DECISIONFLOW

# Benefits Realization and Delivery Review

*Planned versus delivered scope, KPI movement, AI and operating outcomes, variances, lessons, and investment decision*

**DELIVER PHASE | Version 1.0 | Illustrative 90-day controlled-scale program**

**EVIDENCE & PUBLISHING INTEGRITY**

This public portfolio edition distinguishes actual evidence, assumed continuity steps, simulated case-study outcomes, targets, and recommendations. Simulated customers, commercial results, benchmarks, incidents, quotations, and later-lifecycle outcomes are illustrative and must not be interpreted as results from a live commercial product. Evidence labels are retained throughout the package to support transparent and responsible interpretation.

## Document Control

| Control | Definition |
|---|---|
| Document code | DEL-07 |
| Version / status | 2.0 — Final Public Portfolio Edition — Simulated Delivery Scenario |
| Baseline / author | 14 July 2026 - Min Thu Kyaw |
| Lifecycle continuity | Actual MVP -> DEV-06 assumed productionization -> simulated launch -> Conditional Scale -> Deliver phase. |
| Delivery decision | Execute a 90-day controlled-scale program; keep collaboration, integrations, richer ingestion, and enterprise features evidence-gated. |
| Evidence rule | Actual, assumed, simulated, target, and recommendation statements are explicitly separated. |

### Actual Project Artifacts

| Artifact | Reference |
|---|---|
| Low-Fidelity Prototype | View prototype |
| High-Fidelity Prototype | View prototype |
| MVP Frontend Repository | Open repository |
| MVP Backend Repository | Open repository |

# Contents

# Review Purpose and Evidence Boundary

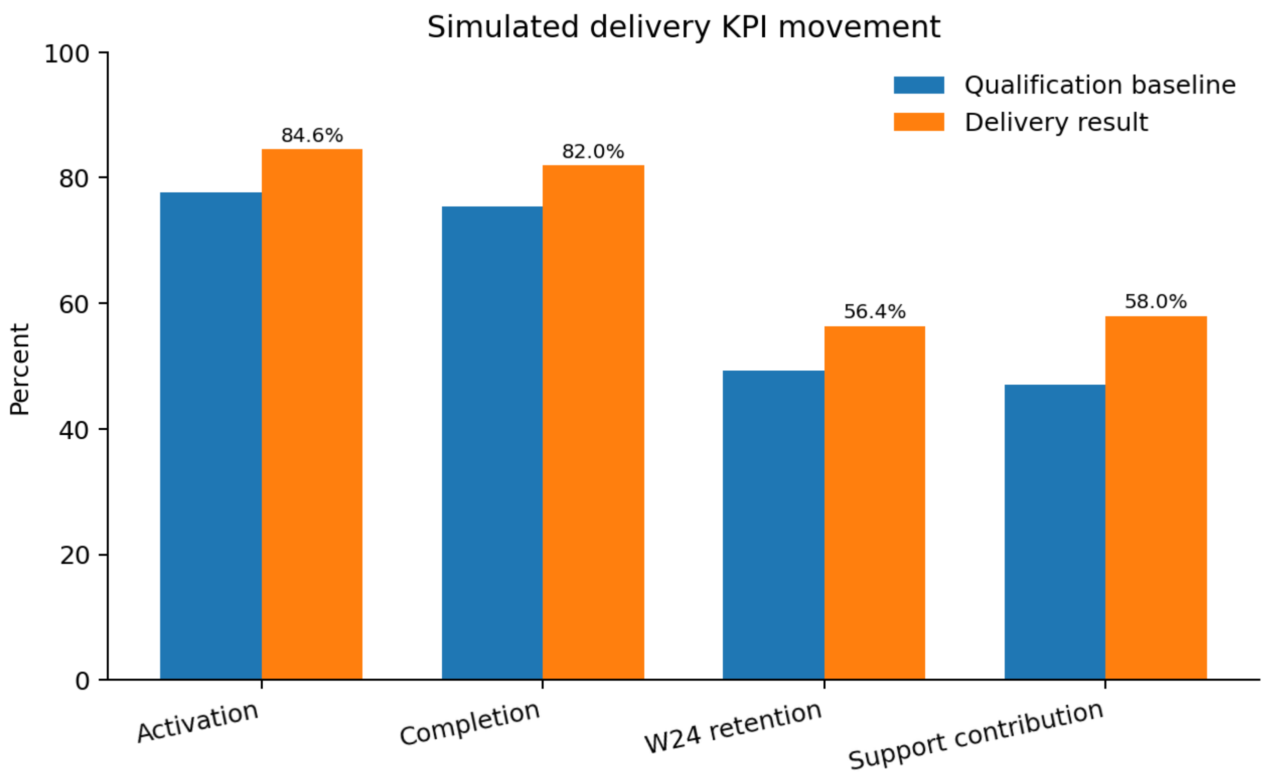
This review assesses whether the illustrative 90-day program delivered the conditions attached to the Qualify-phase Conditional Scale decision. It does not claim broad product-market fit or enterprise readiness. All delivery and outcome data in this report are simulated for portfolio demonstration.

