Langkraft AI

FREE CHECKLIST

The production AI checklist: 12 things to fix before you launch

Most AI pilots work in the demo and fail with real users, real costs and real data. Use this list to check an AI agent, assistant or LLM app before it goes live. Each point comes from systems we have built and run.

- [ ] **01 Test against real cases before every release**

  Keep a set of real inputs with hand-checked answers, run it before every change, and block the release if it fails. A demo that worked once proves nothing.

  Seen in: Front desk agents

- [ ] **02 Flag what the AI isn't sure of**

  Give every extracted field or answer a confidence check. Uncertain results are highlighted for a person instead of being guessed and passed on.

  Seen in: Front desk agents

- [ ] **03 Answer only from your own sources**

  Ground answers in your documents or records, and have the AI say so when a detail isn't there. Prices, times and numbers must appear in the source.

  Seen in: Product discovery · Front desk agents

- [ ] **04 Keep a person in the loop for anything risky**

  Money, cancellations, refunds, complaints and anything clinical or legal go to a person. These rules live in code, not in a prompt the model can ignore.

  Seen in: Live demo · Front desk agents

- [ ] **05 Check every AI output in code**

  Validate format, fields and values before anything touches your systems. If the output doesn't pass, it is retried or sent to a person, never used as is.

  Seen in: Live demo

- [ ] **06 Plan for a provider outage**

  Route calls through a fallback chain across providers, with retries and circuit breakers. If every provider fails, the work goes to a person rather than stopping.

  Seen in: Live demo

- [ ] **07 Put a ceiling on cost**

  Set limits per query, per user and per day, and show a plan before anything expensive runs. A runaway loop or a bad question should never become a surprise bill.

  Seen in: BigQuery analytics

- [ ] **08 Never act twice**

  Make every action idempotent: a retry, a crash or a double click can't send the same email, create the same record or book the same slot twice.

  Seen in: Vendor negotiations · Front desk agents

- [ ] **09 Test without touching real customers**

  Use test modes that redirect outgoing emails, dry runs that skip paid calls, and plan-only runs. You should be able to try everything safely on real data.

  Seen in: Vendor negotiations · BigQuery analytics

- [ ] **10 Keep sensitive data away from the model**

  Mask customer figures, card and bank numbers before a model sees them, and keep personal data out of logs. Only send data to providers covered by an agreement.

  Seen in: BigQuery analytics · Front desk agents

- [ ] **11 Limit each user to their own data**

  Every request carries who is asking, and access is enforced by your data layer, not by asking the model to behave.

  Seen in: BigQuery analytics · Power BI hub

- [ ] **12 Log every step with a trace ID**

  Record each request, model call, tool call and decision under one trace ID, so you can explain any answer and fix problems fast.

  Seen in: Vendor negotiations · Sales assistant

Ticked fewer than 9? Let's look at it together.

An AI System Review tests your system on real cases in one to two weeks and gives you a ranked fix list with effort and cost.
Book a free 30-minute call: amanpreet@langkraft.ai · langkraft.ai

© 2026 Langkraft AI Private Limited · Delhi, India · Privacy