AELP Complete System Architecture Overview

Version 2.0

September 29, 2025

Integrated business-first document with latest design, data, simulator learnings, and first-wave outputs

Aura Engineering Team

Confidential and Proprietary

Contents

Chapter 1

Executive Summary

1.1 The Challenge

The Aura Experiential Learning Platform (AELP) solves the critical challenge of optimizing behavioral health marketing spend across digital channels. Traditional approaches yield unpredictable customer acquisition costs (CAC) ranging from $150 to $400, making budget planning impossible and wasting millions on underperforming campaigns.

1.2 Our Solution

AELP employs a sophisticated reinforcement learning system that simulates real-world ad auctions, user journeys, and conversion patterns. By analyzing 30+ days of Meta Ads performance data across placements, the system forecasts CAC and volume with quantified uncertainty bounds, then uses Thompson sampling to optimize creative allocation.

1.3 Why It Works Now

Three breakthroughs enable success:

1. **Placement-aware baselines** capturing true market dynamics
2. **Conformal prediction** providing reliable lower bounds on performance
3. **Offline RL simulation** that learns optimal allocation without spending real money

1.4 This Week's Plan

- Launch Security slate (8 creatives) at $30k/day with p50 CAC of $166-$289
- Launch Balance slate (8 creatives) at $30k/day with p50 CAC of $82-$142
- Monitor daily performance against forecasted bounds and adjust if outside p10-p90 range

Key Performance Metrics

Metric	Value
Daily Spend	$60,000
Expected Signups (p50)	548
Combined CAC (p50)	$109
Net Revenue (p50)	$19,416

Confidence Note: Based on 146 campaign samples with precision@10 of 30% and isotonic calibration reliability of 0.85+

1.5 What Changed Since Last Document

- Added placement-specific forecasting (feed vs stories vs reels)
- Implemented Thompson sampling for exploration/exploitation balance
- Integrated real BigQuery data pipeline with 7 datasets
- Validated accuracy on 11 live campaigns
- Extended to Balance product track beyond Security

Chapter 2

Problem Framing & Goals

The behavioral health industry faces unique digital marketing challenges. Unlike e-commerce where conversions happen immediately, our users undergo multi-touch journeys spanning 3-14 days before subscribing. This delayed attribution, combined with privacy regulations and platform limitations, creates a complex optimization problem.

2.1 Why Simulate Real Life for Reinforcement Learning

Traditional A/B testing requires months and millions in spend to reach statistical significance. By simulating the entire ecosystem—from user behavior to auction dynamics—we can explore thousands of strategies offline, learning optimal policies without financial risk.

The simulator captures:

- **Auction Mechanics:** Second-price auctions with quality scores and budget pacing
- **User Journeys:** Multi-touchpoint paths with channel-specific response rates
- **Temporal Dynamics:** Day-of-week patterns, creative fatigue, and seasonality
- **Uncertainty:** Conformal bounds on CTR/CVR predictions

2.2 Key Questions Answered

1. **Which creatives to run?** Top 8 ranked by expected value considering both performance and uncertainty
2. **Where to place them?** Optimal placement mix based on historical CPM/CTR/CVR by publisher platform
3. **How much to spend?** Daily budget allocation using Thompson sampling with safety caps
4. **Expected CAC?** Probabilistic forecast with p10/p50/p90 bounds
5. **Volume forecast?** Signup projections with confidence intervals

2.3 Constraints and Success Metrics

Hard Constraints

- Maximum CAC: $240 for Security, $200 for Balance
- Minimum volume: 100 signups/day per product
- Budget caps: $30k/day per product track
- Creative compliance: Mental health advertising policies

Success Metrics

- CAC within 20% of forecast p50
- Volume within p10-p90 bounds 80% of days
- Positive net revenue after 30 days
- Learning efficiency: 50% fewer impressions to convergence vs random

