

Answer Key — Tabletop Exercise

Reference solution · compare to your own plan before reading

Below is a first-hour response plan for the scenario "engineer pasted a production DB connection string into personal ChatGPT." It is written the way you would actually write it in a pinned Slack message to yourself at 15:42 on a Tuesday: terse, ordered, and explicit about timing. Your plan does not have to match this line for line — it has to match the shape.

First-hour timeline

- **T+0 (15:42):** Acknowledge to engineer; instruct them to screenshot the prompt + response, note the exact time, and NOT delete the chat. Tell them to stop using the tool.
- **T+2:** Pull the DB connection string details — which DB, which credentials, which network exposure (internal-only / public).
- **T+5:** Rotate the DB password immediately via admin console or IAM. Revoke the old credential.
- **T+8:** Check DB audit log for any connections using the old credential after the paste timestamp. If any from unfamiliar IPs, escalate to incident.
- **T+12:** Rotate any other credentials the engineer pasted in the same chat session (scroll up — the blast radius is the whole session, not just the one message).
- **T+15:** Audit for any service using the old credential; update deployments, rotate downstream.
- **T+20:** Notify security lead and engineering manager.
- **T+25:** If PII on that DB was accessible with that credential → notify legal/privacy.
- **T+30:** Draft a timeline doc, assign postmortem owner.
- **T+45:** If required by contract/law/policy → notify affected customers.
- **T+60:** Submit data deletion request to the AI provider (knowing it's not a guarantee).

The most common mistake on this scenario is rotating the DB password at T+5 and stopping there. The whole chat session is the blast radius. If the engineer pasted other context into the same conversation — schema snippets with table names, a second credential from the same `.env` file, the migration script with embedded secrets — all of it was exposed to the same provider on the same account at the same time. Rotate all of it.

Questions for the engineer

Send these immediately after the acknowledgement at T+0. You need the answers before most of the rotation decisions can be made correctly.

- **Exact prompt text** — copy-paste the full message you sent, verbatim, into secure notes (not into this DM). Include every character, including what you think is "just the query."
- **Exact response text** — the model's reply, also verbatim, in secure notes. Some models echo the secret back in formatted output, which is a second copy.
- **Time** — the precise timestamp you sent the message, to the minute if possible, and your time zone.
- **Is the chat still open?** — we need it open for evidence. Do not close the tab, do not start a new chat.
- **Personal or corp account?** — this changes the paper trail, the contract posture, and the notification path.
- **Any other secrets in the same chat session?** — scroll up through the whole conversation and list anything that looks like a credential, key, path, hostname, or identifier.
- **Was chat history enabled on the account?** — if yes, the data is retained for training and review by default on most consumer plans, and may also be synced to other devices logged in as the same user.
- **Which database** was it? Name, environment (prod/stage), and roughly what data is on it.
- **Was the DB reachable from the public internet**, or only from inside the VPC?

Instructions to engineer

Send this as a short numbered list. It is the engineer's half of the response, in parallel with your rotation work. Keep it short enough to follow under stress.

- **Stop using the tool.** No more prompts, no follow-ups.
- **Keep the chat open** for evidence. Do not delete it. Do not close the tab. Do not start a new chat.
- **Take screenshots** of the prompt, the response, the account indicator, and the timestamp. Save them locally first.
- **Do not upload the screenshots to personal cloud storage** (iCloud, Google Photos, Dropbox). They contain the same secret you just leaked.
- **Send the screenshots only via the corp secure channel** — the incident DM we are in now, or the corp ticketing system. Not SMS, not personal email.
- **Do not discuss on public Slack channels.** Keep it in this DM until told otherwise.
- **Scroll up** through the full chat session and list every other secret, identifier, or path that appeared. Send the list to me in the next message.
- **Stay reachable** for the next ninety minutes.

Notification list

- **Security lead** — immediate (by T+20). "Confirmed credential exposure to consumer AI provider, rotation in progress, no evidence of misuse yet, will update in 15 minutes."
- **Engineering manager** (of the affected engineer / owning system) — immediate (by T+20). Same summary; they need to know before anybody escalates over their head.
- **DBA team / data platform team** — immediate for rotation coordination; they own the downstream services that will break when the old credential is revoked, and they are the fastest way to check the DB audit log.
- **Legal / privacy** — if PII was reachable with the leaked credential. Within the hour. They start the regulatory clock on their side.
- **Affected customers** — if your contract, your regulatory regime, or your internal policy requires it. Within the SLA stated in the contract, not later.

- **AI provider data deletion request** — advisory, by T+60. File the form, keep the reference number, but do not let any other remediation wait on it.

If your plan did most of this — evidence before rotation, whole-session blast radius, notification with a clock on it, postmortem owner assigned inside the first hour, provider deletion filed but not trusted — you did the exercise correctly. If it rotated only the DB password and stopped, or if it asked the engineer to delete the chat, re-read the lesson: those are the two mistakes the playbook exists to prevent.