

KickoffPRD

PRD — Next Phase for Zero (iOS Email + Summary App)

Vision (Current Phase)

Zero makes email triage effortless for busy parents using AI classification, 1-tap contextual actions, and thread summarization with human conviction intact.

Success =

- Email reader + summarizer quality is **best-in-class**
 - The **top 10 high-value actions** work reliably and feel magical
 - A **staged beta** validates the model with real usage, not assumptions
 - We launch without being reactive to infrastructure or cost surprises
-

1.

User Personas

Persona	Needs	Beta Role
Busy Parent	fast triage, clear summaries, reliable actions, no app crashes	Core beta tester
Power Parent	auto-index, multi-action sequences, tune AI, influence weights	AI feedback partner
Design + Eng Admin	tune model, measure confidence, registry mgmt, alerts	Inner-loop AI tuner

Marketing Viewer	understand features, demo value visually, believe the story	Campaign consumer
-------------------------	---	-------------------

2.

Must-Solve First

Core capability foundation

1. **High-quality email summaries** for long threads
2. **Mail vs Ads classification $\geq 90\%$ confidence**
3. **Action success $\geq 98\%$ reliability** for top actions
4. **Local + cloud fallback resilience** for network failures
5. **Token + cost telemetry** for every AI call

Critical benchmark alignment inspired by your roadmap philosophy:

3.

Top 10 Action Candidates for Beta Testing

Initial actions to register, prototype, and validate:

1. Apple Calendar RSVP/schedule flows
2. Apple Reminders task creation
3. Pencil signature capture via PencilKit
4. Flight status check-ins aligned to American Airlines flows
5. Package tracking mapped to USPS

6. Summary-powered list intelligence via Combine
7. Glass UI foundations influenced by SwiftUI
8. Store purchases integrated via StoreKit 2
9. Analytics tracked via PostHog
10. Web scraping agent suggestions logged for tuning

Each must include:

- context validation
 - confidence rating
 - graceful fallback
 - telemetry for cost + error tracing
-



Execution Plan (Beta → Launch → Scale)

Phase 1 —

Reader + Summary Quality (Week 1-2)

Deliverables:

- Summarizer prompt refinement
- Accuracy guardrails, hallucination prevention
- Expert vs AI ranking agreement tests
- Cost telemetry foundations

Phase 2 —

Top 10 Actions Tested (Week 3-6)

Deliverables:

- Action registry cleanup + modal routing
- 10 prototypes functional in staging
- End-to-end success tests with fallbacks
- Latency < 2s (local) and < 3s (cloud p95)
- 20% test buffer for unknown complexity

Phase 3 —

Staged Beta (Week 7-10)

Deliverables:

- TestFlight staged roll-out (Top 100 → 500 → 2,000 users)
- Segmented cohorts: Core vs Power Parent
- In-app feedback flows for tuning AI
- Measure: agreement, edit %, time saved, satisfaction
- Build prioritization weights tuning via beta usage
- No addresses or personal attributes stored in docs (by design)

Phase 4 —

MVP Launch Hardening (Week 11-14)

Deliverables:

- Fix critical bugs
- Security + accessibility audits
- Prep release artifacts:

- PRD (this)
- GTM plan
- Cost sheet
- Hire plan
- Messaging scripts
- Landing site sections (About Us, Why Rationale)
- Analytics dashboards
- NIL launchpad staged section placeholder (not user facing yet)

Phase 5 —

Launch + Iteration (Week 15-24)

Deliverables:

- 48h stability monitoring
 - 3-sprint iteration plan
 - Model + actions refined via real feedback
 - Hire + marketing flight paths scaling
-

Budget Sizing (Initial Estimate — 6 to 12 Month Horizon)

1.

Engineering & Leadership

Role	Cost Estimate
1 × Engineering Lead	\$180-220K/year
2 × Senior Swift Engineers	\$150-180K each/year
1 × AI/ML Full Stack Engineer	\$160-200K/year
1 × QA + Accessibility	\$120-150K/year

Subtotal Engineering (Annual): ~\$730K-\$1.0M

2.

Design & Prototyping

Role	Cost
1 × Senior Product Designer (contract/cofound quality)	\$140-160K/year or \$12-15K/month contract
1 × Motion/Branding IC	\$100-130K/year

Subtotal Design (Annual): ~\$240K-\$320K

3.