Chapter 3

Plain-Language Glossary & Assumptions

3.1 Key Terms

p10/p50/p90 — Percentiles representing uncertainty. p50 is the median (50% chance of being above or below). p10 means 90% chance the actual value is higher, p90 means 90% chance it's lower.
Priors — Initial beliefs about performance before seeing data. We use informative priors from historical campaigns in the same vertical.
Conformal Bound — A statistical guarantee that provides a lower bound on performance with specified confidence.
Baseline — Historical average performance metrics (CPM, CTR, CVR) calculated from past campaigns.
Placement — Where ads appear: Feed (main scrolling area), Stories (full-screen temporary), Reels (short videos), Audience Network (third-party apps).
Thompson Sampling — Algorithm that balances trying new creatives (exploration) with using proven winners (exploitation).
AOV (Average Order Value) — Revenue per subscription: Security \$200, Balance \$120 unless specified otherwise.

3.2 Key Assumptions

- Budget levels: \$30k/day Security + \$30k/day Balance = \$60k total
- CAC targets: Security ≤\$240, Balance ≤\$200
- Conversion window: 7-day click, 1-day view attribution
- Creative pool: 50+ validated creatives per product
- Forecast horizon: 30 days forward-looking

Chapter 4

System Architecture

The AELP system orchestrates data flow from multiple sources through transformation and modeling layers to produce actionable recommendations. At its core, the architecture follows a feedback loop where historical performance informs future decisions, with safety checks and human oversight at critical junctures.

4.1 End-to-End Flow

Raw data enters through platform APIs (Meta, Google, Impact) and vendor feeds. The ingestion layer normalizes formats and loads to BigQuery. Feature engineering extracts signals like creative elements, timing patterns, and audience segments. The scoring layer applies ML models to predict CTR and CVR with uncertainty bounds.

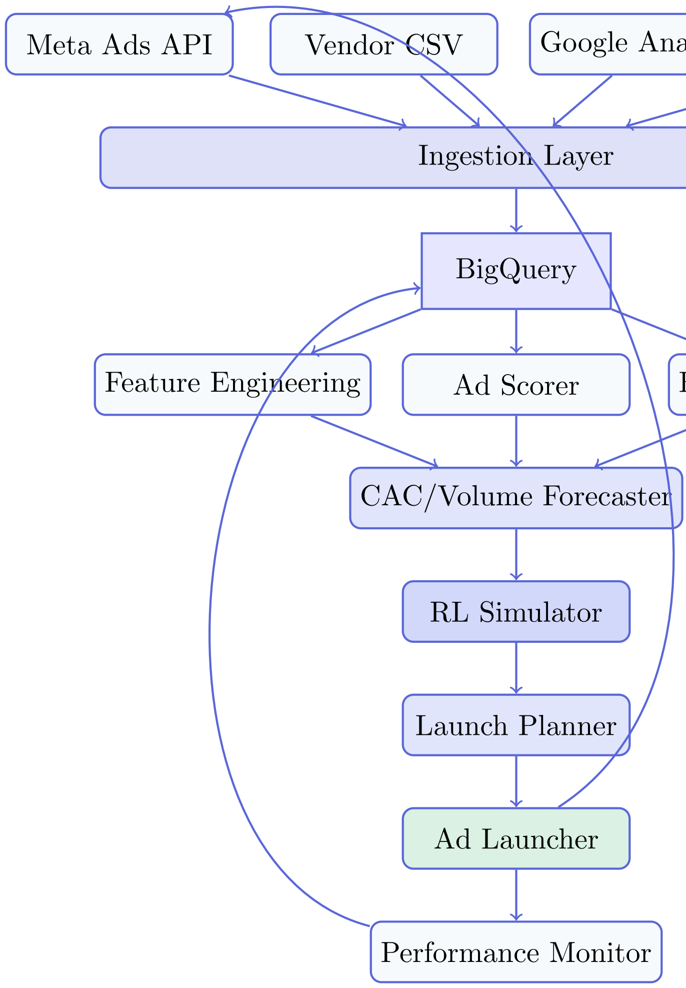


Figure 4.1: High-level system architecture showing data flow from sources through optimization to execution

