

Babylon 2k

Beth 2.0 - team scope

Revised 6 October 2026. Commercial figures omitted. Planning requirements, not production functionality already delivered.

Pilot: 381 planned days / 3,048 hours / 16-21 weeks. Broader launch, cumulative: 575 days / 4,600 hours / 26-32 weeks. Includes 20% reserve; one day is eight hours. Three core engineers and part-time supporting roles assumed.

Controlled production pilot

One Babylon 2k organisation; founder, specialist, manager and expert reviewer accounts with real permissions.

Three journeys: starting a straightforward locally owned business, routine bookkeeping/compliance, and assessing or clearing an ordinary records backlog.

Four packages assembled from approximately 12-15 validated service items; the remaining catalogue stays in draft.

English-first conversation, with basic Taglish handling tested on the pilot cases. Complex VAT, official notices, foreign ownership and regulated activities receive a specialist review route.

Approximately 8-12 onboarded specialists with verified service capabilities, coverage, capacity and agreed pricing responsibilities.

Live AI, approved source citations, conversational profiling, deterministic estimates, saved chats, Babylon 2k reports and secure document collection.

Configurable routing across three tested model task profiles from one approved AI provider; shared confirmed context, escalation, spending caps and tested fallbacks.

A small reviewed FAQ library plus question-time web research using approved BIR, CTA and Supreme Court source families. Bounded source caching, background checks, evidence validation and an outage/manual-review fallback; no complete tax corpus acquisition.

A real Beth case portal with saved cases, secure files and additional-work approval; staff-confirmed appointment requests; Chatwoot shared specialist inbox with one Telegram founder channel and email alerts. Specialists work in the shared inbox. Personal Telegram reply relays and autonomous multi-day orchestration are deferred.

Business-facing administration for interviews, source/fact review, date applicability, package and pricing rules, specialist data, publication, rollback and audit. Langfuse evaluations, monitoring, backups and measured launch tests.

Sizing assumption: up to 300 monthly active founders, 1,000 conversations per month and 20 simultaneous chat sessions; load tests must validate this envelope.

Langflow serves tested stage flows; LiteLLM owns model routing and budgets; PostgreSQL stores approved knowledge and typed relationships. Managed Langfuse is assumed, subject to acceptable data handling. No graph database, n8n, Appsmith or Kubernetes is required for this pilot.

Pilot cases are supervised, with saved status and staff-owned follow-up. This is not a promise of automatic recovery for a long case spanning multiple systems. If that becomes a pilot requirement, bring the Temporal extension forward and replan before work starts.

Broader launch, cumulative

Expand to the ten founder pain-point journeys, ten packages, 36 service cards and 11 specialist role definitions already proposed. All require approval before publication.

Add reviewed routing for VAT, payroll, inventory, multiple entities and foreign founders, including counsel escalation. Advice and estimates remain bounded by validated scope.

Add WhatsApp and LINE through Chatwoot, one connected calendar, and one documented Babylon directory/jobs adapter. This replaces the earlier WhatsApp-alert-only extension with richer channels.

Improve English/Filipino/Taglish evaluation, specialist capacity and reassignment, reminders, delivery analytics and comparison of quoted versus actual work.

Strengthen failure recovery, load and permission testing for a proposed envelope of 2,000 monthly active founders and 100 simultaneous sessions. Actual infrastructure sizing follows tests.

Extend live-source coverage and model/source regression tests for the remaining approved founder journeys; each new source family or model release must pass acceptance.

Temporal owns one standard durable case lifecycle: document wait, approval, reminder, fulfilment and recovery. Add explicit idempotency, activity receipts and worker compatibility tests; ordinary chat remains on the fast API path. Managed Temporal is assumed.

Dedicated graph engines, personal specialist Telegram relay, full self-hosted observability, Appsmith and n8n remain separately assessed options, not included merely because they are available.

Lean stack and ownership

Component	Stage	Responsibility
Beth web app + API	Pilot	One business admin surface; own identity, cases, quotes, knowledge and permissions. React/TypeScript and validated existing backend modules.
Langflow + native MCP client	Pilot	One bounded AI-stage runtime and approved tool connections. Editor in development; pinned flow/connector versions in production.
OOMOL OpenConnector	Pilot proof	Approved external-app actions and encrypted connections through MCP; six added base days. Shares app VM. No duplicate gateway; not installed in this demo.
IBM ContextForge	Replacement candidate	Use only if OpenConnector fails agreed compatibility/governance needs. Re-estimate the replacement; no parallel connector gateway.
LiteLLM	Pilot	One owner of model routing, call budgets and approved fallback; Beth tools enforce business decisions.
PostgreSQL, managed identity and object storage	Pilot	Separate platform database roles. Reviewed facts/typed relationships, source cache, private files and immutable release records.
Chatwoot + required Redis	Pilot	Shared specialist inbox, web API-inbox bridge and one Telegram founder channel. Staff use the inbox; no personal Telegram relay.
Langfuse + hosting health/log alerts	Pilot	Managed AI evaluation/tracing with redaction; operational alerts through hosting. Defer a full self-hosted monitoring stack.
Temporal	Broader launch	One standard multi-day case lifecycle, action receipts and recovery. Bring forward only if the pilot needs autonomous long-lived workflows.
WhatsApp, LINE, calendar and Babylon adapter	Broader launch	Add through the existing inbox/API boundary; prove identities, retries and receipts.
Neo4j / Graphiti / n8n / Appsmith / Kubernetes	Deferred	Only after a demonstrated unmet need and a separate estimate.
CopilotKit / AG-UI	Optional web adapter	Use if the compatibility proof fits the quote/profile UI. Choose one web transport; preserve Beth UX.

Reuse and narrower pilot automation reduce provisional planned effort about 14% / 13%. This is not measured productivity. Beth-specific rules, verification and permission code remain. No second reuse discount applies.

Licences and paid editions

Most selected engines have open-source cores; n8n is source-available and deferred, and LangSmith is commercial and not selected. Product-level check: 6 October 2026. Record exact releases/dependencies before implementation.

