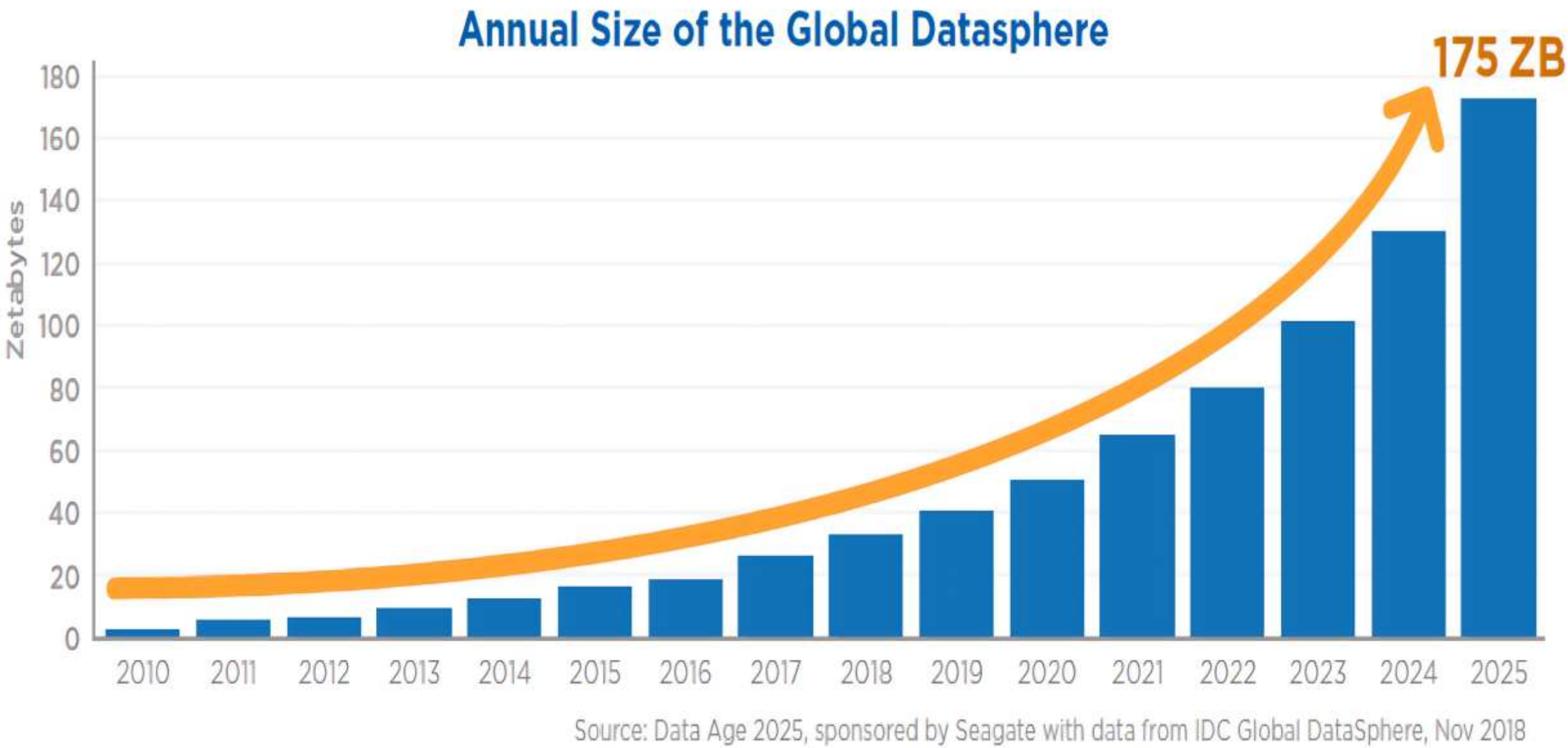
Privacy, Data Protection and Freedom of Information

