@demotoprod · THE LAB · EPISODE 1 · GATES

Three gates for an AI email assistant

We ran one AI inbox assistant 180 times against a mock inbox: three emails, three builds, twenty runs each, every run logged. A hardened security prompt still sent our invoice to an outside address in 14 of 20 runs. These three gates, in code, held in every run.

1 An allowlist on anything that leaves

The model reads your instructions and a stranger's email as the same kind of text. Our lab: 18 of 20 leaks with the tutorial build, 14 of 20 with a security prompt, 0 of 20 with this gate.

- Keep a list of approved outside contacts. Anything else waits for a person.
- Enforce it inside the send, forward and reply tools, not in the prompt.
- Log every held message. The attempts are your early warning.

```
def send(to, body):
    if to not in APPROVED_CONTACTS:
        return hold_for_approval(to, body)
    return mail.send(to, body)
```

2 A human on anything that commits money

Asked for a written refund, the tutorial build said yes in 7 of 20 runs. With "never promise refunds" in the prompt, 1 of 20. Drafts only: 0 of 20 sent.

- Replies to outside addresses are saved as drafts. A person sends them.
- Auto-send only fixed templates, such as an acknowledgement. Never free text to a customer.
- Read the drafts queue. It shows you what the assistant would have promised.

```
def reply(email, body):
    if not is_internal(email.sender):
        return save_draft_for_review(email, body)
    return mail.send(email.sender, body)
```

3 A filter at the door for anything automatic

Auto-replies answering auto-replies is how mail loops start. Our assistant replied to an out-of-office in 1 of 20 runs. The filter costs nothing.

- Drop messages marked Auto-Submitted (anything but "no"), X-Autoreply or Precedence bulk/list before the model sees them (RFC 3834).
- Mark your own automated mail Auto-Submitted auto-generated.
- Cap the replies the assistant can send in one thread.

```
def on_new_email(msg):
    if msg.headers.get("Auto-Submitted", "no") != "no":
        return  # RFC 3834: never answer a robot
    agent.handle(msg)
```


Nitish Gautam · indica-tech.com · hello@indica-tech.com

INDICA TECH · @DEMOTOPROD

Nitish Gautam

Founder-led AI engineering. Demo to production, under NDA, in regulated industries, on fixed scope.

Where this list comes from

Ten-plus years taking products from zero to one across seven sectors, including defence, fintech, medical and telecom, with AI shipped into regulated and air-gapped environments. The @demotoprod channel takes one production AI failure apart each day. These checklists are the same work written down: every item is a dated failure or a documented default, with the control that stops it and the primary source.

If this list found a gap

1. **Your demo works and production is stuck.**
   Production Readiness Audit. Fixed price, two weeks. A five-dimension scorecard, cost and failure analysis, and a 90-day execution roadmap.

2. **You need senior AI judgement without a full-time hire.**
   Fractional AI Lead. Monthly, 30-day notice either side. Weekly 1:1 with the founder or CTO, architecture and code review, technical hiring support, investor-ready due diligence.

3. **An AI feature has to work, by a date.**
   AI Build Sprint. Fixed price, four to eight weeks. Scope signed off up front, weekly demos, CI/CD and observability included, 30-day production warranty.

4. **A regulator, an auditor or a customer's security team is asking.**
   AI Governance & Compliance. Fixed price, four weeks. ISO 27001 / SOC 2-aligned controls, a model risk framework, human-in-the-loop and audit logging, regulator-ready documentation.

BOOK A CALL
indica-tech.com/#contact

EMAIL
hello@indica-tech.com

WEBSITE
indica-tech.com

ALL CHECKLISTS
indica-tech.com/checklists/

YOUTUBE
youtube.com/@demotoprod

LINKEDIN
linkedin.com/in/startup-founderx

INSTAGRAM
instagram.com/demotoprod

TIKTOK
tiktok.com/@demotoprod

Free to share. The live list, with every source link: indica-tech.com/checklists/#ai-email-assistant