Engine / candidate	Licence position	Payment / conditions
Langflow	MIT - open source	Self-hosted engine has no licence fee; hosted services/support can cost.
OOMOL OpenConnector	Apache-2.0 - open source	Selected external-tool connector candidate for the scoped pilot proof. Self-hosted core has no licence subscription; upstream APIs, hosting and operations can cost. Not installed in the demo.
LiteLLM	MIT core - open source	Enterprise directory is separately licensed. SSO, audit and advanced governance may require payment; verify the selected feature set.
PostgreSQL	PostgreSQL Licence - open source	Engine has no licence fee. Managed database, backups, identity and storage are separately billed services.
Chatwoot	MIT community core - open source	Enterprise code is separate. Custom roles, branding, SSO and SLA features are paid in current plans; Beth API access checks do not protect direct inbox access.
Langfuse	MIT core - open source	Enterprise directories have separate terms. Managed hosting and enterprise features/support can cost.
Temporal	MIT server - open source	Self-hosting has operating costs; Temporal Cloud and support are paid services.
CopilotKit / AG-UI	MIT repositories - open source	Open web adapter/protocol. Optional managed products and underlying model calls may cost.
Redis	Depends on version/licence	Redis 8+ offers open-source AGPLv3 alongside source-available options. 7.4 is source-available; 7.2 and earlier BSD. Select an exact supported release and terms.
Valkey - candidate for cache/queue	BSD-3-Clause - open source	Prefer evaluating it for a permissive cache/queue option; prove Chatwoot and other clients against the chosen version before substitution.
IBM ContextForge - optional MCP gateway	Apache-2.0 - open source	Self-hosted registry/proxy; upstream tools and hosting may cost. Documentation currently shows release-candidate versions; prove maturity and exact compatibility.
n8n - deferred	Sustainable Use Licence - source-available	Do not label it OSI open source. Permitted internal use differs from embedding/reselling; some uses/features need commercial terms.
Appsmith - deferred	Apache-2.0 community core - open source	Advanced business features can be paid. Beth guided administration remains the baseline.
Neo4j / Graphiti - deferred	Neo4j Community GPLv3; Graphiti Apache-2.0	Community/library are open source; Neo4j Enterprise/cloud and Graphiti model calls can cost. No dedicated graph service in the pilot.
LangSmith - not selected	Commercial platform	Self-hosting requires an Enterprise add-on. Langflow does not require LangSmith; use Langfuse for the selected evaluation/tracing role.

Engine / candidate	Licence position	Payment / conditions
Kubernetes - deferred	Apache-2.0 - open source	Engine has no licence fee; cluster hosting and operations cost. Single-host containers are sufficient for the pilot architecture.

Open source does not mean the complete service is free. Model/search calls, messaging, hosting, storage, backup and staff remain operating expenses. Managed identity/storage providers are services, not necessarily open-source engines.

No promise that every future release or hosted plan will remain free. Retain pinned code/images and licence notices, record each edition and required feature, review upgrades and keep export/replacement paths. Rights for an already licensed release depend on its actual terms; future versions can have different terms.

Before choosing an edition, prove the required permissions, trace retention, branding and audit functions without a trial key. If community features cannot meet an access requirement, either approve a paid edition or change the architecture explicitly.

This is a product-edition inventory, not a completed dependency audit. Check every connector, container and transitive dependency at implementation. Default to community engines with no mandatory licence subscription; all paid exceptions require an explicit decision.

Open MCP connections

Selected proposed tool path: Beth authorises -> Langflow MCP client -> OOMOL OpenConnector -> external app. Not installed in the demo. LiteLLM owns model routing. ContextForge and LiteLLM MCP aggregation are replacement options.

Choice	Job	Scope boundary
Pilot: Langflow MCP client + OOMOL OpenConnector	Langflow owns AI steps; OpenConnector exposes approved external-app actions and connections. Beth API authorises each action.	Add 4 / 6 / 8 low/base/high person-days to the AI/integration workstream, with QA/platform role allocation. Prove up to two existing provider integrations; no new provider SDK from scratch.
Alternative: IBM ContextForge	Registry/proxy and REST-to-MCP adapters if OpenConnector cannot satisfy the agreed tool or governance requirements.	Replacement option, not an additional gateway. Previously scoped 5-8 base days; rebaseline the selected design after proof rather than adding both allowances.
Alternative: LiteLLM MCP gateway	Possible MCP aggregation or OpenAPI conversion if that is all the tool layer needs. Verify edition boundaries.	LiteLLM remains the model gateway. Its MCP layer is not enabled alongside OpenConnector for the same connections.

Model calls: Langflow -> LiteLLM -> selected model. Tool calls: Beth authorises -> Langflow MCP client -> OpenConnector -> external app. Keep one connector/credential owner; no ContextForge or parallel LiteLLM MCP gateway in this selected path.

MCP reduces repeated connector work where a suitable server exists; it does not make every provider free or remove provider authentication, permissions, API fees, field mapping and reliability tests.

Beth API remains the authority for identity, quotes, consent, booking and case actions. Begin with read-only tools; write tools use approved scopes, user confirmation where needed and duplicate-safe receipts. Tool results are evidence/input, never instructions overriding Beth policy.

Business managers can enable/disable an approved connector and inspect status through guided Beth controls. Engineers approve new servers, maintain credentials and validate upgrades. A connector gateway is not a replacement for Langfuse evaluation or Temporal durable cases.

Two provider integrations are a planning bound, not two integrations already delivered. A listed catalogue action is not proof it works. Check catalogue, local execution and real-provider verification separately; select one calendar and one document/storage provider during discovery.

Proof gate: Langflow compatibility with OpenConnector stateless POST MCP transport; per-founder scoped connection tokens, explicit connection allowlists, OAuth expiry/refresh, encrypted credentials, revocation, timeout and duplicate-safe action receipts. Never expose an unrestricted bootstrap/admin token to founders. Beth owns booking consent, identities and case access.

Engineers register approved actions; managers use guided Beth controls for connection status and approved enable/disable operations. If either chosen provider action fails the proof, retain manual fulfilment or re-estimate the replacement; do not declare every catalogue item production-ready.

OpenConnector adds six base days once in AI/integration, covering up to two existing providers with platform and QA allocation. Revised baseline: 381 planned pilot days / 575 cumulative broader-launch days. Alternative gateways are replacements, not stacked additions.

