



THE PROCESS SIX STAGES, TWO OF THEM GATES

01Intake Problem, sponsor, users, baseline	02Score Value, readiness, risk, scale	03Design Gate Architecture, controls, approvals	04Pilot Gate Limited release, logs, rollback	05Prove Value Outcome, adoption, quality, economics	06Decide Fund, condition, redesign, or stop
---	--	--	---	--	--

01 Intake scoring

CRITERION	WT. 5 = STRONG EVIDENCE
Strategic alignment	15 Advances an approved enterprise priority
Economic value	20 Quantified revenue, cost, capacity, or risk benefit
Process pain	10 Frequent, costly or high-friction task; clear users
Data readiness	15 Approved, accessible, current, representative data
Technical feasibility	10 Viable workflow, integrations and performance
Risk and control fit	10 Material risks understood and controllable
Adoption readiness	10 Owner, users, workflow change and training ready
Measure and scale	10 Baseline/target/owner defined; reusable pattern

SCORING Score each criterion 1 to 5; weighted points = weight x rating / 5. 75–100 pilot candidate · 60–74 discovery or remediate · <60 backlog or stop

AUTOMATIC HOLD No sponsor or baseline; prohibited use or data; or high risk that cannot be controlled. A strong score never overrides a failed gate.

02 Governance gates and required evidence

GATE	DECISION QUESTION	MINIMUM EVIDENCE PACK
G0 Intake	Is this a real, owned problem?	Sponsor and accountable owner; users and workflow; problem statement; current baseline; preliminary data classification; intake score.
G1 Design	Is it safe, lawful and feasible?	Risk tier; architecture and data flow; model and vendor rationale; privacy, legal and security reviews; human oversight; control plan; success and stop criteria.
G2 Validate	Does it perform under normal and hostile conditions?	Instructions and system prompt; data and knowledge inventory; access model; test set and results; quality and grounding thresholds; red-team tests; failure modes; version record.
G3 Pilot	Can a limited group use it safely?	Pilot charter and approved users; training and disclosures; monitoring and logs; incident, escalation and rollback plan; support owner; change control; release approvals.
G4 Scale	Should the organization fund production?	KPIs versus baseline; adoption, quality and risk results; control effectiveness; TCO, ROI and payback; production architecture and SLA; operating owner; funding request.

GATE RULE – NO EVIDENCE, NO PASSAGE Evidence must be stored, versioned and attributable to an owner and approver. Risk, legal, privacy or security may veto any score.

03 Custom GPT standards checklist

☐ OWNERSHIP AND LIFECYCLE Business owner, technical steward, risk tier, intended users, review date and retirement trigger.	☐ ACTIONS AND AUTHENTICATION Approved endpoints and schema; managed OAuth or key; scoped permissions; validation, timeout and error handling.	☐ PURPOSE AND TRANSPARENCY Name, description and starters match the approved use; limitations and human responsibility are explicit.	☐ CONSEQUENTIAL ACTS Read-only default; explicit user confirmation or human approval before writes, transactions, messages or changes.
☐ INSTRUCTIONS Role, allowed and prohibited tasks, source rules, uncertainty, refusal and escalation, approvals and output format are defined.	☐ SECURITY Prompt-injection, data-exfiltration and secret-disclosure tests passed; external content is treated as untrusted.	☐ KNOWLEDGE Files are approved, classified, current and owned; source and expiry recorded; no secrets or unnecessary personal data.	☐ QUALITY Representative and edge-case tests pass defined grounding, accuracy, completeness, safety, latency and cost thresholds.
☐ ACCESS AND SHARING Least privilege; workspace or invite-only by default; sharing level and settings permissions are approved.	☐ OPERATIONS Version history, change approval, monitoring and logs, incident response, rollback, support and quarterly review assigned.	☐ CAPABILITIES, APPS, ACTIONS Only required connections are enabled; purpose, owner, data exchanged and dependency are documented.	☐ EVIDENCE PACK Configuration record, knowledge inventory, test report, approvals, known limitations, release note and owner attestation.

04 Pilot measurement to funding recommendation

DIMENSION	WT. DEFAULT FUNDING EVIDENCE / TARGET
Business outcome	30 Approved KPI improves materially versus baseline (default 20% or more) and benefit is attributable to the pilot.
Adoption	15 60% or more of eligible users active; 50% or more repeat use; workflow completion and user acceptance meet target.
Quality	20 90% or more task success or accepted output, or approved task-specific threshold; rework under 10%; grounding target met.
Risk and reliability	20 Zero critical incidents; 95% or more of required controls executed; error, latency and availability thresholds met.
Economics and scale	15 Annualized net benefit at least 2x annual run cost; payback 12 months or less; production owner and capacity confirmed.

FUNDING RECOMMENDATION

Fund 80–100 All gates pass; no critical event; intake score 75 or above; named benefits owner.	Conditional 65–79 Mandatory gates pass; gaps have owner, budget and remediation under 90 days.	Stop / Redesign <65 Failed mandatory gate, material harm or data event, weak adoption, or no credible economics.
---	---	---

STANDARD

A GPT is not production-ready until every applicable item is evidenced. Public or link sharing, external actions and sensitive data require explicit exception approval.

PRE-REGISTER BEFORE PILOT

Baseline, target, data source, sample, duration, comparison method and metric owner. Do not invent success criteria after results are known.

ANNUALIZED NET BENEFIT

(validated time saved x loaded labor rate x annual eligible volume x adoption) + revenue or risk benefit – annual run cost