**Decision standard -** Continue investment only where customer value, AI trust, economics, reliability, security, and delivery discipline improve together. A strong metric in one dimension does not override a failed safety or data-integrity gate.

# Planned Versus Delivered Scope

| Planned outcome | Delivered evidence | Variance |
|---|---|---|
| Recurring decision templates and second-decision journey | Templates, lifecycle prompts and retrieval onboarding released. | No material scope variance. |
| AI owner/date and calibration improvement | Owner 92.3%, due date 91.2%, ECE 0.08. | Passes gates. |
| Asynchronous processing and 40-job capacity | 39s p95 and 0.7% failures at 40 jobs. | Pass; 80-job profile still fails. |
| Support efficiency | Contacts/activation reduced to 0.24; contribution reaches 58%. | Pass. |
| Professional package validation | 15 paying accounts; 86.7% logo retention; 9.8m payback. | Pass with small sample. |
| Export / webhook evidence | Prototype and interviews completed. | Discovery only; no production integration. |
| Collaboration evidence | Jobs, permission concerns and willingness-to-pay themes documented. | No build commitment, as intended. |

# Benefits Realized




*Delivery KPI movement*

| Benefit | Baseline | Simulated result | Interpretation |
|---|---|---|---|
| Faster value | 77.7% activation; 8m 42s | 84.6%; 6m 55s | Examples and templates improve self-service first use. |

| Benefit | Baseline | Simulated result | Interpretation |
|---|---|---|---|
| More complete workflows | 75.4% start-to-finalize | 82.0% | Fewer state and comprehension barriers. |
| More durable use | 49.3% W24 retention | 56.4% | Passes controlled-scale condition, not broad-market proof. |
| More accountable AI | Due date 87.8%; ECE 0.12 | 91.2%; 0.08 | Key qualification gaps close. |
| Lower variable AI cost | $0.38/finalized decision | $0.29 | Routing and prompt efficiency improve margin. |
| Better service economics | 47% support-adjusted contribution | 58% | Self-service operating model passes gate. |
| Greater capacity | 40 jobs: 56s / 1.9% | 39s / 0.7% | Controlled-growth ceiling passes. |

# Commercial and Service Outcome

| Measure | Simulated result | Assessment |
|---|---|---|
| Paying accounts | 15 | Meets qualification condition; still a limited sample. |
| 90-day paid logo retention | 86.7% (13/15) | Passes 80% condition. |
| MRR | $6,840 | Growth reflects controlled cohort and Professional packaging. |
| Observed gross-profit payback | 9.8 months | Passes <=12-month condition; channel mix is narrow. |
| Gross margin | 84% | Healthy software margin. |
| Support-adjusted contribution | 58% | Passes scale gate. |
| Self-service activation | 72% | Improves operating leverage; needs larger-cohort validation. |