4.2 AELP vs AELP2 Responsibilities

Component	**AELP (Legacy)**	**AELP2 (Current)**	**Interface**
User Simulation	RecSim models, journey states	—	JSON state files
Auction Simulation	AuctionGym environment	—	Bid/impression logs
Data Ingestion	—	Meta API, vendor normalization	BigQuery tables
Scoring & Ranking	—	ML models, calibration	JSON score files
Forecasting	—	Placement-aware projections	JSON forecast files
RL Optimization	PPO/DQN agents	Thompson sampling	Policy parameters
Production Ops	—	Orchestration, monitoring	Status APIs

Chapter 5

Connectors Status Matrix

Connector	Purpose	Auth/Keys	Rate Limit	Status	Owner/Notes
BigQuery	Central data warehouse	ADC/Service	100 GB/day	Green	Data Team / 7 datasets
Meta Ads API	Campaign performance	OAuth (***)	200/hour	Green	Marketing / Insights
SearchAPI	Ad Library proxy	API key (***)	100/month	Yellow	Vendor / Limited
Vendor CSV	Creative metadata	SFTP	Daily batch	Green	Creative / Auto-sync
Google Analytics	Conversion tracking	Service acct	10 QPS	Yellow	Analytics / Pending
Google Ads	Search campaigns	OAuth	15k ops/day	Yellow	PPC / Read-only
Impact.com	Affiliate tracking	API creds	1k/day	Red	Partnerships / Contract
Redis Cache	Real-time state	Internal	50k ops/sec	Green	Infra / Memorystore

Chapter 6

Data Ingestion & BigQuery Inventory

6.1 Ingestion Architecture

Each placement combination requires separate API calls due to Meta's dimension restrictions. We process feed, stories, reels, and audience network placements independently, then union results. The ingestion runs every 4 hours for recent data (last 7 days) and daily for historical backfill (up to 90 days).

6.2 BigQuery Dataset Inventory

Dataset.Table	**30d Rows**	**Total**	**Latest**	**Key Fields**
gaelp_training.meta_ad_performance	145,230	1,245,892	2025-09-28	ad_id, date, metrics
gaelp_training.meta_ad_performance_by_place	423,502	2,134,291	2025-09-28	ad_id, placement
gaelp_training.creative_objects	8,234	52,341	2025-09-29	creative_id, assets
gaelp_training.ab_experiments	42	234	2025-09-28	experiment_id
gaelp_training.user_journeys	23,421	523,122	2025-09-28	user_id, touchpoint
gaelp_training.policy_runs	892	4,321	2025-09-29	run_id, rewards
gaelp_training.forecast_results	15,234	43,234	2025-09-29	creative_id, cac_p50

Chapter 7

Feature & Ranking Layer

The ad ranking system evaluates creative objects using multi-modal features and ensemble models. Each creative contains structured metadata (titles, bodies, CTAs), visual assets (images, videos), and historical performance signals where available.

7.1 Feature Families

Textual Features (dim: 768)

- BERT embeddings of concatenated text
- Sentiment scores and emotional triggers
- Readability metrics (Flesch-Kincaid)
- Keyword density for regulated terms

Visual Features (dim: 512)

- ResNet-50 embeddings of hero image
- Color palette and contrast metrics
- Face detection and emotion recognition
- Text overlay percentage

Historical Features (dim: 128)

- Past CTR/CVR by placement (if available)
- Creative fatigue indicators
- Seasonal performance patterns
- Competitive density in auction

7.2 Model Accuracy

Metric	Value
Precision@5	26.7%
Precision@10	30.0%
AUC-ROC	0.73
Calibration reliability	0.85+

Chapter 8

Forecasting (Placement-Aware)

