

Lesson 05 — Identity Depersonalization

How an AI account accumulates a fingerprint of you, and how to design a persona that does not leak back to your real identity.

Audience: Devs + Operators

AI OpSec

Personas

Notes: Slightly more technical lesson. Covers stylometry, tool-chain telemetry, and persona discipline.

What you will learn

- **The fingerprints an AI account leaves** — account metadata, device, network, behavioral, and linguistic
 - **How to design a clean persona** — device, browser profile, payment trail, voice, cadence
 - **Cross-contamination avoidance** — how personas get linked and what to do when it happens
-

What an AI account knows about you

Every AI account quietly collects far more than the text you typed. Assume the following are all tied to your account server-side:

- **Billing info** — card number, billing name, billing address, BIN country
 - **Linked OAuth** — every "log in with X" you ever clicked
 - **IP addresses** — every network you signed in from, with timestamps
 - **Device fingerprint** — user agent, fonts, screen size, GPU, timezone
 - **Account age, chat history, writing style, typing cadence**
-

Writing-style fingerprinting

Two prompts with identical meaning but different phrasing:

Prompt A

"hey can you stub this out real quick? just the skeleton, no impl, ta"

Prompt B

"Please generate a function signature and docstring only, without an implementation body."

Same request. Completely different authors — and an LLM or an analyst can often re-identify the author across accounts from a handful of examples.

Tool-chain fingerprinting

The *set* of tools you use is itself a signature — even if no single tool is unusual.

- Claude + Cursor + fish + iTerm2 + 1Password + Raycast + macOS 15.1 on Apple Silicon
- Your IDE extensions, your shell aliases, your keyboard layout, your preferred model
- Crash reports, auto-update pings, analytics beacons — all carry the same fingerprint

A persona that uses "your" tools is not a persona. It is you with a different name.

Linking through OAuth

One click, two worlds joined

"Log in with GitHub" on a persona account links that persona to every other account on both sides of the GitHub identity. The provider now has a durable, provider-attested link between the two.

- OAuth is a **correlation primitive**, not a convenience feature
 - Every scope you grant is a new joint the investigator can pivot on
 - There is no "just for login" — the link is permanent on both sides
-

Persona design principles

- **Separate device or VM** — dedicated hardware or a clean VM that never touches personal accounts
 - **Fresh browser profile** — no extensions, no history, no saved logins from the real identity
 - **Fresh payment trail** — prepaid card, virtual card, or other non-personal funding source
 - **Fresh phone number** — for SMS verification, never your real number
 - **Consistent but distinct voice** — the persona writes the same way every time, and differently from you
 - **No cross-reference with real identity** — no shared names, no shared idioms, no shared timezone chatter
-

The "never two identities in one profile" rule

One browser profile per persona, period.

Not one profile for two "similar" personas. Not one profile "just for login." Not one profile "because the persona is low-risk." One persona, one profile, always.

- Browser profiles are the cheapest isolation primitive you have
 - Mixing two personas in one profile is indistinguishable from having one person with two names
-

Example — the burned persona

A red-teamer maintained a clean persona for an op-related Claude account: dedicated VM, fresh email, prepaid card, the works.

- On their personal ChatGPT they often wrote *"let's stub this out"* when asking for skeletons
- On the op-related Claude they used the exact same idiom, three times in the first week
- Content analysis correlated the two accounts by prose features
- The persona was linked to the real identity before the operation even started

The VM and the prepaid card did nothing. The phrase did everything.

Cross-contamination symptoms

If a persona has been linked to your real identity — by the provider, by a recommendation engine, or by an analyst — you will often see it before anyone tells you.

- **Personal-topic suggestions** appearing on the op account
- **"People you may know"** / "related accounts" surfacing your real contacts
- **Recommendation engines** recommending content tied to your real interests
- **Account warnings** that mention the "other" account by implication

When you see any of these, the persona is burned. Do not rehabilitate it.

Rules to remember

- Your account knows more than you typed: billing, IP, device, OAuth, prompt history, and prose style.
- One persona = one browser profile = one device (or dedicated VM) = one payment trail.
- Never use "log in with <real-account>" on a persona.
- Your writing style is a fingerprint. Persona = persona voice, consistently.
- If contamination is suspected, burn the persona. Do not rehabilitate it.