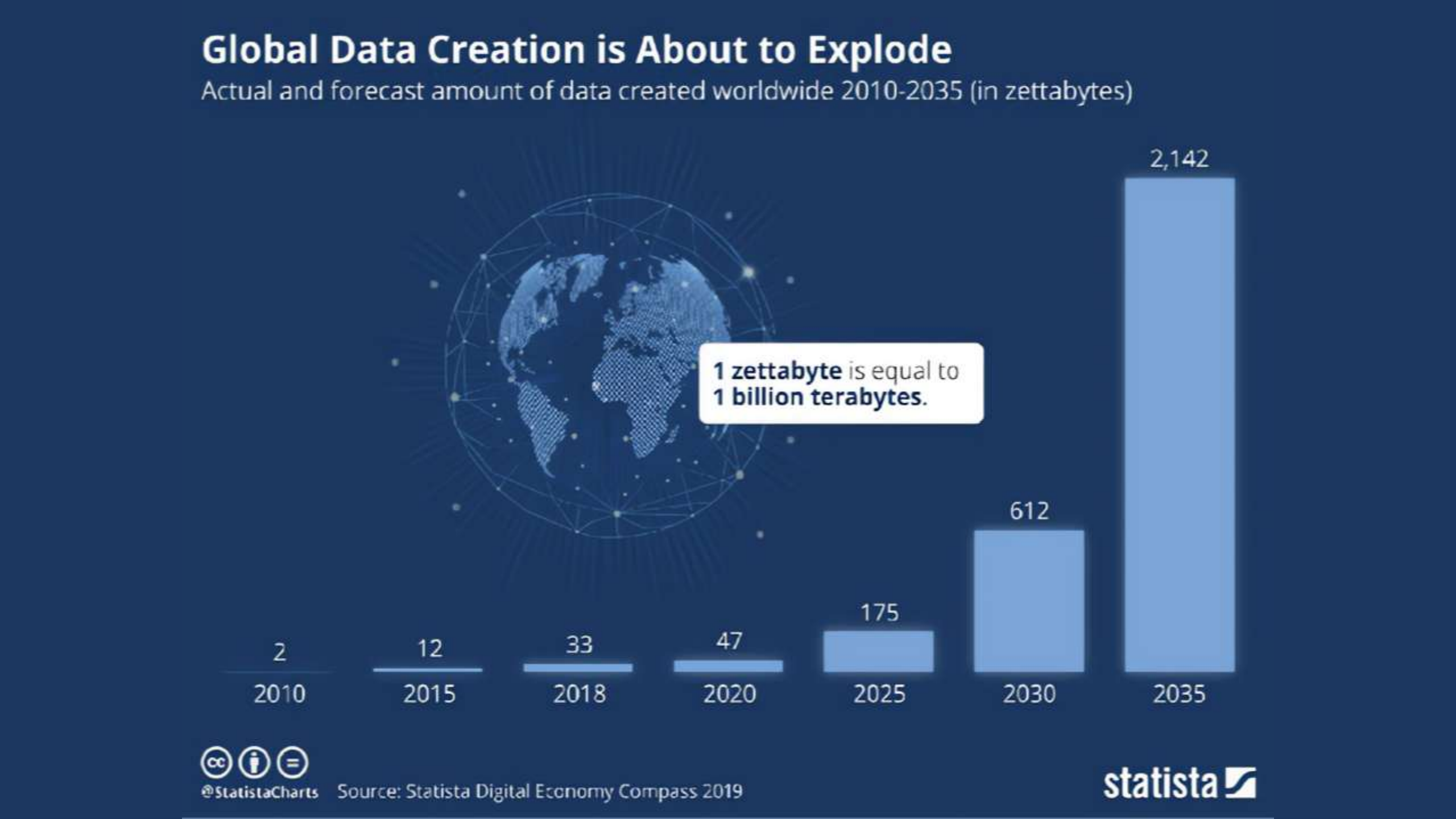
Why Data Protection Matters

- Every modern system you build will handle personal data
- Privacy rules now shape software design, especially for AI and cloud systems
- Global clients expect compliance with GDPR, CCPA, PDPL, and similar laws
- Data leaks damage users, companies, and your professional credibility


Annual Size of the Global Datasphere
175 ZB
Zetabytes
180
160
140
120
100
80
60
40
20
0
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
Source: Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018

