General Data Protection Regulation (GDPR)

The General Data Protection Regulation (GDPR) is a law made by the European Union (EU) that governs how personally identifiable information is collected, processed, and eventually deleted from a computer system.

It is a regulation that requires businesses to protect their personal data. Personal data is defined broadly in GDPR:

- Basic identity information like name, address, and ID numbers.
- Health and genetic data.
- Biometric data.
- Racial data.
- Political opinions.

Purpose: Protects personal data and privacy of EU citizens.

Key Principles:

- Data must be collected lawfully and transparently.
- Individuals must give consent.
- Right to access, rectify, and delete data ("Right to be Forgotten").

Example: A UK-based fitness app collecting data from EU users must ask for explicit consent and allow users to delete their profiles completely upon request.

Examples of GDPR

📱 1. WhatsApp Privacy Update (2021)

WhatsApp announced a privacy policy update requiring users to share data with Facebook.

🧾 2. Google Fined €50 Million in France (2019)

Google failed to properly inform users about how their data was being collected and used for personalized ads.

Violation: Lack of transparency and valid consent.

GDPR Action: The French data regulator (CNIL) fined Google €50 million.

Lesson: **Clear privacy policies** and user consent are mandatory under GDPR.

🛬 3. British Airways Data Breach (2018)

Hackers accessed personal and payment data of 400,000 customers.

GDPR Action: UK’s ICO issued a **record fine of £20 million**.

Violation: Failure to protect user data through adequate cybersecurity measures.

FERPA (Family Educational Rights and Privacy Act - USA)

Real-World School Policy Response
Many U.S. schools now:

- ➢ Use **secure online portals** (like PowerSchool, Blackboard, or Canvas)
- ➢ Require **student PINs or logins** to access grades
- ➢ Train staff to **never disclose records over phone/email** without verification

Purpose: Protects the privacy of student education records.

Applies to: Schools and colleges receiving U.S. Department of Education funds.

Rights Given:
- ➢ Parents/students can review and request changes to records.
- ➢ Institutions cannot share educational data without consent.

- **Example**: A teacher cannot share student grades publicly with names attached.

Examples of FERPA

📚 Example 1: Unauthorized Grade Sharing

Scenario: A teacher at a U.S. high school accidentally sends a group email to parents containing all students' grades and personal IDs.

FERPA Violation: This is a breach of student educational records. FERPA mandates that grades and personal data cannot be shared without student or parent consent.

Action Taken: The school must notify the affected families, report the breach, and take corrective steps (e.g., retrain staff, update email practices).

🧾 Example 2: Posting Report Cards in Public

Scenario:
A college posts student report cards on a bulletin board with full names and student IDs.

FERPA Violation:
Educational records are being shared **publicly** without consent, which is a direct violation of FERPA.

- **Correct Practice**:
  Schools must only post anonymized student info or get written permission from students.

HIPAA (Health Insurance Portability and Accountability Act - USA)

In the United States, **HIPAA** is what is considered the **standard in protecting patient information**. Any company dealing with covered **health information (PHI)** is mandated to have and enforce in practice all process, network, and physical security necessary.

Purpose: Protects sensitive patient health information.

Key Provisions:

- ➢ Healthcare providers must secure patient data.
- ➢ Data cannot be shared without patient consent.

Example: A hospital storing patient records in cloud storage must ensure it's encrypted and HIPAA-compliant.

Examples of FERPA

📞 Example 1: Leaving Voicemail with Sensitive Information

Scenario: A nurse leaves a voicemail on a patient’s phone with detailed test results.

HIPAA Violation: Yes, if the patient did not authorize communication this way, it's a violation.

Proper Practice: Obtain patient consent on how they wish to receive information (email, text, voicemail, etc.)

Privacy Rule – Communications must respect patient preferences and confidentiality.

🖥️ Example 2: Lost Laptop Containing PHI

Scenario: A healthcare worker's laptop, which contained unencrypted patient data, was stolen from a car.