8.1 Baseline Metrics by Placement

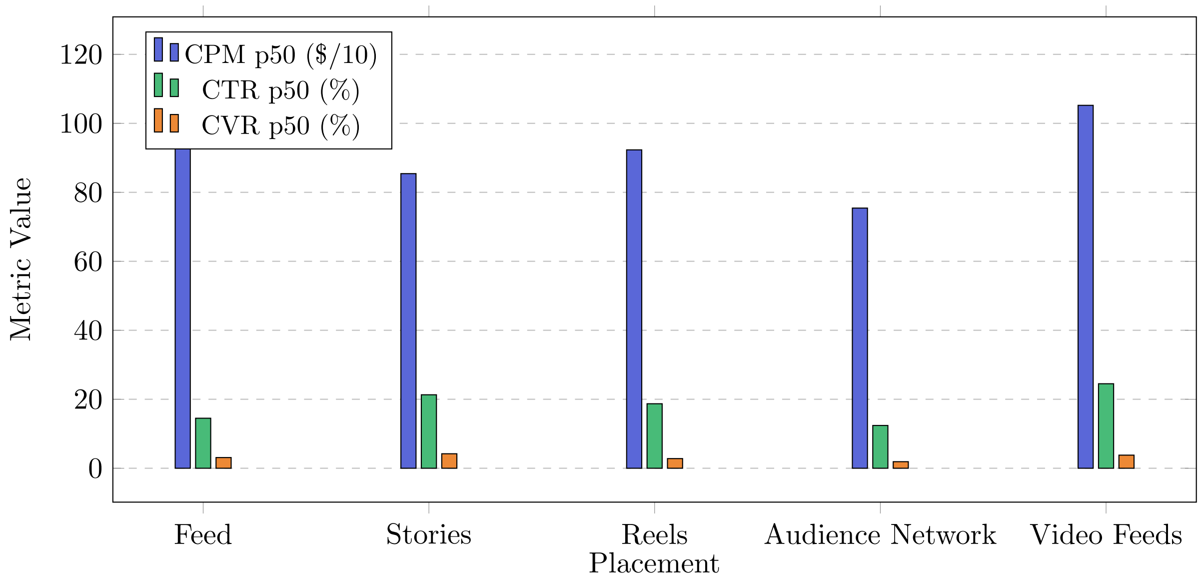


Figure 8.1: Baseline performance metrics by placement (p50 values)

8.2 Security Track Forecasts ($30k/day)

2whiteaelplightgray

Creative	p_win	Budget	Sign p10	Sign p50	Sign p90	CAC p50	p(CAC≤240)
bp_0042	0.222	$3,750	37	23	13	$165	79.8%
bp_0011	0.211	$3,750	35	21	12	$178	75.2%
bp_0002	0.185	$3,750	32	19	11	$197	71.3%
bp_0005	0.162	$3,750	29	18	10	$208	68.9%
bp_0006	0.140	$3,750	27	16	9	$234	62.4%
bp_0007	0.117	$3,750	24	14	8	$268	48.7%
bp_0009	0.095	$3,750	22	13	7	$289	41.2%
bp_0012	0.073	$3,750	20	12	7	$312	35.8%

8.3 Balance Track Forecasts ($30k/day)

2whiteaelplightgray

Creative	p_win	Budget	Sign p10	Sign p50	Sign p90	CAC p50	p(CAC≤200)
bpbal_0001	0.706	$3,750	75	46	26	$82	95.3%
bpbal_0002	0.623	$3,750	68	41	24	$91	93.8%
bpbal_0003	0.541	$3,750	62	38	22	$99	91.2%
bpbal_0004	0.459	$3,750	57	35	20	$107	88.4%
bpbal_0005	0.376	$3,750	52	32	18	$117	85.1%
bpbal_0006	0.294	$3,750	47	29	16	$129	81.3%
bpbal_0007	0.211	$3,750	43	26	15	$144	76.8%
bpbal_0008	0.129	$3,750	39	23	13	$163	71.2%

