PIIANO

PII by Design

This one-pager aims to help developers and other privacy professionals easily understand what sensitive personal data like PII (personally identifiable information) can refer to and to provide some privacy engineering tips.
This way you can prioritize what's important in order to protect it better, reducing its exposure and complying with privacy regulations. Normally, some PII like identifiers are more sensitive than others.
If you are a developer working on code that includes handling PII or other sensitive data, you should understand the required privacy policies that apply to it. Consult with a privacy professional within your organization, such as a DPO, Chief Privacy Officer, Privacy Engineer, or your legal department.

Privacy by Design (PbD)

PbD is the planning – taking all privacy considerations as part of the development - and implementation of a system that fully supports individuals' rights and protects their data. Learn more here.

Privacy Engineering Tips

1. Never log PIIs in plaintext
2. Don't pass PIIs in HTTP Get params in URLs – they might get logged
3. Don't expose internal objects and IDs to external clients (e.g a JSON obj might contain more fields than you expect)
4. Omit identifiers in analytics pipelines – exposed to all employees
5. Login system shouldn't leak intel about existing users
6. Centralize & segregate what matters - identifiers & sensitive data – gain more control
7. Encrypt, mask, tune access policies & audit PIIs
8. Masking must happen in server side
9. Keep an updated data inventory map of where to find your customers' PIIs
10. Tokenize PIIs instead of passing them directly across systems in plaintext

Disclaimer

Related regulations: GDPR, CCPA, CPRA, FERPA, HIPAA.
There are some differences in the definition of PII in different privacy laws. Please consult with your privacy focal point to know what applies to you.

Privacy Principles

When handling PII, you have to follow obligations regarding the data.

Limits on collected data:

- **Transparency** – You need to tell people what you do with it and get a consent
- **Purpose limit** – You must use sensitive information only for the core purpose you specified
- **Consent** – You have to honor users' consent preferences about using and sharing their data
- **Minimization** – Collect and process the minimum amount of data required
- **Retention** – Define policies of when to delete data; automatically

Data subject access rights:

- **Deletion** – You have to delete data of the user when requested
- **Access** – You have to provide a copy of the data of the user when requested
- **Accuracy** – It has to be correct, and you need to correct it if requested

Security:

- **Data protection** – You must keep data safe from threats
- **Security** – ACL, encryption, masking, tokenization, auditing

Misc:

- **No discrimination** – You can't punish people for exercising their rights
- AI should be trained neutrally

Key Identifiers

We believe it's most effective to segregate, protect, encrypt, and tokenize the following identifiers:
full names, display names, emails, addresses, phones, tax-id/SSN/national-ids, payment info and URLs to profile pictures or other unique personal linkers. Or **anything that** fully helps to **identify a person**.
Sometimes the combination of innocent types of data might identify a person, and thus should be protected too.

De-Identification Importance

- De-identified data becomes out of scope, and enables business activities
- Stolen customer records without PII don't require a breach notification
- Customers aren't hurt when data is pseudonymized, trust is preserved
- A cost of up to $180 per lost/stolen record of personally identifiable information, IBM 2021

In Case of a Breach

If the stolen information contains PII, you will need to report it to the relevant individuals and sometimes to regulators (usually, there are some thresholds). Note that in some cases, even if you are not sure which information was stolen exactly, you're still obliged to report a breach.

Categories

Personal Identifiers

- **Full name**, initials, alias, maiden name, mother maiden name, **nickname, handle**
- **Phone numbers**, fax number, business phone numbers
- **Personal email**, business email, family emails
- **Date of birth**, marriage date, death date
- Emergency contacts & persons
- Location: **Full address, GPS coordinates**, longitude, latitude, geolocation, **street**, state, country, county, city, **zip code**, area code, district, province, region, town, village, township, community, parish

Education Information

Any information related to an enrolled student that is directly related to the student, such as grades, enrollment information, conduct history, etc.

Financial Information

Credit card number (PAN), financial assets/investments, **tax information**, NPI: payment history, **bank account number** (BAN), account balances, loans, CC & DC purchases, court records, consumer report, **routing number**, credit score, debit card number, cardholder PIN, card expiration date.

Insurance Information

Medical insurance, life insurance, supplemental insurance.

Sensitive Personal Information

- **Passport information, driving license,** (KYC related docs), **SSN, government IDs, personal tax ID**, vehicle identifiers, **birth certificate**, visa information, green card information, naturalization certificate, tribal identification number
- A consumer's **account login information**, financial account, **debit card**, or **credit card number** in combination with any required security or access code, password, or credentials allowing access to an account
- **Precise geolocation & time**
- A consumer's racial or ethnic origin, religious or philosophical beliefs, or union membership
- The contents of a consumer's mail, email, and text messages unless the business is the intended recipient of the communication
- A consumer's genetic data

Commercial Information

Records of **personal property, products or services purchased**, obtained, or considered, or other **purchasing or consuming histories** or tendencies.

Consumer Preferences

Characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities and aptitudes, sexual preferences, race, gender, religious views, political views.

Professional or Employment-Related Information

Employment status, **employer name**, employment verification, **employer identification number, job position**, grievance information, payroll and benefits information, leave-of-absence reason, disciplinary records.

Medical Information

Account number, medical histories, test and laboratory results, mental health conditions, treatment information.

Biometric Information

Facial geometry, fingerprint, retina scan, voice signature, gait.

Internet and Other Electronic Network Activity Information

IP & MAC addresses, cookies stored on a web browser, **usernames, passwords**, device information, ecommerce order ID, internet account numbers, device fingerprinting, **browsing history, search history, pictures**, audio, video, **online identifier, signature/avatars** & handwriting, electronic, visual, thermal, olfactory, or similar information.