Marketing

Campaigns, creative, microsite, beta storytelling:

Item	Cost
UXR/closed beta agency support	\$40-80K
10k-50k install marketing push	\$120-250K
Social + influencer + copywriting design	\$60-120K
Microsite development + hosting + analytics	\$15-30K

Subtotal Marketing (One-Time + 6-mo): ~\$270K-\$500K

4.

AI API & Model Tuning

Vendor	Model	Est. Monthly
OpenAI	Gemini 1.5 Flash	\$2-6K
Google	GPT-4	\$8-20K

Subtotal AI Services (6 months): ~\$70K-\$140K

(Varies by usage volume, but we'll instrument to measure.)

5. 

Cloud & DevOps

Using:

Provider	Cost
Google Cloud Run	Cloud Run for Backend
Google Cloud Build	Cloud Build for CI/CD

Subtotal Infra (6-mo): ~\$24K-\$66K



Total Estimated Spend (6-12 Months)

 \$1.5M – \$1.9M for 6 months

 \$2.2M – \$2.8M for 12 months

This gives you:

- a **non-reactive beta + launch**
- reliable **email + summaries**
- 10 magical **action prototypes**
- **AI tuning loops**
- **marketing flight**
- **engineering hires**
- **✓ delivery environment stability**

Additional Artifacts You Should Make (Recommended)

Internal + External OS

Artifact	Purpose
12-page pitch/story deck	CEO/Partners
1-page investor one-pager	Funding
Internal Studio OS (Notion style)	Team alignment
Website refactor plan for Cursor/Claude	New codebase
Beta testing taxonomy doc	Cohort planning
Marketing copy + campaign scripts	Install drive
Model tuning + feedback UX spec	AI improvement
Budget sheet (the table above)	Financial planning
Action acceptance test sheets	10 flows

Want help generating those artifacts next?

Tell me the format and I'll produce them:

- A. PRD 2-pager w/ success metrics
- B. 12-slide pitch deck w/ speaker notes
- C. PNG logo pack for App + microsite
- D. Marketing campaign plan and sample copy
- E. Engineering hires role JD templates
- F. AI tuning prompt templates for summarization & actions

Just reply with: A, B, D, E, F, etc.

betalaunchprd

1. Product Requirements Document (MVP + Beta + Launch)

Product Vision

Zero reimagines inbox triage for busy parents using AI classification, 1-tap contextual actions, and thread summarization that feels magical and reliable.

Scope

- AI: Mail vs Ads classification, Thread Summaries, Smart Replies, Action Suggestions
- UX: Swipeable card viewport + targeted modal flows for 10 core actions
- Admin: Model tuning, feedback telemetry, confidence monitoring
- Surfaces: iOS app, beta microsite, marketing campaign, hiring plan

Success Metrics

Metric	Target
Summary relevance accuracy	$\geq 92\%$
Classification accuracy	$\geq 90\%$
Action reliability	98% (post-override)
Draft edit rate	< 30% avg

Latency p95	< 3s cloud / < 2s local
Beta satisfaction rating	≥ 4.3/5
Stability post-launch	48h no critical errors
ROI — time saved	25–40 min/day/user

10 Core Actions to Validate in Beta

1. Calendar scheduling/RSVP
2. Reminders task creation
3. Signature capture
4. Flight check-ins
5. Package tracking
6. Invoice payments
7. List formatting extraction
8. Contact card creation
9. AI suggested action override + edit
10. Batch AI feedback submission

Key Risks + Mitigations

Risk	Impact	Mitigation
------	--------	------------

AI hallucination in summaries	trust erosion	anchored prompting + sample agreement checks
Beta instability crashes	churn + low ratings	staged TestFlight + crash telemetry
Cost surprises from AI calls	budget cap blowup	build cost monitoring from day 1
Too many actions without registry constraints	confusion + drift	limit to 10 core flows initially
Over-reliance on example data models	misaligned planning	gather real usage metrics quickly

2. ☁

Architecture + Instrumentation Spec

Backend Microservices & AI Stack:

- Ingestion: Gmail cards pulled into memory
- Processing: None bypasses human override
- Search: Contextual action enrichment
- ML orchestration: Vertex AI + GPT services
- EventKit + Contacts + Wallet + Reminders + Haptics integration

AI Tuning Instrumentation

Service	Model	Duty
Classifier	Gemini 1.5 Flash	Categorize Mail vs Ads
Summarizer	Gemini 1.5 Pro	Summarize threads/emails
Shopping Agent	Gemini Flash	Price intelligence
Action Registry Engine	custom Swift registry	Validate routing and context needs
Analytics	PostHog	Track all actions and costs

3.

