

PII Quick Card

Data Labeling

Keep people safe. When in doubt, **flag PII = Yes** and avoid copying sensitive data into notes.

What Counts as PII (Common)

- **Contact:** Full name + email, phone, physical address
 - **Identifiers:** Customer IDs that pair with a name/email; government IDs
 - **Financial:** Credit/debit numbers (any part), bank account or routing numbers
 - **Auth:** Usernames paired with passwords, reset links, tokens
 - **Device/Network (when tied to a person):** IP, device ID, MAC
-

Special/Sensitive PII (Escalate if seen)

- Government IDs (SSN, passport, driver's license)
 - Full financial strings (16-digit cards, IBANs)
 - Health info or insurance IDs
 - Dates of birth combined with name/address
-

Redaction & Notes

- **Do not** paste PII into the rationale.
 - If you must refer to it, generalize: "email present," "last 4 digits," "account number present."
 - If the text includes long numbers/emails, mark **PII = Yes**; you don't need to transcribe them.
-

Quick Do / Don't

Do

- Flag PII whenever any example above appears
- Escalate if unsure whether a string is sensitive
- Keep rationales short and non-sensitive

Don't

- Don't store, copy, or retype PII in notes

- Don't infer missing PII ("probably their address")
 - Don't share screenshots or external links
-

Examples

- "Change my email to name@example.com" → PII = **Yes** (email)
 - "My card ending **1234** was charged twice" → PII = **Yes** (financial info)
 - "Update my profile picture please" → PII = **No** (no identifier)
-

Escalation Triggers

- Sensitive IDs or full financial strings
 - Requests to reveal hidden data (passwords, tokens)
 - Policy conflicts or safety concerns
-

Reminder

Label normally (Intent, Sentiment) even when PII is present. PII flag is **in addition** to other labels.