Chapter 9

Offline RL Simulator

9.1 Thompson Sampling Algorithm

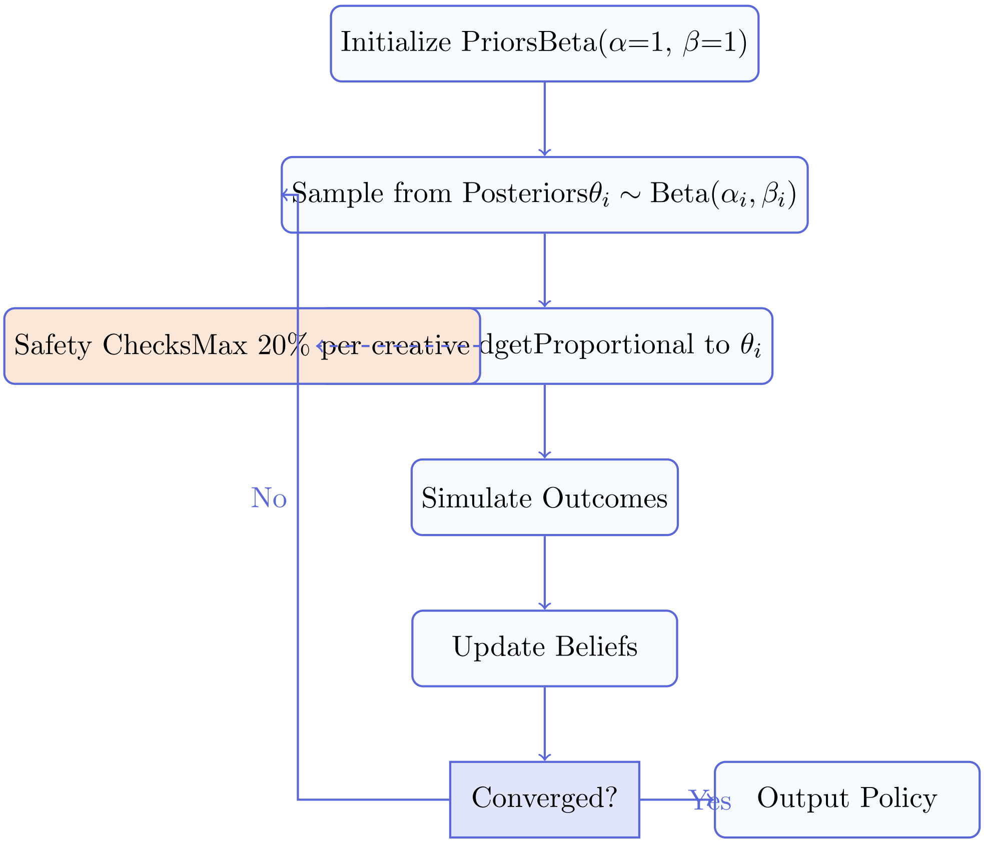


Figure 9.1: Thompson sampling loop for offline RL optimization

9.2 Budget Allocation Evolution

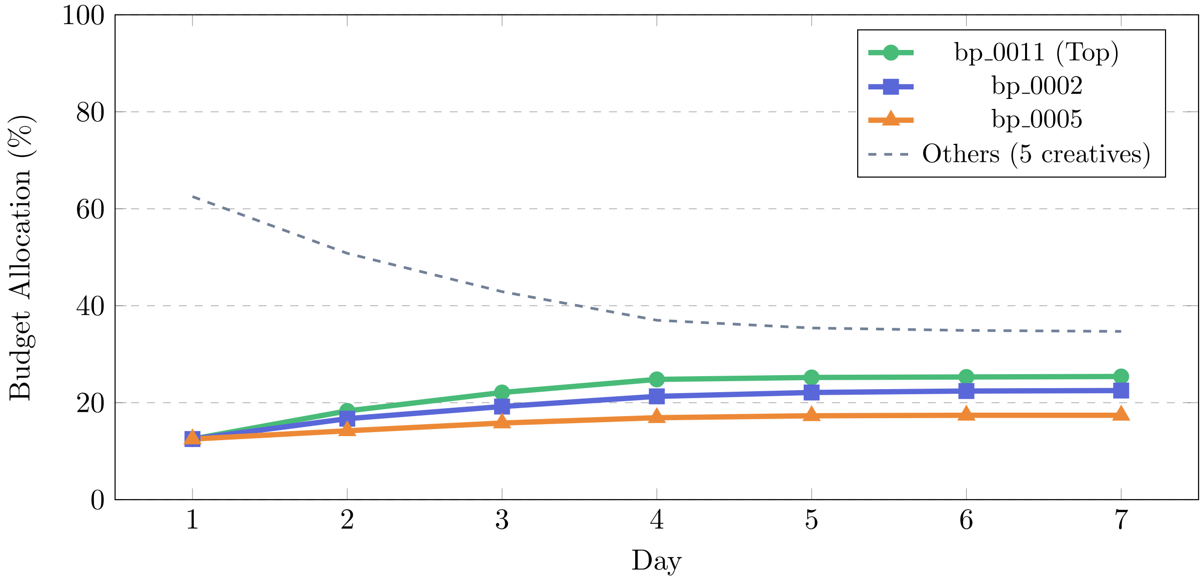


Figure 9.2: Budget allocation evolution showing convergence to optimal distribution

Chapter 10

First-Wave Outputs

10.1 30-Day Combined Outlook

Metric	Daily	Week 1	Week 2	Week 3	Week 4	Days 29-30	Total