Pilot effort

Workstream	Low	Base	High	Result
Discovery and operating design	12	16	22	Journeys, responsibility, integration discovery, privacy/retention decisions and acceptance criteria
UX and interface design	9	12	16	Chat, mobile layouts, interviews, client/staff flows and accessibility
Accounts, data and permissions	17	22	30	Managed auth, shared schema, business memberships, staff roles, recovery and access policies
AI conversation, Langflow and LiteLLM	23	30	40	Stage flows, streaming adapter, memory and validated Beth tools; LiteLLM model policy, bounded retries, budgets and fallback. Langflow MCP client to OOMOL OpenConnector: up to two agreed existing providers; stateless POST transport, scoped per-founder tokens, explicit connection allowlists, encrypted credentials, OAuth expiry/refresh and duplicate-safe action receipts. Pin flow/model/prompt/connector versions. Six added base days include AI, platform and QA allocation once. No parallel connector gateway or catalogue-based discount.
Living knowledge, live evidence and cache	27	34	45	Three approved source families; relevant passage extraction, question-time verification, bounded cache and outage/manual review. PostgreSQL facts/typed links, effective and recorded dates, review states, immutable releases, retirement and cache invalidation; no complete tax corpus or separate graph engine.
Profiling and eligibility rules	9	12	16	Question selection, routing, explicit unknowns, risk and scope confirmation
Pricing and package engine	16	20	27	Work units, costs, bundle logic, rule versions, assumptions and replayable quotes
Directory and matching	7	9	12	Verified capabilities, fit reasons, availability and conflicts
Appointment requests	6	8	11	Staff confirmation, time-zone handling, cancellation and reminders
Beth client and specialist case portal	14	18	24	Saved case state, assignment, milestones, secure files and additional-work approval. Portal integrates with the Chatwoot conversation identity; staff supervise long case follow-up during the pilot.
Chatwoot inbox and human takeover	9	12	16	One Telegram founder channel, website API-inbox bridge, account linking, deduplication, explicit conversation ownership, AI pause and late-reply suppression. Staff use the shared inbox; personal Telegram relay is excluded.
Email delivery and operational alerts	4	5	7	Use Chatwoot notifications and one transactional email connector. Save delivery IDs/status, expose failed sends and use bounded retry/reconciliation. No new general-purpose workflow engine in the pilot.
Business admin and reviewed publication	19	24	32	Guided interviews; knowledge/change comparison; service units, exclusions, pricing drivers and specialist editor; author/reviewer/publisher rights; preview cases, publish/rollback and audit. All edits use Beth API checks. No direct flow/code/database editing for routine business changes.
Reports and history controls	6	8	11	Babylon 2k reports, recall/export and policy-aware deletion
QA, Langfuse and acceptance evaluation	19	24	32	Langfuse datasets/traces and deterministic quote checks. Expert-labelled held-out cases; role isolation, model/source/cache/outage tests, live handoff, concurrency, load and regressions. Managed observability does not remove Beth evaluation work.

Workstream	Low	Base	High	Result
Deployment, access separation and recovery	16	20	27	Container deployments on one consolidated application host; separate staging/production and platform DB roles, secrets, health alerts, log redaction, backups/restore and access reviews. Record edition/licence/dependency inventory, prove required community features, pin upgrades and test compatibility. Managed Langfuse, database and identity; no Kubernetes build. Protect commercial scope server-side. OpenConnector shares application-host headroom; separate encryption key and least-privilege database role.
Coordination, staff training and pilot rollout	12	15	20	Decisions, UAT, specialist onboarding, business-admin exercises, runbooks and release ownership. A manager must perform a normal rule correction and rollback without developer intervention.
Domain preparation and validation	22	28	37	CPA/expert work to validate pilot services, effort, sources and reference cases; focused privacy/legal review; FAQ risk/freshness policies, amendment/applicability review and outage escalation cases

Additional launch effort

Workstream	Low	Base	High	Result
Remaining journeys and catalogue	25	32	43	Reviewed services and rules for the ten journeys, ten packages and 36 service cards
WhatsApp and LINE channel expansion	9	12	16	Chatwoot connectors, business accounts, templates/consent, attachment and outbound tests, callbacks, identity linking and failure handling; external provider approval time is a calendar dependency.
One external calendar	9	12	16	OAuth, free/busy, event creation, reschedule/cancellation and reconciliation
Existing Babylon backend interface	14	18	24	One documented directory/jobs API adapter and agreed system-of-record mapping
Temporal case lifecycle and capacity	16	20	27	One durable document/approval/reminder case flow with bounded Langflow activities, receipts and duplicate prevention; workload, reassignment and recovery. No migration of every chat turn into Temporal.
Language and journey refinements	9	12	16	Reviewed English/Filipino/Taglish coverage and broader edge-case UX
Outcome and pricing analytics	11	14	19	Quote versus actual effort, rework, profitability and candidate policy adjustments
Additional QA and load work	17	22	30	Expanded holdouts, integration failures, recovery and the larger load envelope; expanded source coverage and model routing/fallback tests
Additional coordination and expert review	16	20	27	Launch preparation, content validation, staff training and integration review; approvals for expanded source/model coverage

Team and ongoing ownership

Role	Pilot	Full	Skills
Technical lead / backend engineer	54	80	TypeScript/Python API contracts, PostgreSQL, identity/permissions, pricing rules, knowledge versions and architecture decisions
Frontend / full-stack engineer	65	93	React/TypeScript, accessible chat, guided business admin, portal, preview/release UI and Chatwoot bridge
AI / integration engineer	66	94	Langflow, LiteLLM, evidence retrieval, typed tools, citations, Langfuse evaluation and model policy
QA / automation engineer	33	51	End-to-end, data isolation, source/quote regression, webhook/handoff, load and staff UAT
Platform / security engineer	23	37	Containers, managed services, Redis/PostgreSQL, CI/CD, observability, secrets, backups and recovery
UX / product designer	12	16	Founder journeys, readable forms, responsive interaction and business-admin usability
Product owner / delivery manager	24	40	Scope, source/service ownership, integration accounts, decisions, staff training and acceptance
Philippine CPA / tax and focused privacy experts	40	68	Approved service units, legal applicability, reviewed facts, example cases, pricing inputs and release review

Responsibility	Monthly allowance	Coverage
Technical maintenance and platform ownership	6-10 person-days/month	Pinned dependency upgrades, staging compatibility checks, gateway/source health, alerts, backups/restore, credential rotation and incidents. Managed hosting reduces infrastructure work; it does not remove an engineer. Includes OpenConnector token/connection isolation checks, OAuth refresh/revocation and upstream action regressions.
Expert knowledge and pricing review	6-10 person-days/month	Source changes, applicability, retired facts, service effort, quote variance, candidate releases and evaluation examples. Routine edits occur in the admin workspace.
Business admin and case coordination	3-5 person-days/month plus actual case workload	Staff input, specialist onboarding, availability, failed alerts and support queues. Does not include professional accounting delivery or 24-hour coverage.

Pilot: 15-25 person-days/month / 120-200 hours. Broader launch: 25-40 days / 200-320 hours, excluding accounting delivery and 24-hour coverage. Role totals and workstream totals are the same effort, not additive.

Hardware and hosting

Yes: multiple engines can share one server. A component is software; a container is its isolated running copy; the server supplies CPU, memory and disk to those containers. You can rent one cloud VM instead of buying many physical servers.