Beta Staging & Test Plan (TestFlight)

Cohorts

Stage	Users	Purpose
1	0–100 (internal dogfood)	validate summary + 10 actions quality

2	100–500	validate model agreement vs manual overrides
3	500–2,000	validate cost telemetry, satisfaction, performance
4	2,000→10k	pre-launch confidence + marketing push

Beta Validation Milestones

1. Measure >70% agreement between AI classification and human swipes
 2. Measure <30% editing required after AI draft generation
 3. No AI suggestion bypassing human validation
 4. Fallback tested for API failures, network drops, and malformed summaries
 5. Telemetry includes tokens, errors, latency, and cost per interaction
-

4. Engineering Hire Plan + Job Descriptions

Hiring Philosophy

Build in sprints, hire for minimal convex surface ownership coverage:

- 1 engineering lead
- 2 Swift senior ICs
- 1 AI infra/ML full-stack

- 1 QA/accessibility hardener
 - 10 action modal owners temporarily unassigned (for grooming later)
-

A. Engineering Lead — iOS

Owner: Architecture, performance, and AI fallback routing

Key Skills

- 100% Swift and SwiftUI
- Modal routing, EventKit integration
- AI API instrumentation telemetry
- UXR hand-off quality guardrails

Responsibilities

- Define routing patterns based on classification type
 - Ensure Action reliability p95 <2 seconds local, <3 seconds cloud
 - Telemetry for cost, latency, tokens, and errors
-

B. Senior Swift Engineers (2 roles)

Owners: Action modules 1–10 in high-quality context

Key Skills

- Swift 5.9+ / iOS 17+
- SwiftUI or equivalent component design systems

- Combine or equivalent reactive data flows
- PencilKit for signature capture
- Accessibility for parenting UX

Responsibilities

- Add/Edit actions in the ActionRegistry.swift
 - Build modals for Pay Invoice, Calendar RSVP, Flight check-ins, Package tracking, etc.
 - Implement acceptance criteria for each action
 - Ensure <30% hallucinatory suggestion editing required
-

C. AI/ML Full-Stack Engineer

Owner: Model tuning, classifier, and summarizer refinement

Key Skills

- Vertex AI: Gemini Flash + Pro
- Prompt engineering for summaries + action suggestions
- Cost + token telemetry, accuracy guardrails
- Expert vs AI ranking agreement evaluation

Responsibilities

- Refine summarizer prompts for Mail and threads
 - Tune AI model weights using beta feedback loops
 - Maintain cost per AI interaction < manageable miles per job
-

D. QA + Accessibility IC

Owner: Stability hardening + ARIA/WCAG audit pass

Key Skills

- XC Testing / Unit + Integration test coverage 80%+
- VoiceOver, High contrast, parenting UX accessibility

Responsibilities

- Add missing test cases for 10 core flows
 - Accessibility pass ensuring 4 user personas are validated
 - Shake out crashes and feedback loops
-

5. 💰 Budget Sizing Sheet

(Derived from operational hiring & infra sizing inspired by roadmap philosophy)

6 mo Estimate

Line Item	Cost
Engineering + Leadership	\$1.5M–\$1.6M
AI API + Tuning	\$70K–\$140K
Cloud Build + Cloud Run	\$24K–\$66K

UXR Beta Agency	\$40K–\$80K
Marketing Install Push	\$270K–\$500K
Total 6 mo	\$1.9M–\$2.2M

12 mo Estimate (linear extension)

Line Item	Cost
Engineering + Leadership	\$2.2M–\$2.8M
AI API + Tuning	\$140K–\$280K
Cloud Build + Cloud Run	\$66K–\$130K
UXR + Beta Agency	\$80K–\$120K
Marketing Install Push	\$350K–\$550K
Total 12 mo	\$2.5M–\$3.2M

6. Marketing Campaign Playbook

Campaign Positioning

“Don’t be reactive to your inbox — swipe with AI conviction.