HIPAA Violation: Failure to encrypt devices or apply strong security measures.

Result: The healthcare organization paid a **fine of $1.5 million** and updated its device policies.

Security Rule – ePHI must be encrypted and protected on mobile devices.

IRB (Institutional Review Board)

An Institutional Review Board (IRB) is a committee that reviews and monitors biomedical and behavioral research involving human subjects to ensure the research is ethical and safe. IRBs are crucial for protecting the rights and welfare of human research subjects, ensuring that appropriate steps are taken to safeguard their interests.

The Food and Drug Administration (FDA) has also been recommended to conduct an annual risk assessment to determine whether the agency is conducting an adequate number of routine IRB inspections and to optimize the use of IRB inspections in the oversight of IRBs and protection of research participants.

Purpose: Reviews research involving human subjects to ensure ethical treatment.

Responsibilities:

- ➢ Ensure informed consent.
- ➢ Minimize risk to participants.

Example: A university study on stress among students must pass IRB review before data collection begins.

Examples of IRB

🧬 1. Facebook Emotional Contagion Study (2014)

Scenario: Researchers manipulated the news feeds of ~700,000 Facebook users to study the effect of emotional content on user behavior — without informed consent.

Ethical Issue: No IRB reviewed the study beforehand, and participants were unaware they were part of an experiment.

IRB Relevance: If conducted under an academic institution, an IRB would have likely rejected or modified the study due to lack of consent and potential emotional harm.

👶 2. Havasupai Tribe Case – Genetic Research Abuse

Scenario: Arizona State University researchers collected blood samples from the Havasupai Tribe to study diabetes, but later used the samples for unrelated research like schizophrenia and inbreeding, without tribal consent.

Violation: This was a major **ethical breach**. The tribe sued ASU, and the case highlighted the **importance of IRB approval and culturally sensitive informed consent**.

Outcome: The university settled and returned the samples. It also led to reforms in how IRBs handle Indigenous community research.

🧪 3. COVID-19 Vaccine Trials (2020–2021)

Scenario: All clinical trials for vaccines like Pfizer, Moderna, and Johnson & Johnson required **IRB review** before human testing.

IRB's Role:

- Ensured **informed consent** for participants
- Reviewed **risk vs. benefit** analysis
- Oversaw **participant safety protocols**

Outcome: IRBs ensured the trials followed **ethical standards**, and participants were properly protected during expedited research.

🧠 4. University Psychology Study – Deception

Scenario: A university psychology department conducts a study where students are told they're taking a memory test but are actually being observed for **stress response**.

IRB Involvement: This involves **deception**, which is only allowed if:

- It's essential for the study
- Participants are debriefed afterward
- There's minimal risk

Result: The IRB approved it with strict conditions — including a **debrief session** immediately after the test.

PCI DSS (Payment Card Industry Data Security Standard)

The (PCI DSS) is **an information security standard used to handle credit cards from major card brands**. It was created to better control cardholder data and reduce credit card fraud. Validation of compliance is performed annually or quarterly with a method suited to the volume of transactions. The standard is administered by the Payment Card Industry Security Standards Council, and its use is mandated by the card brands.

Purpose: Ensures safe handling of credit card information.

- **Requirements**:
  - Encryption of cardholder data.
  - Regular security testing.
- **Example**: An online store must follow PCI DSS if it stores or transmits card payment data.

Examples of PCI DSS

🏪 Example 1: Target Data Breach (2013)

Scenario: Target Corporation, a major U.S. retailer, was the victim of a massive data breach where attackers stole **credit and debit card information of over 40 million customers**.

Cause:

- Hackers gained access through a third-party HVAC vendor with weak security.
- Card data was collected through malware installed on POS (Point-of-Sale) systems.

Violation of PCI DSS: Yes — failure to:

- Segment internal systems
- Maintain secure configurations
- Monitor access and system logs

Outcome:

- Target paid over $18 million in settlements.
- It was a wake-up call for retailers to strictly adhere to PCI DSS compliance.