Deployment choice	Server allowance	What it means
Recommended pilot: one app server + managed services	1 application VM; start around 12 vCPU / 32 GiB RAM / 200 GiB SSD.	Beth web/API, Langflow, LiteLLM, Chatwoot app/workers, OpenConnector and Redis or validated Valkey share the VM as isolated containers. Managed PostgreSQL/identity, private object storage/backups and Langfuse stay outside. No dedicated physical machine per engine.
One larger self-hosted pilot server	1 VM; provisional 24 vCPU / 64 GiB RAM / 500 GiB SSD.	Adds local PostgreSQL/identity and Langfuse web/worker/ClickHouse/storage. Use separate users, databases, volumes and queue/cache policies. Still needs off-host backups and external AI APIs. Excludes local model hosting, Temporal and high availability. Extra deployment/recovery work: provisional 8-15 base days plus reserve, subject to provider/identity choice.
Broader launch or higher availability	Split or replicate only when load, uptime and recovery targets justify it.	Move busy inbox/AI workers or analytics first; add redundant application capacity and database availability as required. Temporal Cloud plus an activity worker is the baseline at wider launch; self-hosting Temporal needs a separate operational estimate.

These are engineering starting estimates for 300 active founders/month, 1,000 conversations/month and 20 simultaneous sessions, not tested capacity. A 12-core/32-GiB app host includes headroom over roughly 10 vCPU/20 GiB app allocations plus 1-2 GiB Redis. The managed database allocation is outside that host, so do not add it again.

Do not add the deployment option totals to the component rows: the rows explain how the same capacity is distributed. Choose ONE deployment option. OpenConnector adds a provisional 1 vCPU / 1-2 GiB allocation within the 12-vCPU/32-GiB host headroom; resize if the compatibility/load proof requires more.

One host has one failure point: if it stops, all local services stop. Off-host backups support recovery, not uninterrupted availability. Managed services reduce local maintenance but do not eliminate application downtime.

Staging is a separate environment, not necessarily a separate physical server. Prefer a small temporary second VM for release/load tests, with separate credentials and data; its allowance is outside the live-host figure. Never test with live customer records.

No GPU is required with external model APIs. Test CPU, RAM, database/queue latency, disk growth and recovery before admitting the agreed pilot load; resize or split services from measured results.

Component allocations within the chosen deployment

Layer	Starting allocation	Assumption
Beth API / web	2 vCPU / 4 GiB	Isolated container allocation; static assets through standard hosting/CDN.
Langflow runtime	2 vCPU / 4 GiB	Pilot single-instance assumption, editor separate. Replicas/shared events for the agreed availability and load target.

Layer	Starting allocation	Assumption
LiteLLM gateway	2 vCPU / 4 GiB	Planning allocation, not a vendor minimum; bound workers, streams and database connections.
Chatwoot app + workers	4 vCPU / 8 GiB initially	Current deployment guide recommends 4+ cores, 8 GB minimum and 16 GB recommended. Increase after workload tests.
OpenConnector	1 vCPU / 1-2 GiB initially	Provisional container allocation inside shared-host headroom; existing managed PostgreSQL, separate role/key. Prove transport, auth and load.
Managed PostgreSQL	2 vCPU / 4-8 GiB; 50 GiB initial storage	Separate platform databases/users, growth limits, point-in-time recovery and restore tests.
Redis	1-2 GiB initial memory	Separate non-evicting job pools from evicting caches. Test each platform against the selected version.
Private files / bounded source cache	100 GiB initial allowance	Retention/lifecycle caps, upload limits, signed access and malware checks; not a full tax-law archive.
Staging / development	Separate reduced instances; 2-4 vCPU / 4-8 GiB shared test capacity	Never use live accounts as fixtures. Temporary load workers may need more resources.
Langfuse	Managed baseline	Self-hosting adds web/worker, ClickHouse, PostgreSQL, Redis and object storage; separate estimate.
Temporal	Managed service + worker at broader launch	Worker sizing follows activities; self-hosted server/persistence adds operational work.
GPU	None	External model APIs perform inference; local model hosting/training changes the scope.

Pilot allocations are estimates, not a measured benchmark or high-availability design. No GPU is needed with external model APIs. Final sizing follows load/outage tests and provider selection.

Business administration and handover

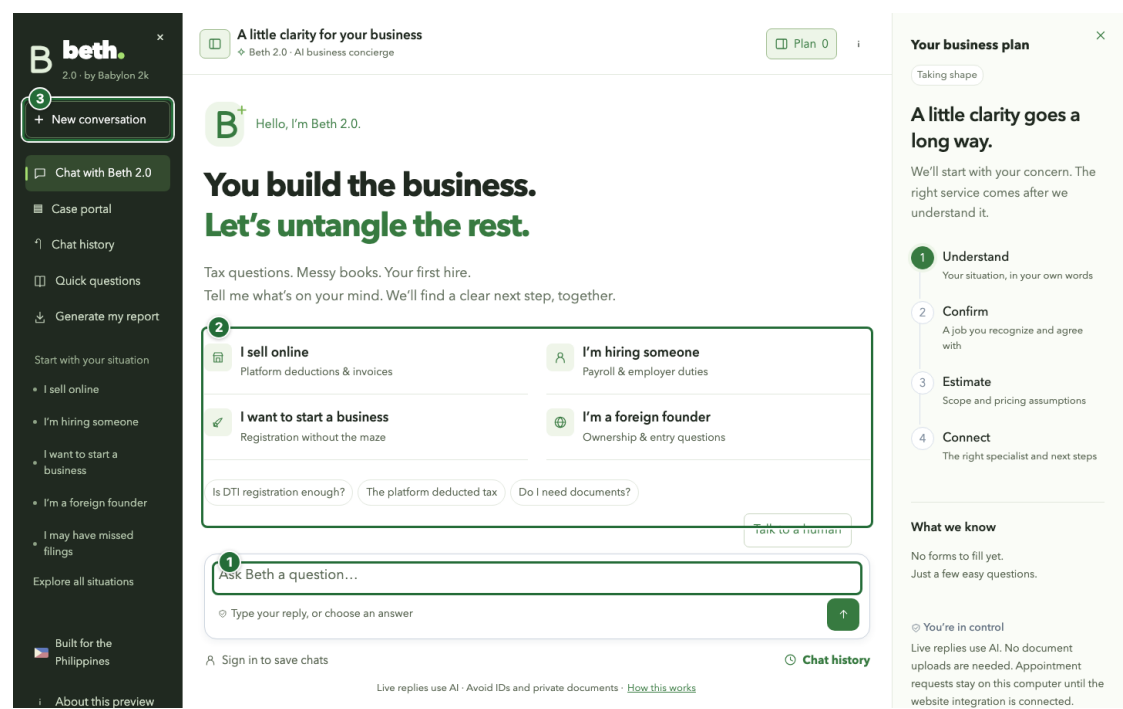
Role	Can do	Boundary
Specialist / respondent	Submit structured interview answers/evidence and permitted profile changes.	Cannot publish shared law or live pricing.
Knowledge editor	Draft sources, facts, applicability dates, amendment and retirement links.	Preview affected answers; reviewed publication required.
Expert reviewer	Approve evidence/interpretation and historical/current test examples.	Unresolved conflicts remain quarantined.
Operations manager	Maintain service units, exclusions, package/pricing drivers and specialist coverage.	Use an authorised approval and publish step.
Authorised publisher	Preview founder examples, publish and roll back approved releases.	Server records exact versions and reviewer/publisher identity.
Decision maker	Review commercial assumptions through separate protected access.	Founder/staff login does not grant commercial access.

