

AI Configs Demos Brainstorm

Detailed AI-generated ideas. May want to scale back the scope and tech-stack. Work on 2-3 in 2025?

AI Configs as the source of truth: Model, prompt, tool scopes, safety thresholds, and rollout segments live in LD and are editable by authorized non-dev stakeholders.

Progressive delivery for AI: Canary releases, percentage rollouts, and environment-specific overrides for AI behaviors.

Online evaluation: Use LD Experimentation (or export metrics) to detect winner configs; regression tests via promptfoo/Phoenix before promotion.

Cost/latency governance: Route through OpenRouter or Portkey to enforce budgets; surface per-config cost in dashboards.

Auditable changes: Every config change has history and approvals, making LLM behavior traceable and defensible.

1) "Switchboard Support"

E-commerce AI agent with progressive delivery

Pitch: A storefront chat agent that can hot-swap models, prompts, tools, and retrieval strategies during a live pilot using AI Configs. We progressively roll out better variants to segments (loyalty tier, geography) while tracking cost and CSAT guardrails. This demo highlights "flip a switch, change the brain" without redeploying code.

- **Target audience:** VP/Head of CX & Support in retail/e-commerce; Platform & ML engineers.
- **Tech stack:** LaunchDarkly **AI Configs** + Experimentation; Node/TypeScript service; RAG over product catalog (Weaviate or pgvector); OpenRouter or Portkey gateway; Prompt/response eval via promptfoo; observability via Helicone or Langfuse; front-end widget in React.
- **Time to build:** ~1.5–2 weeks.
- **Downsides/risks:** Needs clean product data; guardrails for discounts/returns; measuring CSAT reliably takes traffic.
- **Why it's compelling:** Competes head-to-head with CX platforms by demonstrating **dynamic model/prompt rollout** as a product ops capability vs. static bot builders.
- **GTM:** Live webinar with a Shopify sample store; publish a public repo + Cloud demo; co-marketing with a payments partner.

2) FinServ "Policy Pro"

Compliant RAG copilot for regulated answers

Pitch: An internal copilot that answers policy and product questions with source citations and an always-updatable safety layer in AI Configs. We can instantly switch between high-accuracy and low-cost model profiles based on user role or time of day.

- **Target audience:** Product & Compliance leaders in banking/fintech; Risk/Legal; Platform teams.
- **Tech stack:** LaunchDarkly **AI Configs** (policy gates, model routing); Python FastAPI; RAG over approved corpora (Azure Blob + FAISS); Prompt eval with Phoenix (Arize) or promptfoo; audit logs via Langfuse; SSO (OIDC).
- **Time to build:** ~2 weeks.
- **Downsides/risks:** PII handling; retention/purge policies; evaluation datasets must be curated.
- **Why it's compelling:** Shows compliance toggles centralized in LaunchDarkly, so risk teams can approve/delay rollout without code changes.
- **GTM:** Joint webinar with a GRC vendor; write a "regulator-ready RAG" blog; targeted ABM to banks.

3) BrandVoice Studio

Controlled content generation for Marketing Ops

Pitch: A marketer-facing UI that lets non-technical users select a brand voice, audience, and campaign objective-backed by AI Configs versions and approvals. Run A/B variants by segment and promote winners with one click.

- **Target audience:** Marketing leaders, content ops, growth teams.
- **Tech stack:** LaunchDarkly **AI Configs** (prompt templates & versions); React app; serverless API (Vercel/Cloudflare); OpenRouter models; promptfoo for quality gates; analytics export to GA/StatSig-style metrics; CMS handoff.
- **Time to build:** ~1 week.
- **Downsides/risks:** Brand drift without strong guardrails; risk of off-brand claims.
- **Why it's compelling:** Elevates AI configs from dev tool to **marketing control surface** with auditability and rollback.
- **GTM:** Interactive landing page with downloadable “brand voice pack” starter; partner with a CMS vendor.

4) IT Helpdesk Copilot with Change-Control Safety Rails

Pitch: A Slack/Teams assistant that resolves tickets, runs automations (reset passwords, provision access), and gates risky actions via LaunchDarkly approvals. Ship new skills as feature-flagged “capabilities.”

- **Target audience:** CIO/ITSM leaders; Internal Tools teams.
- **Tech stack:** LaunchDarkly **AI Configs**; Python orchestration; Skills via workflow engine (n8n or custom); OpenRouter/Portkey; ticketing integrations (ServiceNow/Jira); eval via promptfoo; observability via Helicone.
- **Time to build:** ~2 weeks.
- **Downsides/risks:** Security scopes; brittle integrations; change windows.
- **Why it's compelling:** Demonstrates “agent skills as flags” with instant disable/rollback per environment and group.
- **GTM:** Conference talk + demo booth for ITSM events; tutorial series with sample playbooks.