# Delivery Variances and Lessons

| Observation | Lesson | Decision implication |
|---|---|---|
| Templates improved activation and repeat use most in frequent-decision roles. | Recurring jobs matter more than generic feature breadth. | Keep the beachhead narrow and template strategy role-specific. |
| Mandatory owner/date confirmation added friction but reduced accountability errors. | Trust and accountability may justify deliberate friction. | Optimize the evidence display; do not remove confirmation. |
| Model routing reduced cost without a measured quality regression. | Economics can improve through governance, not only pricing. | Continue quality-gated routing and regression tests. |
| Async architecture passed 40 jobs but not 80. | Capacity is a tested boundary, not a permanent property. | Keep the growth cap and fund the next test only when demand requires it. |
| Professional demand is promising but cohort and channel evidence remain small. | Early revenue is not repeatable GTM proof. | Expand one controlled channel at a time. |
| Integration and collaboration requests persist. | Requests indicate jobs and workflow friction, not automatic build priority. | Advance discovery with quantified decision gates. |

# Residual Risks and Next Conditions

| Residual risk | Current evidence | Next condition |
|---|---|---|
| Retention may fall outside the concentrated beachhead. | 56.4% W24 in controlled cohort. | Repeat >=55% in another cohort and segment analysis. |
| Observed CAC is channel-specific. | 9.8-month payback in controlled acquisition. | Second repeatable channel with <=12-month payback. |
| 80-job performance remains weak. | 68s p95; 1.4% failures. | New architecture/capacity gate before higher ceiling. |

| Residual risk | Current evidence | Next condition |
|---|---|---|
| Provider concentration remains material. | Alternate procedure exists; one primary route dominates. | Provider game days and acceptable failover quality/cost. |
| Collaboration could create authorization and privacy risk. | Discovery demand only. | Validated jobs, willingness to pay, threat model and multi-user prototype. |
| Enterprise expectations exceed current scope. | No SSO/admin/compliance product. | Separate enterprise qualification; no current marketing claim. |

# Investment Options

| Option | Rationale | Assessment |
|---|---|---|
| Accelerate broad scale | Most delivery conditions passed. | Rejected: evidence remains controlled, capacity bounded, and GTM narrow. |
| Continue controlled scale | Preserves momentum while retaining evidence gates and scope discipline. | Selected. |
| Continue with corrective conditions only | Would reduce growth risk. | Too conservative because all primary delivery conditions pass. |
| Pause expansion | Would protect operations. | Not supported by current value, economics and reliability evidence. |
| Pivot to meeting summarization | Could broaden top-of-funnel demand. | Rejected: weakens differentiated durable decision value. |
| Stop | Avoids future AI and commercial risk. | Not supported by the simulated evidence. |

# Final Decision and Next Roadmap

**Final delivery decision -** CONTINUE CONTROLLED SCALE. The program closes the qualification conditions for the individual core workflow and demonstrates better customer value, AI accountability, economics, and 40-job operational capacity. It does not qualify broad horizontal, enterprise, collaboration, integration, multilingual, or richer-ingestion expansion.

| Next horizon | Priority | Decision gate |
|---|---|---|
| Next 90 days | Repeat retention and channel economics; improve provider resilience; preserve AI quality and support margin. | Another evidence gate with larger observed cohorts. |
| Discovery | Collaboration jobs, permission model, export/webhook value, and richer-ingestion risk. | No production build until quantified value and safety gates pass. |
| Operational scale | 80-job capacity only when observed demand approaches the current boundary. | p95 <=45s and failures <1% at the new approved ceiling. |
| Enterprise | Security, tenancy, SSO, audit, retention and procurement discovery. | Separate strategy and qualification package; not part of current product claim. |

- Keep AI assistance advisory and human finalization mandatory.
- Keep Starter and Professional packaging aligned to released capabilities.
- Publish metric definitions, operating boundaries, limitations, and case-study disclosure consistently.
- Treat delivery operations, AI evaluation, and benefits review as recurring systems rather than a one-time phase artifact.

*Evidence discipline: replace all simulated delivery values with observed, reproducible results before presenting them as actual product performance.*