Acceptance: a trained manager can edit, preview, publish and roll back a package without a programmer. An expert retires a fact while historical applicability continues to work. New connectors and upgrades still require engineers.

Delivery sequence

Phase	When	Exit
Discovery and compatibility proof	Weeks 1-3	Exit: contracts and operating owners accepted; re-estimate if the proof fails.
Shared foundation and business admin	Weeks 3-7	Exit: a manager can edit and publish an approved test rule with audited permissions.
Advice, profiling and reproducible quotes	Weeks 6-12	Exit: expert-reviewed held-out cases pass the agreed tests.
Cases, human inbox and recovery rehearsal	Weeks 10-16	Exit: one founder journey works end to end with a separate specialist account.
Supervised pilot and handover	Weeks 16-21	Exit: named staff can manage routine content, packages and corrections independently.
Broader launch	Weeks 26-32 cumulatively	Exit: approved expanded journeys and failure paths meet acceptance. Provider approvals may extend elapsed time.

Start with the founder's concern

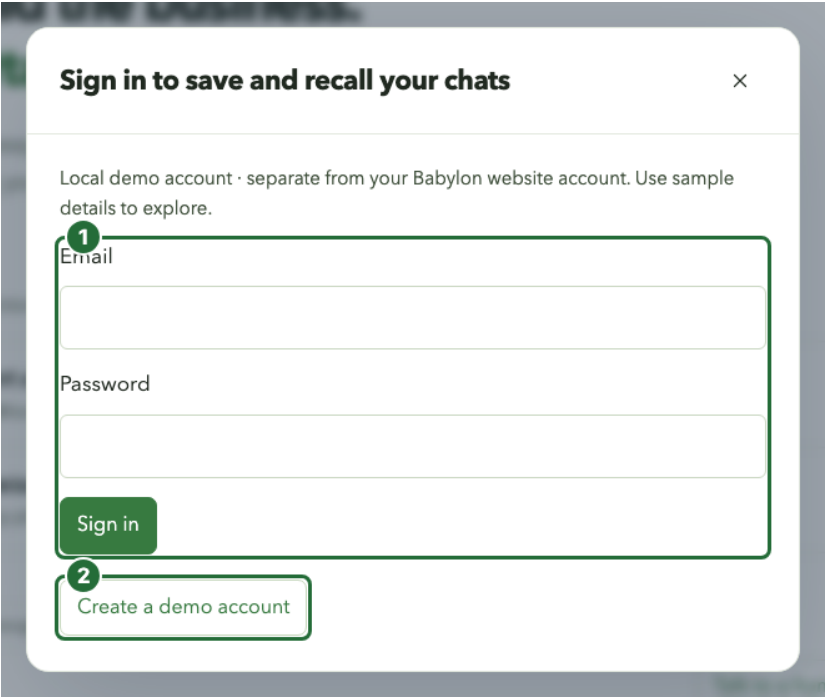


Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Ask Beth a question	Let a founder describe a problem in their own words and receive a useful next step. Maintain context across tested task models and validated AI tool calls. Configure model tiers, approved fallbacks and spend limits; stream and recover without unapproved tax downgrades. Evaluate Philippine founder questions and supported languages.	30 / 42
Choose a starting situation	Use a recognisable concern to start relevant questions, without asking for documents immediately. Define concern-to-question and urgency rules with experts. Record unknowns separately from confirmed facts. Handle conflicting answers and reasons to escalate.	12 / 12
Start a new conversation	Keep unrelated issues separate while retaining each conversation for later. Create independent conversation records and reliable identifiers. Prevent context leaking between conversations.	30 / 42

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Sign in and save the work

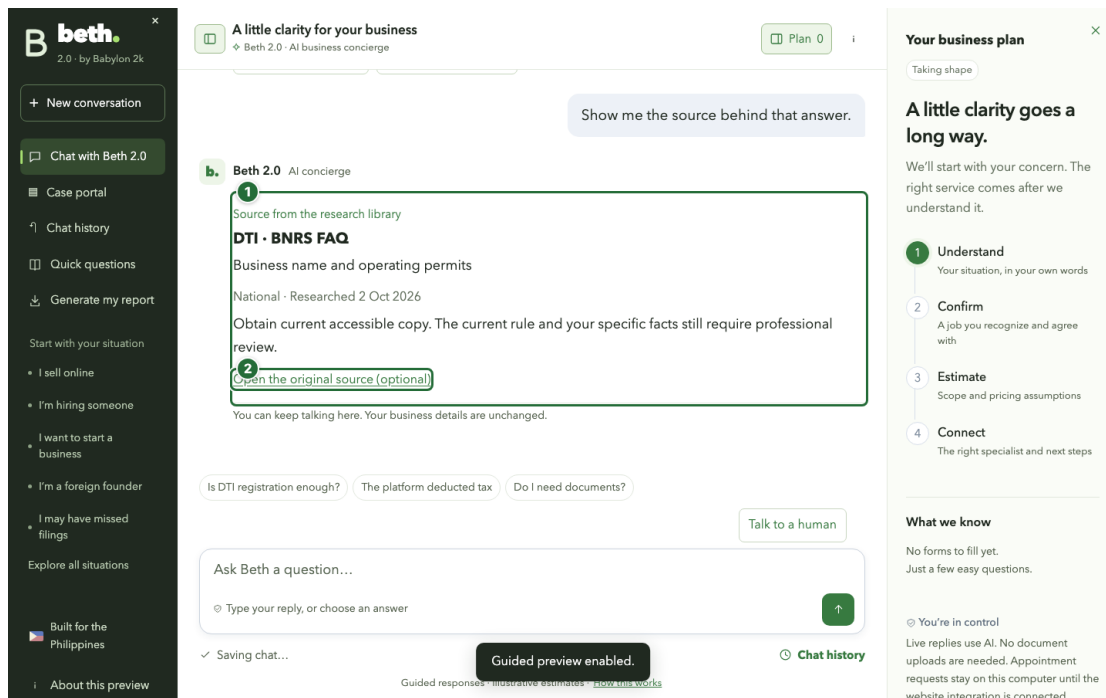


Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Sign in	Associate saved chats, interview answers and cases with the correct account. Implement managed sign-in, recovery and secure sessions. Define organisation membership and verified staff roles. Enforce account and case permissions server-side.	22 / 22
Create an account	Let a first-time founder register before saving their journey. Validate registration, verification and duplicate-account handling. Define consent and account lifecycle policies.	22 / 22

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Advice with a source the founder can inspect



Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Read the cited guidance	Explain an answer with supporting evidence and the limits of that evidence. Search and fetch relevant official passages on demand; retain dates, scope and version IDs. Use small reviewed FAQs and a bounded cache; quarantine changed guidance and provide outage/manual-review paths. Check citations against the claims actually made.	34 / 34
Open the original source	Let the founder verify the original reference rather than rely only on Beth's summary. Check original URLs, source-copy identity and amendment dates; disclose when only cached evidence is available. Escalate when evidence is missing, outdated or contradictory.	34 / 34

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Confirm what Beth understood

b. Beth 2.0 AI concierge

Here’s what I think you need.