5) Field Service Guide - voice + vision copilot for utilities/telecom

Pitch: A mobile assistant that understands pictures of equipment, proposes steps, and logs work orders. AI Configs control the allowed tools (vision, RAG, work-order writeback) per region and season.

- **Target audience:** Field ops leaders; COO; Digital transformation teams.
- **Tech stack:** LaunchDarkly **AI Configs**; React Native; vision-capable model via OpenRouter; on-device caching; RAG over manuals (Weaviate); eval with promptfoo; cost/latency tracing via Portkey.
- **Time to build:** ~2–3 weeks.
- **Downsides/risks:** Safety and liability; variable connectivity; model hallucinations on rare equipment.
- **Why it's compelling:** Moves beyond chat-shows **device features and tool access** governed centrally with LD.
- **GTM:** Pilot with a regional utility; case-study focusing on truck-roll time reduction.

6) Healthcare Payer Answers

Benefits & coverage explainer with guardrails

Pitch: Member services copilot that answers benefits questions with plan-specific policy, always citing sources and deferring when confidence is low. AI Configs let compliance freeze or hot-patch risky intents.

- **Target audience:** Health plan CX leaders; Compliance; Contact center ops.
- **Tech stack:** LaunchDarkly **AI Configs**; Python API; HIPAA-aware hosting; RAG over plan docs; Phoenix or promptfoo for hallucination checks; audit via Langfuse.
- **Time to build:** ~2–3 weeks.
- **Downsides/risks:** PHI handling; accuracy requirements; seasonal surge traffic.
- **Why it's compelling:** Puts compliance in the driver seat with an “approve-to-ship” configuration lifecycle.
- **GTM:** Co-sell with a BPO; thought-leadership on safe healthcare RAG.

7) Revenue Copilot

Sales email & sequence generator with territory toggles

Pitch: SDRs generate first-touch emails/sequences that respect region, industry, and product lines. AI Configs choose models and prompts per market to balance quality and cost.

- **Target audience:** CRO/RevOps; Sales enablement leaders.
- **Tech stack:** LaunchDarkly **AI Configs**; TypeScript API; CRM plug-ins (Salesforce/HubSpot); OpenRouter; promptfoo regression suite; analytics export for conversion tracking; red-team checks.
- **Time to build:** ~1 week.
- **Downsides/risks:** Spam/compliance; measuring causal impact requires traffic.
- **Why it's compelling:** Treats territory and ICP as **configuration**, not code, enabling rapid go-to-market iteration.
- **GTM:** “Cold-start to first meeting in a week” challenge campaign; partner with a data provider.

8) Claims Triage Summarizer

Insurance intake with risk scores

Pitch: A triage assistant that structures FNOL (first notice of loss), summarizes evidence, and flags potential fraud cues. AI Configs control thresholds and tool access (image analysis, external lookups) per line of business.

- **Target audience:** P&C claims leaders; SIU; Data platform teams.
- **Tech stack:** LaunchDarkly **AI Configs**; Python; RAG over policy language; OpenRouter + tool use; promptfoo for regression tests; Langfuse for traces.
- **Time to build:** ~2 weeks.
- **Downsides/risks:** Bias/fairness; explainability; integration with legacy claims systems.
- **Why it's compelling:** Makes governance visible-risk toggles and skill scopes become governed settings.
- **GTM:** Whitepaper on loss-adjuster productivity; pilot with a mid-market carrier.

9) HR Policy Copilot

Employee handbook Q&A with tiered confidence

Pitch: Employees ask sensitive questions (leave, benefits, travel policy). The copilot answers with links and confidence bands; low confidence routes to HR automatically. AI Configs govern answers by country and business unit.

- **Target audience:** CHRO/People Ops; IT for internal tools.
- **Tech stack:** LaunchDarkly **AI Configs**; Slack/Teams bot; RAG over HR docs; promptfoo suites; Portkey for cost/latency; SSO.
- **Time to build:** ~5–7 days.
- **Downsides/risks:** Policy drift; privacy; change-management.
- **Why it's compelling:** Demonstrates **policy-by-segment** and non-disruptive rollout (pilot to one BU, expand).
- **GTM:** Internal champion program; open-source starter with anonymized sample policies.

10) Travel Concierge

Airline/hospitality rebooking agent

Pitch: A traveler-facing agent that checks inventory, proposes alternates, and triggers vouchers-with all actions gated by AI Configs to prevent over-issuance. A/B model profiles to control cost during peak disruptions.