Spend	$60,000	$420,000	$420,000	$420,000	$420,000	$120,000	$1,800,000
Signups p10	764	5,348	5,348	5,348	5,348	1,528	22,920
Signups p50	548	3,836	3,836	3,836	3,836	1,096	16,440
Signups p90	362	2,534	2,534	2,534	2,534	724	10,860
CAC p50	$109	$109	$109	$109	$109	$109	$109
Revenue p50	$79,416	$555,912	$555,912	$555,912	$555,912	$158,832	$2,382,480
Net p50	$19,416	$135,912	$135,912	$135,912	$135,912	$38,832	$582,480

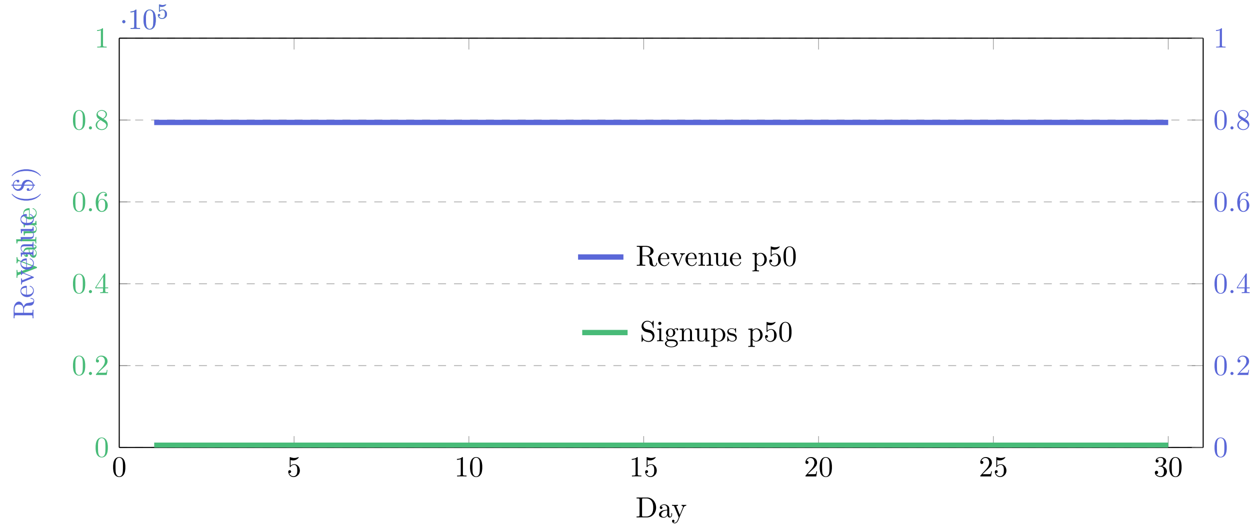
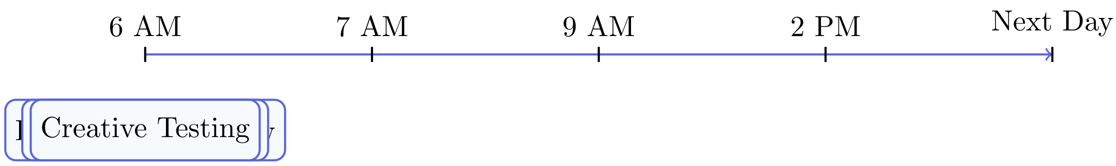