Keep your books reconciled and your agreed filings supported by clear evidence.

1

Business structure

Sole proprietor

Monthly activity

Up to 100 entries

Business location

Cebu

Record quality

Organized and up to date

- Reconcile one business’s books each month
- Prepare the agreed filing schedule with CPA review
- Provide monthly financial reports and completion evidence

Does this describe the help you’re looking for?

2

Yes, that sounds right

Let me adjust a detail

Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Review the business facts	Turn a conversation into understandable facts: business type, location, workload and registration status. Define canonical facts, units and unknown values. Show corrections without silently overwriting confirmed information. Link each fact to a question, eligibility rule or price driver.	12 / 12
Confirm the proposed scope	Get the founder’s agreement that Beth has identified the right problem before preparing an estimate. Persist a confirmed profile snapshot. Require reconfirmation for material changes. Route unresolved facts to a specialist.	12 / 12

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Explain the package and indicative price

1

≡ Indicative range · not a firm quote

Everyday business care

₱10,000-₱14,500

per month

2

Monthly bookkeeping	₱6,339-8,761
Ledger and reconciliation; complex inventory excluded	
Income tax return assistance	₱1,024-1,618
Reviewed return and filing evidence if engaged	
VAT or percentage tax assistance	₱1,319-2,219
Reviewed return and submission evidence	
Withholding return assistance	₱1,318-1,902
Return and certificates within stated scope	

Babylon 2k pilot estimate Pilot 0.1 · 2026-10-02. Based on assumed delivery and review effort, not an approved rate card or a national average. Professional fee estimates only. Government fees, taxes, penalties, contributions, notarial/courier/travel costs, software and other third-party expenses are separate. Fee-tax treatment must be confirmed before a binding offer.

Always separate: professional fees, taxes, penalties, government charges and third-party costs. No completion date is promised before documents and dependencies are checked.

Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Indicative range and billing period	Show what the founder may pay, whether the work is one-off or recurring, and how certain the estimate is. Version approved costs, workload bands and complexity rules. Calculate fees deterministically with explicit unknowns and exclusions. Save the inputs and rule version so a quote can be replayed.	20 / 52
What the package includes	Translate service codes into clear line items and their purpose, without charging twice for shared work. Define service bundles, eligibility and dependency rules. Deduplicate common work by entity and period. Separate statutory charges, fees and out-of-scope work.	20 / 52

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Keep a Babylon 2k recommendation

1

BABYLON 2k

Business and accounting support

Business clarity report

Prepared by Beth 2.0 for Babylon 2k · Demo draft

I need help with my books

Business structure

Sole proprietor

Monthly activity

Up to 100 entries

Business location

Cebu

Record quality

Organized and up to date

₱10,000-₱14,500 per month

► Read the full report

2

Download Babylon 2k report

Print / save as PDF

Save chat

Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Babylon 2k report and scope	Give the founder a branded record of the situation, recommendation and estimate. Render confirmed facts and exact saved quotation versions. Include assumptions, source references and report identity. Apply sharing consent and access rules.	8 / 8
Download, print or save	Allow the founder to retain or share the recommendation without repeating the interview. Provide reliable export and saved report recall. Protect downloads and apply retention/deletion policy.	8 / 8

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Babylon 2k - team scope - 6 October 2026

19

Choose a suitable specialist

1

MA

Mary Ann Atillo Vaflor

4Yr Finance or Adequate Exp.
Cebu City, Cebu, Philippines

₱100/hour · published directory rate

Published Cebu location. Service specialty, active credentials, remote service and availability still need confirmation.

2

Select specialist & arrange consultation

[View published listing](#)

Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Compare suitability	Explain why a specialist fits the location, service and complexity of this business. Verify skills, credentials, coverage and seniority. Rank eligible specialists and record reasons and conflicts. Add capacity-aware ranking at the broader launch.	9 / 9
Select this specialist	Carry the selected specialist and business brief into booking or case assignment. Persist specialist selection and exact profile identifiers. Support rejection, reassignment and an unavailable-specialist fallback.	9 / 9

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Request a consultation time

Request a consultation

Mary Ann Atillo Vafior

[Open this specialist's exact profile](#)

Choose your preferred time. The specialist's availability and consultation fee need confirmation.

Preferred date and time · Philippines (UTC+8)

1

05/10/2026, 10:00 AM

Length

Format

30 minutes

Video call

What would you like to discuss?

Discuss monthly books, tax filing scope and the records to prepare. Sample founder walkthrough.

2

Saved under Guide Founder's local demo account. No message is sent to the specialist.

Save appointment request

Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Preferred date and time	Collect a convenient time and meeting format without suggesting the specialist has accepted it. Use explicit time zones, validation and staff confirmation. Handle cancellations, rescheduling and duplicate requests. Integrate one external calendar at the broader launch.	8 / 20
Save appointment request	Persist the request, brief and selected specialist so the founder can follow up. Save request status transitions and an auditable booking record. Trigger confirmation/reminder events through the alerts module.	8 / 20

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Ask for human help

Speak to a human specialist

A case is created immediately for the chosen specialist. If they accept the live request, they can reply here. Otherwise, your contact and case brief stay in their jobs queue.

Choose the specialist

Mary Ann Atillo Vaffor · Cebu City, Cebu, Philippines

Not live

1

Demo availability. If not live, your case is kept for follow-up.

Published directory candidates. Detailed fit, credentials and actual availability still need confirmation.

Guide Founder

beth-guide-1790954123722@example.test

Phone / WhatsApp for follow-up (optional)

Preferred follow-up

Case portal

2

I agree to share my volunteered contact details, conversation and case brief with this chosen specialist in the local demo.

Email, WhatsApp and Telegram notifications are simulated; no external messages are sent. This request accepts no paid engagement.