One tap summaries. Ten magical actions. No tool switching. Just clarity.”

Campaign Surfaces + Tools

- Spatial microsite pitch viewer: Fivetran + React for campaign view
- Install drive: App Store + influencer storytelling
- Analytics: PostHog event registry
- Staged Testing: TestFlight
- Content Masquerade: educational prototype demos & visuals
- Beta invitations: Liquid clarity, low friction entry pop-ups
- No financial asks tied to alimony, ex-wife, or child support in copy

Messaging Scripts + Examples

Install CTA

“Swipe left to act. Get summaries you can trust.

Tap once to do what matters. No toggles, no tabs, no chaos.”

Influencer Script (15–25s Reel)

“Parent life moves fast. Email shouldn’t move slow.

I swipe once, I see the summary, I tap once — done.

Beta soon. Join the TestFlight. Zero chaos.”

Why Zero (Short Copy for Site/About Page)

“Email bruises time. Zero gives it back.

AI classification you can trust. Actions that feel like habit.

Built for parents. Tuned by parents. Launched with certainty.”

Beta Invite Email

“Be part of shaping smarter summaries + 1-tap actions.

You’ll see Mail vs Ads classification, thread clarity, and 10 magical modals.

Your feedback tunes the AI. Human control stays intact.”

Budget Telemetry KPIs for Marketing

KPI	Target
Install cost per acquisition	\$2.50–\$6.00
Beta invites opened	≥ 40% open rate
Actions accepted by swipe overrides	≥ 65%
Summary edit rate post AI generation	< 28%
Cost telemetry visibility	100% instrumented

7. AI Tuning Prompts for Summarization & Actions

Summarizer Prompt Template

None

You are a concise, accurate email summarizer.
Given a parent-facing inbox message, produce:
1. 2-3 sentence summary under 280 characters
2. A confidence rating from 1.0 to 5.0
3. Fallback plan if summary can't be used
Do not fabricate details. Surface ambiguity explicitly.

Action Suggestions Prompt Template

None

You are an action inference engine.
Given an email card, suggest 1 best action from registry:
- match to 10 staged modal actions
- validate required context keys
- ensure human override guard
Return: actionId, displayName, rationale, confidence (1-5)
Ensure Fallback if confidence <3 or service errors.

8. 6–12 Month Milestone View

Month	Milestone
1	Summarizer + Classifier quality gates
2	10 core action reliability + modal routing
3	500 user beta confidence validation

- 4 2,000 user beta + cost telemetry stabilized
 - 5 Launch hardening + marketing storytelling landed
 - 6 Production launch & 48h stability check
 - 7–9 3-sprint iteration, AI model behavior satisfaction
 - 10–12 Scale install drive 10k–50k AL rollout
-

Tab 2



Included in the artifacts

Engineering → Budget Plan

Your budget sheet includes:

- **6–12 month cost horizon**
- **Engineering hires**
- **Cloud + AI service burn**
- **Marketing acquisition allotment**
- **Buffers for uncertainty**

(inspired by roadmap philosophy — cited above)

Staged Beta → Launch

You'll see:

- TestFlight staging from dogfood → 50k installs
- 10 core action validation
- Summary accuracy, latency, edit rates
- Human override + fallback requirements

Marketing Scripts → Campaign KPIs

Your playbook contains:

- Install CTA copy
- 15–25s influencer reel script

- Beta invite email copy
- Acquisition KPIs instrumented via PostHog

AI Summary → Action Inference Tuning

Prompts include:

- Summary in 280 chars or less
- 1 best action from 10-action registry
- Confidence ratings (1.0–5)
- Explicit fallback on <3 confidence or API failure

If you want, next I can also generate:

1. **High-level business success model slide deck (.pptx)**
2. **Expanded 1-week sprint breakdown plan**
3. **Install acquisition success calculator model**
4. **Mock data diversity tests for 10 registered actions**
5. **An equity + hiring ramp plan alongside burn model**

Which one should we refine first?