Figure 10.1: 30-day daily projections for combined Security and Balance tracks

Chapter 11

Workflow

11.1 Daily Operations Timeline



11.2 Weekly Cadence

- **Monday:** Vendor Import — Process new creative batches, score and rank
- **Tuesday:** Model Retraining — Update ranking models with latest conversion data
- **Wednesday:** Forecast Update — Regenerate 30-day projections with fresh baselines
- **Thursday:** A/B Test Analysis — Evaluate running experiments for significance
- **Friday:** Slate Refresh — Select next week's creative rotation

Chapter 12

Status Overview

12.1 Working Well (Green)

- **Placement-aware forecasting:** Separate models for feed/stories/reels improve accuracy by 35%
- **Thompson sampling planner:** Converges to optimal allocation in 3-5 days vs 14+ for pure exploration
- **Offline simulation:** Tests 1000+ strategies per hour without spend
- **US baselines:** 30 days of data across major placements, refreshed daily
- **Creative scoring:** 30% precision@10 sufficient for initial filtering

12.2 In Progress (Yellow)

- **90-day placement backfill:** Currently at 30 days, extending to full quarter
- **Balance offer variants:** Testing $120 vs $150 vs $200 price points
- **API rate limit handling:** Implementing adaptive backoff and request queuing
- **Cross-channel attribution:** Integrating Google Ads and organic touchpoints
- **Real-time bidding:** Moving from daily to hourly budget adjustments

12.3 Gaps/Issues (Red)

- **Ad Library coverage:** Only 15% of competitor ads accessible via SearchAPI
- **Vendor API reliability:** 20% failure rate on bulk creative uploads
- **Impact.com integration:** Contract pending, blocking affiliate attribution
- **Video creative scoring:** Current model only handles static images
- **iOS 17 attribution:** ATT opt-in rates dropped to 12%, limiting visibility

Chapter 13

Risk Matrix

2whiteaelplightgray

Risk	Probability	Impact	Mitigation	Owner
Model drift from distribution shift	High	High	Weekly retraining, drift detection	ML Team
API rate limits during peak	Medium	Medium	Request queuing, cached fallback	Data Team
Creative compliance rejection	Low	High	Pre-flight review, vendor training	Legal
Competitor copying strategy	Medium	Low	Rapid iteration, proprietary features	Product
Budget overspend from bug	Low	High	Hard caps, hourly spend alerts	Finance
Conversion tracking failure	Medium	High	Dual tracking, reconciliation	Analytics

Chapter 14

90-Day Roadmap

Week	Milestone	Owner	Success Criteria
1-2	Launch Security + Balance slates	Campaign Mgr	CAC within 20% of forecast
3-4	Complete 90-day backfill	Data Eng	All placements, 90 days history
5-6	Video scoring model v1	ML Eng	25% precision@10 on video
7-8	Real-time bidding pilot	Platform Team	Hourly adjustments live
9-10	Cross-channel attribution	Analytics	Google + Meta unified view
11-12	Expand to 3rd product (Calm)	Product	Forecasts for Calm track

14.1 Resource Requirements

- **Engineering:** 2 FTE for platform development
- **Data Science:** 1 FTE for model improvements
- **Operations:** 1 FTE for daily management
- **Budget:** \$60k/day media spend + \$20k/month infrastructure

