

Checklist for your RAG-powered bot

This checklist is designed to catch the most common killers of launch success from subtle UX issues to silent cost explosions and compliance landmines.

Use it as the last gate before flipping the switch. Every item here has bitten real teams in the past, checking them ruthlessly.

Final Checklist

Business & Success Definition

Define chatbot primary goal (sales support / lead gen / customer support / onboarding / automation)

Identify key conversion points (pages & steps where chatbot can drive uplift)

Set measurable success metrics (conversion uplift, AOV, ticket deflection, CSAT, lead quality, cost reduction etc.)

Map chatbot vs human responsibilities + clear escalation boundaries

Document performance baselines (current stats) to prove future ROI

Data & Retrieval Quality

- | | |
|--|--------------------------|
| Data freshness, PII redaction, and validated chunking/overlap | <input type="checkbox"/> |
| Audit existing website content (accuracy, consistency, up-to-date) | <input type="checkbox"/> |
| Prepare objection-handling content (pricing, trust, comparison objections) | <input type="checkbox"/> |
| Tag content by intent (sales / support / onboarding / educational) | <input type="checkbox"/> |
| Retrieval quality: RAGAS faithfulness > 0.85 on 30+ queries | <input type="checkbox"/> |
| Hybrid + reranking beats baseline (A/B result) | <input type="checkbox"/> |
| Empty retrieval fallback smooth | <input type="checkbox"/> |

Humanization & Brand Voice

- | | |
|--|--------------------------|
| Persona prompt locked & consistent | <input type="checkbox"/> |
| A/B: humanized vs plain → +20% engagement/satisfaction | <input type="checkbox"/> |
| AI disclosure in first messages/footer | <input type="checkbox"/> |
| Tone + cultural fit for target country checked | <input type="checkbox"/> |
| Questions asked delicately & optionally | <input type="checkbox"/> |
| Add empathy patterns (recognize frustration, urgency, confusion) | <input type="checkbox"/> |

Performance & Reliability

- | | |
|---|--------------------------|
| Response time: < 4s (95% requests) test with real network | <input type="checkbox"/> |
| Rate limits & quotas monitored | <input type="checkbox"/> |
| Error handling in place: graceful fallbacks (API, Timeouts, Empty) | <input type="checkbox"/> |
| Session memory working (conversation history preserved across 5–10 turns) | <input type="checkbox"/> |
| Caching enabled for repeated queries → cost savings 30–70% | <input type="checkbox"/> |
| Verify stability with load testing under peak concurrency (e.g., 50+ users) | <input type="checkbox"/> |

Security & Compliance

- | | |
|--|--------------------------|
| Prompt injection / jailbreak attempts blocked (use guardrails or custom filters) | <input type="checkbox"/> |
| API keys stored securely (never hardcoded, use env + secret manager) | <input type="checkbox"/> |
| Logging anonymized: no user messages stored with PII | <input type="checkbox"/> |
| Legal disclaimers in place for high-stakes niches (fintech: “not financial advice”, healthcare: “not a diagnosis”) | <input type="checkbox"/> |
| Data residency/location compliant (e.g., EU servers if needed) | <input type="checkbox"/> |
| Scan final output for toxicity/bias before sending to user (Output Moderation) | <input type="checkbox"/> |

Multilingual & Localization

English\multi language support in prompts & responses

Mixed-language queries handled correctly

Cultural/local meme fit checked

Edge Cases & Adversarial Testing

10+ adversarial/jailbreak attempts blocked

Toxic/offensive input → safe response

Very long/short, typo-heavy, slang queries tested

Cost & Monitoring

Cost projection: calculated for AI, LLM calls & integrations.

Budget alerts: notifications set (e.g., >\$500 to Slack).

Monitoring stack: LangSmith/Arize (tracing, hallucinations, latency).

User feedback: thumbs up/down + comment field.

Optimization loop: regular iteration based on real usage data.

UX & Conversion

- | | |
|---|--------------------------|
| Responsive UI: verified on Mobile/Desktop (emojis, carousels, markdown). | <input type="checkbox"/> |
| Onboarding / Starter Chips: clickable suggestions to guide initial user intent. | <input type="checkbox"/> |
| Clear CTAs: distinct action paths ("Add to cart", "Checkout") visible. | <input type="checkbox"/> |
| E2E Testing: 3–5 full journeys (Discovery → Purchase) validated. | <input type="checkbox"/> |
| Human Off-ramp: seamless escalation button if the bot gets stuck. | <input type="checkbox"/> |
| Rate Limiting: safeguards enabled to prevent abuse & billing spikes. | <input type="checkbox"/> |

Launch Readiness

- | | |
|--|--------------------------|
| Staging Parity: matches production (same models, index, prompts). | <input type="checkbox"/> |
| Post-deploy Smoke Test: verify critical flows (e.g., top 5 queries) immediately. | <input type="checkbox"/> |
| Soft Launch: 5–10% traffic for first 24–48h with close monitoring. | <input type="checkbox"/> |
| Rollback Protocol: ability to revert to previous version in < 5 mins. | <input type="checkbox"/> |
| On-Call Schedule: team assigned for immediate support (Week 1). | <input type="checkbox"/> |
| Rate Limiting: safeguards enabled to prevent abuse & billing spikes. | <input type="checkbox"/> |

Quick Launch Score – How ready is your RAG bot?

Count how many items you've already checked (✓):

Readiness Level	Checked	Percentage	Recommendation
Excellent – green light	43–49	88–100%	Soft launch (5–10% traffic, 48–72 hours), full monitoring, on-call rotation. Very high probability of smooth & successful rollout.
Very strong – almost ready	37–42	76–86%	You're close to green. Remaining gaps are usually fixable in 3–10 days.
Acceptable foundation only	31–36	63–73%	Missing ~13–18 items → real chance of serious post-launch problems (compliance incidents, hallucinations, poor conversion, latency spikes, bad user experience).
Not production-ready	≤30	≤61%	Too many critical gaps. High probability of failure, poor ROI, compliance/legal issues or reputational damage.

If implementing your RAG bot feels overwhelming or you want to skip the trial-and-error phase just drop us a message.

Our clients have seen 20–45% lifts in conversion rates, average order value, and repeat purchases after launching humanized RAG assistants.

We'll make sure yours becomes one of them!