- **Target audience:** Airline/hospitality digital; Ops & loyalty teams.
- **Tech stack:** LaunchDarkly **AI Configs**; TypeScript orchestration; integrations to booking APIs; OpenRouter; promptfoo guardrails; Helicone for tracing.
- **Time to build:** ~2–3 weeks.
- **Downsides/risks:** Complex inventory rules; fraud/abuse; high stakes UX.
- **Why it's compelling:** Shows how **operational risk** can be managed with LD before enabling new agent actions.
- **GTM:** Joint webinar with a travel-tech ISV; crisis-scenario drill demo.

11) DocsChat

Developer docs assistant with eval-gated releases

Pitch: A developer helper that answers questions using your API docs and examples. Before release, every change must pass a regression suite and quality floor; AI Configs block promotion until evals pass.

- **Target audience:** DevRel; Platform PMs; Open-source maintainers.
- **Tech stack:** LaunchDarkly **AI Configs**; Next.js app; RAG over docs; promptfoo regression; Langfuse traces; OpenRouter.
- **Time to build:** ~1 week.
- **Downsides/risks:** Docs freshness; overlapping sources.
- **Why it's compelling:** Demonstrates a **CI/CD quality gate** for LLMs using configuration, not ad-hoc scripts.
- **GTM:** Embed on a public docs site; blog + “from 0 to DocsChat in 60 minutes” tutorial.

12) Fraud Alert Triage

Fintech case summarization + escalation policy

Pitch: The agent summarizes alerts, surfaces top risk factors, and recommends next steps, but escalation thresholds are runtime-configurable in AI Configs and can be changed instantly during incidents.

- **Target audience:** Fraud Ops; Risk leaders; Data teams.
- **Tech stack:** LaunchDarkly **AI Configs**; Python; OpenRouter; promptfoo red-teaming; Portkey cost controls; SIEM webhook.
- **Time to build:** ~10–12 days.
- **Downsides/risks:** Needs red-team & bias reviews; alert fatigue.
- **Why it's compelling:** Treats incident settings as toggles for real-time control without redeploys.
- **GTM:** Incident-response tabletop video; partnership with a fraud-intel feed.

13) Content Migration Assistant

Rewrite and classify legacy docs

Pitch: A bulk processor that rewrites tens of thousands of legacy articles into standardized style/reading level, with AI Configs managing model profiles for cost vs. quality and automatic backoff on edge cases.

- **Target audience:** Knowledge management; Support ops; Docs teams.
- **Tech stack:** LaunchDarkly **AI Configs**; Python workers; batch via queues; OpenRouter; promptfoo tests; object storage.
- **Time to build:** ~1–2 weeks for MVP.
- **Downsides/risks:** Quality variance; maintaining golden datasets; throughput costs.
- **Why it's compelling:** Makes content ops measurable and controllable (roll forward/back specific styles with versioned configs).
- **GTM:** Offer a “10k-doc clean-up challenge”; case study on SEO uplift.

14) Global Support Translator

Multilingual CX with per-locale configs

Pitch: A helpdesk translator/augmenter that localizes customer messages and agent replies, controlled per locale via AI Configs (e.g., high-accuracy models for Japanese, low-cost for Spanish).

- **Target audience:** Global CX leaders; BPOs; Enterprise support.
- **Tech stack:** LaunchDarkly **AI Configs**; TypeScript proxy; OpenRouter; promptfoo translation quality checks; Helicone for latency.
- **Time to build:** ~5–7 days.
- **Downsides/risks:** Tone in high-context languages; privacy with transcripts.
- **Why it's compelling:** Shows **locale-aware** configuration without code changes.
- **GTM:** Partner with a BPO; publish BLEU/COMET score benchmarks by language.

15) Eval Lab

Red-teaming & benchmarks wired to promotion gates

Pitch: A demo lab that runs automated evals (factuality, toxicity, jailbreak resistance) and only promotes a config to production when it meets thresholds. Side-by-side dashboards compare cost/latency/quality.

- **Target audience:** ML/MLOps; Platform engineering; Risk.
- **Tech stack:** LaunchDarkly **AI Configs** + Experimentation; promptfoo for automated eval; Arize Phoenix notebooks; Langfuse traces; OpenRouter/Portkey; CI integration.
- **Time to build:** ~1 week.
- **Downsides/risks:** False security without human review; maintaining eval sets.
- **Why it's compelling:** Turns **evaluation into a first-class release gate** visible to product and risk teams.
- **GTM:** Workshop + starter kit; security-focused blog series on jailbreak resilience.