3

Create case & request human

Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Live / Not live indicator	Show whether a verified specialist is available and switch conversation ownership when they accept. Implement verified portal/Telegram identity and presence expiry. Offer human takeover and pause/resume AI replies. Handle delayed replies and disconnection.	12 / 12
Consent to share the brief	Let the founder decide whether the specialist receives the conversation summary and contact details. Record explicit consent and its scope/version. Share only the permitted brief with the assigned specialist.	8 / 8
Create a case and request a human	Keep a task with the chosen specialist even if nobody replies immediately. Create and assign a case with a consistent brief. Persist request status and a no-reply fallback. Enqueue notifications through the alerts module.	18 / 56

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Babylon 2k - team scope - 6 October 2026

22

Track the case and approve more work

B2K-2F0E9D70 · opened Oct 2, 2026, 11:32 PM

Client view

Keep your books in order

1

Open · specialist attending

Assigned to [Mary Ann Atillo Vafior](#) · live lane: Live specialist

Case file

Conversation

Work & approvals (1)

Notifications

Work, scope and approval

Each item covers a stated period. Additional work waits for client approval. Monthly, annual and one-time fees stay separate.

Monthly bookkeeping × 1

BK01 · November 2026 monthly books · MONTH
Ledger and reconciliation; complex inventory excluded · cap: One entity; two accounts; clean digital records
Pricing version: Pilot 0.1

₱7,000-₱9,500 professional-fee budget ·

Proposed

2

Approve this scope & budget

Decline

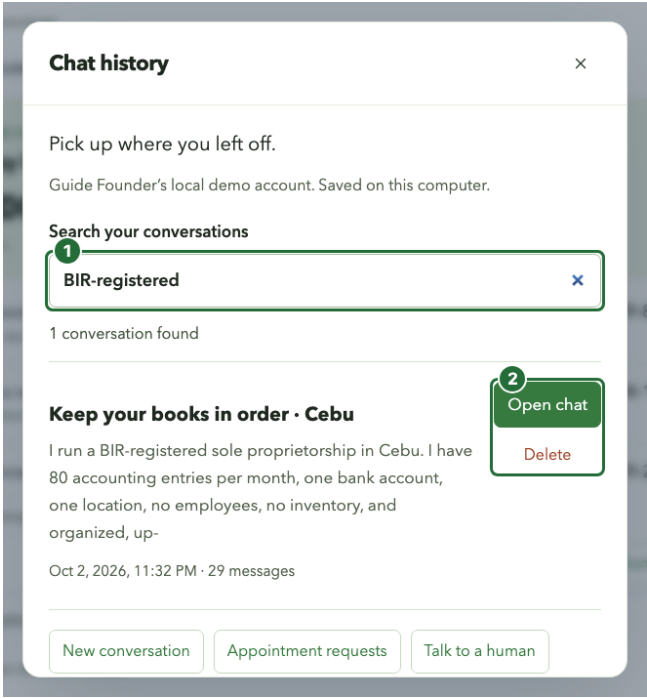
Review inclusions, quantity, period and fee before approving. Approval is simulated and triggers no payment.

Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Case status and owner	Show the founder who is responsible and what happens next. Implement authorised assignment and case status transitions. Secure documents, messages and milestones per case. Connect one documented Babylon backend at broader launch.	18 / 56
Approve scope and budget	Require the founder's approval before a specialist begins additional work. Version proposed scope and fees. Record approval by the authorised client. Prevent starting unapproved work and preserve change history.	18 / 56
Notification delivery status	Make staff and founder alerts traceable rather than assume a message was delivered. Queue email events with retries and deduplication. Record provider acknowledgements and visible failures. Add approved WhatsApp routes at the broader launch.	5 / 17

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Return to an earlier issue



Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Search previous conversations	Save time by finding an earlier business issue and its recommendation. Index only the account's authorised conversation summaries. Connect the saved conversation to report and quote snapshots.	8 / 8
Recall or delete	Resume context or remove the chat with a clear explanation of what is retained in an open case. Implement policy-aware recall and deletion. Distinguish a deleted conversation from required case/audit records.	8 / 8

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

Turn team interviews into reviewed inputs

[illegible]

Demo screenshot. Any founder service fees shown are illustrative professional service charges, separate from software build costs.

Feature	Function and remaining work	Pilot / full base days
Role interviews and service effort cards	Collect expert decisions, effort drivers and cost inputs that explain a founder question or quotation rule. Link each question to a named rule, knowledge item or test. Collect evidence and reconcile conflicting respondent answers. Support shared review, manager/expert sign-off and release history.	24 / 24
Package composition and draft comparison	Show which service items and cost inputs create each package before a manager approves them. Validate bundle composition and shared-work deductions. Compare proposed prices with approved costs and measured jobs. Keep released quotes reproducible.	20 / 52
Readiness and evidence check	Separate completing the survey from proving that Beth quotes and advises accurately. Build held-out founder cases and expected results. Compare quoted effort with actual completed work. Block releases that fail agreed evidence, accuracy or access checks.	24 / 60

Whole-workstream allocations are shared across controls; do not sum repeated allocations. Explore related areas in the interactive scope.

References and limits

OpenAI function calling

<https://developers.openai.com/api/docs/guides/function-calling>

Supports application-defined tools; server code remains responsible for executing and validating actions.

OpenAI structured model outputs

<https://developers.openai.com/api/docs/guides/structured-outputs>

Schema-constrained outputs help structure facts; they do not by themselves prove factual or professional correctness.

OpenAI API data controls

<https://developers.openai.com/api/docs/guides/your-data>

API content is not used for model training by default; retention and feature-specific storage still need review.

Supabase production checklist

<https://supabase.com/docs/guides/deployment/going-into-prod>

Production access policies, authentication, security, availability and backup responsibilities.

Supabase vector columns

<https://supabase.com/docs/guides/ai/vector-columns>

Postgres pgvector can store and query embeddings alongside business records.

Telegram Bot API

<https://core.telegram.org/bots/api>

Founder messaging channel through Chatwoot. Specialists use the shared inbox; personal specialist Telegram reply relay is deferred.

Google Calendar free/busy

<https://developers.google.com/workspace/calendar/api/v3/reference/freebusy/query>

Availability queries for the proposed connected-calendar booking flow.

Google Calendar event creation

<https://developers.google.com/workspace/calendar/api/v3/reference/events/insert>

Creates the consultation event after availability and confirmation checks.

Next.js deployment options

<https://nextjs.org/docs/app/getting-started/deploying>

Node or container deployment supports the proposed managed hosting approach.

OpenAI model selection

<https://developers.openai.com/api/docs/guides/model-selection>

Compare candidate models and reasoning settings on identical tasks; choose the lightest setting that meets the quality bar.

OpenAI models and providers

<https://developers.openai.com/api/docs/guides/agents/models>

Explicit per-agent/run model selection supports task-specific quality, latency and cost profiles. Exact versions are chosen after Beth benchmarks.

OpenAI web search

<https://developers.openai.com/api/docs/guides/tools-web-search>

Responses web search supports domain filtering and clickable source citations; Beth must independently validate original passages, dates and applicability.

