR/CPM → estimate
signups/impressions and CAC 2) Thompson■sample success rate per arm (signups/imps
proxy) 3) Allocate daily budget to top■sampled arms (with caps) 4) Update
posteriors using simulated outcomes 5) Emit daily allocations + expected
CAC/signups Outputs: - rl_offline_simulation.json (ranking, history) - Portfolio
recommendation (8–12 arms + early■stop rules)
```

3) Data pipelines and system flow

• Meta→BQ by placement with backoff/window slicing. • Vendor/Ad Library via SearchAPI or CSV; normalized into creative_objects. • Features + new■ad ranker scoring (p_win + conformal lcb). • Baselines (CPM/CTR/CVR) from US paid data; MC forecasts for CAC/volume. • Offline RL sim (Thompson sampling on signups/imps proxy) to propose a portfolio. • Planner API/UI packages $30k/$50k with playbooks.

4) What we learned (why this design)

• Placement■aware baselines reduce drift; CVR clamps avoid outliers. • Scoring with p_win helps prioritize motifs; conformal lower bounds provide caution. • Offline RL accelerates exploration safely (policies pre■tested); early■stop rules protect downside. • US Ad Library has gaps; augment with SearchAPI/vendor CSV; rely on our own BQ for performance priors.

5) First■wave outputs (recommended slate)

Security $30k/day → signups p50: 153.8, CAC p50≈$114.9 combined

Balance $30k/day → signups p50: 368.4

Combined $60k/day → signups p50: 522.2, CAC p50≈$114.9, revenue p50≈$74,972

Security top 8 (id, p_win, su_p50, CAC p50)

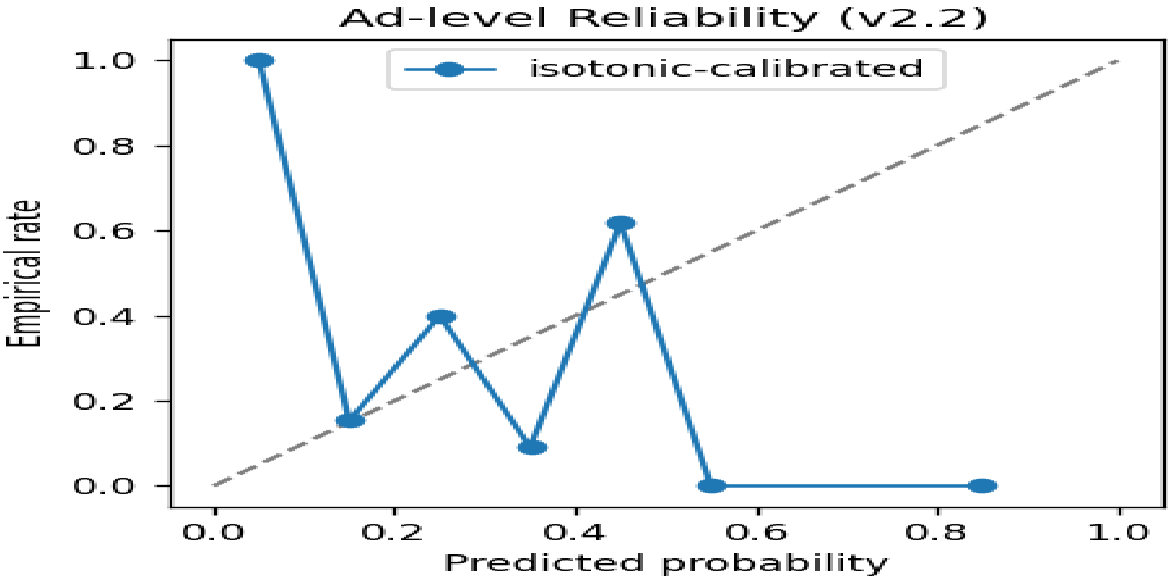
- bp_0042 — p_win 0.222; su_p50 22.7; CAC $166
- bp_0113 — p_win 0.222; su_p50 22.4; CAC $167
- bp_0114 — p_win 0.222; su_p50 22.4; CAC $168
- bp_0117 — p_win 0.222; su_p50 22.6; CAC $166
- bp_0116 — p_win 0.116; su_p50 18.2; CAC $206
- bp_0041 — p_win 0.092; su_p50 17.5; CAC $214
- bp_0001 — p_win 0.000; su_p50 14.0; CAC $267
- bp_0002 — p_win 0.000; su_p50 14.1; CAC $266

Balance top 8 (id, p_win, su_p50, CAC p50)

- bpbal_0001 — p_win 0.706; su_p50 45.9; CAC $82
- bpbal_0002 — p_win 0.706; su_p50 46.1; CAC $81
- bpbal_0004 — p_win 0.706; su_p50 46.2; CAC $81
- bpbal_0006 — p_win 0.706; su_p50 46.5; CAC $81
- bpbal_0009 — p_win 0.706; su_p50 46.3; CAC $81
- bpbal_0010 — p_win 0.706; su_p50 45.7; CAC $82
- bpbal_0014 — p_win 0.706; su_p50 45.6; CAC $82
- bpbal_0017 — p_win 0.706; su_p50 46.1; CAC $81

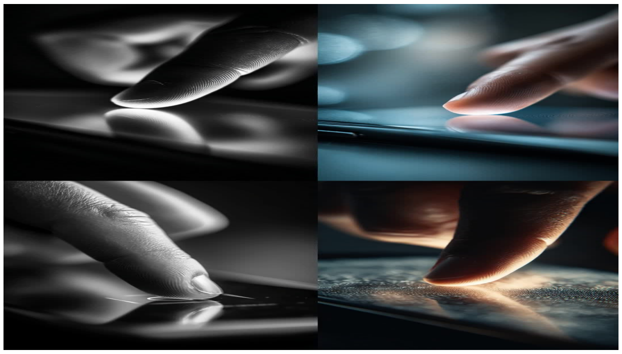
6) Accuracy & coverage

Ranker precision@5: 0.26666666666666667, precision@10: 0.3



7) Example creatives (thumbnails)

cool_preview_v1.jpg



orig_freeze_the_chaos_2.jpg

orig_spot_the_scam_1.jpg

8) Next steps

• Complete 90■day by■placement backfill and recalibrate. • Expand Balance set and run offline RL across offer variants. • Launch controlled live test with early■stop rules and measure lift vs. historical CAC.