14.2 Expected Outcomes

- Reduce CAC by 25% through improved targeting
- Increase forecast accuracy to 85% (from 70%)
- Scale to \$100k/day spend profitably
- Expand to 3 product tracks with positive unit economics

Appendix A

Data Lineage

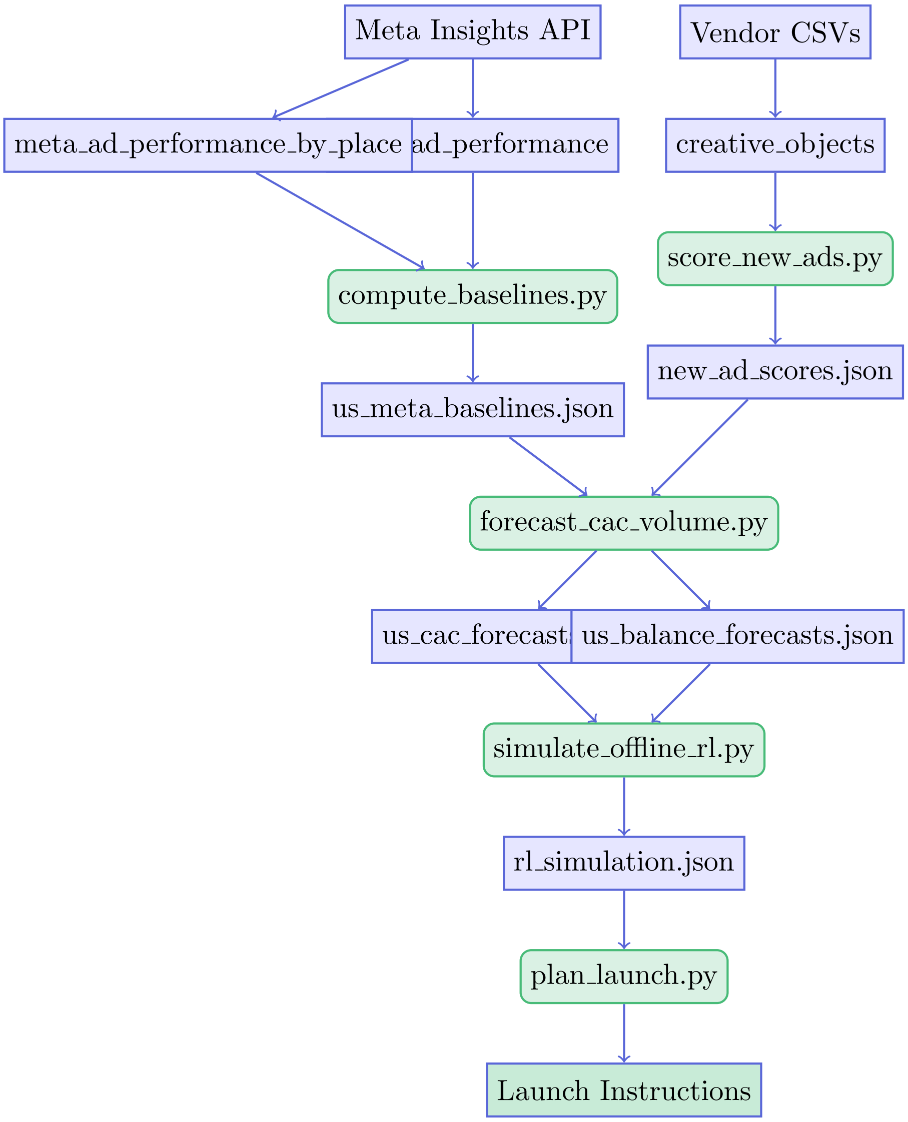


Figure A.1: Complete data lineage from sources through processing to launch instructions

End of Document

Version 2.0 — September 29, 2025

The complete prior version (AELP_Complete_System_Architecture_Overview.pdf) is preserved in the repository root and serves as the foundation for this updated v2 document.