BIR official archives

<https://www.bir.gov.ph/bir-archive>

Starting source inventory for regulations, circulars and orders; access, original documents and version coverage need validation.

CTA decision search

<https://cta.judiciary.gov.ph/decres2>

Official source family for decisions/resolutions. Publication alone does not prove finality or subsequent legal treatment.

NTRC court-decision archive

<https://www.ntrc.gov.ph/fees-and-charges/15-court-decisions>

Government alternative for selected tax decisions; verify exact document identity, period and subsequent treatment before use.

Langflow production resources

<https://docs.langflow.org/deployment-prod-best-practices>

Official guidance on external PostgreSQL, resource sizing and replicas. Beth allocations below are pilot planning recommendations, not measured capacity.

LiteLLM production deployment

<https://docs.litellm.ai/docs/proxy/deploy>

Gateway deployment and shared data stores; use bounded connections and scale based on actual traffic.

Chatwoot deployment

<https://www.chatwoot.com/deploy>

Current guide recommends 4+ CPU cores and 8 GB RAM minimum; 16 GB recommended. Confirm the pinned edition/version.

Langfuse self-hosting resources

<https://langfuse.com/self-hosting/configuration/scaling>

Self-hosting adds web/worker, PostgreSQL, Redis/Valkey, ClickHouse and object storage. Managed hosting is the baseline assumption here.

Temporal durable workflows

<https://docs.temporal.io/workflows>

Broader-launch durable case automation; activities still need duplicate-safe external actions.

Langflow - licence / edition reference

<https://github.com/langflow-ai/langflow/blob/main/LICENSE>

MIT - open source. Self-hosted engine has no licence fee; hosted services/support can cost.

OOMOL OpenConnector - licence / edition reference

<https://github.com/oomol-lab/open-connector/blob/main/LICENSE.txt>

Apache-2.0 - open source. Selected external-tool connector candidate for the scoped pilot proof. Self-hosted core has no licence subscription; upstream APIs, hosting and operations can cost. Not installed in the demo.

LiteLLM - licence / edition reference

<https://docs.litellm.ai/docs/enterprise>

MIT core - open source. Enterprise directory is separately licensed. SSO, audit and advanced governance may require payment; verify the selected feature set.

PostgreSQL - licence / edition reference

<https://www.postgresql.org/about/licence/>

PostgreSQL Licence - open source. Engine has no licence fee. Managed database, backups, identity and storage are separately billed services.

Chatwoot - licence / edition reference

<https://www.chatwoot.com/pricing/self-hosted-plans>

MIT community core - open source. Enterprise code is separate. Custom roles, branding, SSO and SLA features are paid in current plans; Beth API access checks do not protect direct inbox access.

Langfuse - licence / edition reference

<https://github.com/langfuse/langfuse/blob/main/LICENSE>

MIT core - open source. Enterprise directories have separate terms. Managed hosting and enterprise features/support can cost.

Temporal - licence / edition reference

<https://github.com/temporalio/temporal/blob/main/LICENSE>

MIT server - open source. Self-hosting has operating costs; Temporal Cloud and support are paid services.

CopilotKit / AG-UI - licence / edition reference

<https://github.com/CopilotKit/CopilotKit/blob/main/LICENSE>

MIT repositories - open source. Open web adapter/protocol. Optional managed products and underlying model calls may cost.

Redis - licence / edition reference

<https://redis.io/legal/licenses/>

Depends on version/licence. Redis 8+ offers open-source AGPLv3 alongside source-available options. 7.4 is source-available; 7.2 and earlier BSD. Select an exact supported release and terms.

Valkey - candidate for cache/queue - licence / edition reference

<https://valkey.io/>

BSD-3-Clause - open source. Prefer evaluating it for a permissive cache/queue option; prove Chatwoot and other clients against the chosen version before substitution.

IBM ContextForge - optional MCP gateway - licence / edition reference

<https://github.com/IBM/mcp-context-forge>

Apache-2.0 - open source. Self-hosted registry/proxy; upstream tools and hosting may cost. Documentation currently shows release-candidate versions; prove maturity and exact compatibility.

n8n - deferred - licence / edition reference

<https://github.com/n8n-io/n8n/blob/master/LICENSE.md>

Sustainable Use Licence - source-available. Do not label it OSI open source. Permitted internal use differs from embedding/reselling; some uses/features need commercial terms.

Appsmith - deferred - licence / edition reference

<https://github.com/appsmithorg/appsmith/blob/release/LICENSE>

Apache-2.0 community core - open source. Advanced business features can be paid. Beth guided administration remains the baseline.

Neo4j / Graphiti - deferred - licence / edition reference

<https://neo4j.com/legal-terms/>

Neo4j Community GPLv3; Graphiti Apache-2.0. Community/library are open source; Neo4j Enterprise/cloud and Graphiti model calls can cost. No dedicated graph service in the pilot.

LangSmith - not selected - licence / edition reference

<https://docs.langchain.com/langsmith/self-hosted>

Commercial platform. Self-hosting requires an Enterprise add-on. Langflow does not require LangSmith; use Langfuse for the selected evaluation/tracing role.

Kubernetes - deferred - licence / edition reference

<https://github.com/kubernetes/kubernetes/blob/master/LICENSE>

Apache-2.0 - open source. Engine has no licence fee; cluster hosting and operations cost. Single-host containers are sufficient for the pilot architecture.

Langflow native MCP client

<https://docs.langflow.org/mcp-client>

MCP Tools connections, approved tool selection and credential configuration; reuse the existing runtime.

LiteLLM MCP gateway

<https://docs.litellm.ai/docs/mcp>

Alternative connector path, not enabled beside OpenConnector for the same responsibilities.

OOMOL OpenConnector runtime API

<https://github.com/oomol-lab/open-connector/blob/main/docs/runtime-api.md>

Stateless POST MCP actions; verify Langflow client transport and explicit per-user connection allowlists.

OpenConnector credentials and OAuth

<https://github.com/oomol-lab/open-connector/blob/main/docs/credentials.md>

Prove credential encryption, refresh, revocation and scoped tokens before enabling real actions.

OpenConnector verification levels

<https://github.com/oomol-lab/open-connector/blob/main/docs/verification.md>

Listed actions, local executability and real-provider verification are distinct.

AG-UI licence

<https://github.com/ag-ui-protocol/ag-ui/blob/main/LICENSE>

MIT protocol implementation; no separate server required by the protocol itself.

Graphiti licence

<https://github.com/getzep/graphiti/blob/main/LICENSE>

Apache-2.0 library; deferred. Upstream database and model costs remain separate.

IBM ContextForge licence

<https://github.com/IBM/mcp-context-forge/blob/main/LICENSE>

Apache-2.0. Candidate optional gateway; validate pinned release and connector maturity.