Examples of PCI DSS

🛍️ Example 2: British Airways Hack (2018)

Scenario: Hackers used a Mega-cart attack to steal payment card data from the British Airways website and mobile app — affecting over 380,000 transactions.

Violation: British Airways failed to secure its online payment page with proper code monitoring and security scans.

PCI DSS Connection: Lack of regular vulnerability scans and file integrity monitoring violated PCI DSS requirements.

Outcome: UK regulators (ICO) fined British Airways **£20 million** for data protection failures.

🧾 Example 3: Small Business POS Breach

Scenario: A local coffee shop saved **full card numbers** on its system to simplify repeat transactions — a direct violation of PCI DSS.

Result: After a breach exposed dozens of customers' card info:

- The shop was fined by its payment processor.
- It had to adopt a PCI DSS-compliant point-of-sale system, and undergo quarterly scans.

Best Practices for Responsible Data Handling

1. Handling Personally Identifiable Information (PII)

Info that can identify an individual (e.g., name, email, address).

Best Practices:

- ➢Encrypt stored and transmitted data.
- ➢Limit access to authorized users only.
- ➢Use pseudonymization or anonymization.

Example: A survey tool encrypts email addresses so researchers can't see them directly.

Examples of PII

🏢 Example 1: Company Collecting Employee KYC Documents

Scenario:

An HR department collects employee Aadhar cards, PAN cards, and bank statements over email for KYC — and stores them in a shared drive with open access.

📛 What Went Wrong:

- Storing **PII in an insecure location**
- No **access restriction, encryption, or audit logging**

📌 What Should Be Done:

- Store KYC documents in **encrypted, access-controlled folders**
- Use secure **employee onboarding systems**
- Limit file access to authorized HR personnel only

2. Data Security

Methods:

- Use firewalls, multi-factor authentication, role-based access.
- Regular audits and updates.

Example: A cloud database storing financial data uses AWS IAM policies to restrict access by user role.

3. Protecting Anonymity in Small Datasets

Challenge: Even without names, people can be identified if dataset is small.

Solution: Generalize or suppress unique identifiers.

Example: A health study in a small town hides birthdates and uses age ranges instead.

4. Importance of Anonymizing Data

Why:

- Reduces risk of data breaches.
- Often required by law or research ethics.

Trade-offs:

- **Interpretability** ↓ (e.g., can't follow an individual over time).
- **Accuracy** ↓ if important info is removed.

Real-world: Strava's public fitness data exposed secret military bases due to lack of proper anonymization.

5. Pitfalls of Generalizing from Small Data

Problem: Drawing conclusions from unrepresentative data.

Example: A startup surveys 30 users in urban areas and concludes their app works well—may fail in rural areas due to different behavior.

Types of Bias in Data Collection and Interpretation

1. Confirmation Bias

Only seeking data that supports existing beliefs.

Example: A researcher studying meditation only selects participants who already believe in it, ignoring skeptics.

Solution: Random sampling, peer review.

2. Human Cognitive Bias

Definition: Brain shortcuts leading to errors in judgment.

Example: Assuming older people are less tech-savvy during interviews and interpreting their data accordingly.

Solution: Use structured questionnaires and diverse evaluators.

3. Motivational Bias

Definition: Desire to get certain outcomes can influence data.

Example: A pharmaceutical company ignoring data that shows side effects in a trial to protect profits.

Solution: Third-party audits, blind studies.

4. Sampling Bias

Definition: Sample does not represent the population.

Example: Conducting a job satisfaction survey only among full-time employees, excluding part-timers.

Solution: Use stratified sampling; consider demographic diversity.

5. Biased Visualization & Data Representation

Definition: Choosing visuals that mislead or overemphasize.

Example: Using a truncated y-axis in a bar chart to exaggerate differences between products.

Solution: Use consistent scales; explain context; don’t manipulate visuals.