Global Data Creation is About to Explode

Actual and forecast amount of data created worldwide 2010-2035 (in zettabytes)



Source: Statista Digital Economy Compass 2019

The Modern Data Ecosystem
Where User Data Comes From

- Apps, websites, mobile sensors
- Cloud platforms and APIs
- AI models trained on large datasets
- IoT devices, wearables, smart homes
- Social networks and digital services

- Personal data is collected continuously by apps, websites, sensors, and devices
- Information is stored across cloud platforms, often outside your country
- AI systems can analyze behavior, preferences, biometrics, and movement patterns

The Modern Data Ecosystem
Why It's Exploding

- Always-connected devices
- Automated collection (logs, trackers, cookies)
- AI systems that learn user behavior
- Cross-platform data sharing and integration

The Modern Data Ecosystem
What This Means for Developers

- More data means more responsibility
- Privacy and security must be baked into design
- Misuse or leakage can cause real harm
- Global laws restrict how data can be collected and stored

What is Privacy?

Privacy is the ability to control your personal information, your identity, and your personal space. It defines how much of yourself you choose to expose to others or to technology.

Privacy in the Digital World

- **Access Control:** Who is allowed to see or use your data
- **Usage Control:** How your information is processed, analyzed, or shared
- **Exposure Control:** How much apps, websites, or systems can infer about you

Why It Matters in Computing

- Protects users from profiling, discrimination, and unauthorized monitoring
- Builds trust between users and software systems
- Forms the foundation of data protection laws and ethical engineering practices

Elements of Privacy

1. Personal Boundaries

- **Solitude:** The right to be left alone without interference.

  Example: Not being constantly tracked by apps or devices.
- **Anonymity:** The ability to interact without revealing your identity.

  Example: Browsing without being profiled.
- **Intimacy:** Protection from monitoring of private relationships or activities.

  Example: Private conversations, messages, and home spaces.

2. Control Over Personal Information

- **Reserve:** The ability to choose what personal data you share, with whom, and under what conditions.

  Example: Deciding whether an app can access your location, photos, or contacts.

Digital Privacy Today

Common Digital Privacy Risks

- **Tracking:** Cookies, device fingerprints, location services
- **Profiling:** Targeted ads, behavioral predictions
- **Data Sharing:** Apps sharing info with third parties without clear consent
- **Data Breaches:** Leaks exposing passwords, financial info, health data

Why Developers Must Care

- Users trust you with their data
- Laws restrict how data can be collected, used, and stored
- Poor design choices can harm users and create legal liability

Core Data Protection Principles

1. Fair & Lawful Processing

- Collect and use data only when you have a valid reason
- Users must understand what is being collected and why
- No hidden or misleading data practices

2. Purpose Limitation

- Data must be collected for a **specific, clear purpose**
- You cannot later reuse it for something unrelated without consent

3. Data Minimization

- Collect **only the data you actually need**, nothing extra
- Avoid unnecessary fields, logs, or background tracking

4. Accuracy

- Keep user data correct and up to date
- Allow users to fix or update their information

5. Storage Limitation

- Do not keep data longer than required
- Define retention timelines and delete data properly

Minimization, Accuracy & Retention

1. Data Minimization

- Collect **only what is necessary** for the feature to work.
- Avoid excessive fields in forms
- Limit logging of personal information
- Do not collect background data passively unless essential

For Developers:

- Do you really need CNIC, GPS, contacts, photos, or birthdates?
- Minimize permissions for mobile apps and APIs

2. Accuracy

- Ensure user data remains **correct, consistent, and updateable**.
- Provide ways for users to edit their information
- Validate data at entry and periodically review its accuracy
- Sync data properly across microservices / databases

Why It Matters:

- Incorrect data leads to wrong decisions, credit issues, unfair profiling, and system errors.

3. Retention & Deletion

- Keep personal data **only as long as required** for the purpose it was collected.
- Define clear retention timelines
- Automatically delete logs, backups, and inactive user data
- Securely wipe data instead of just hiding or archiving it

For Developers:

- Add data expiry schedules
- Avoid “store forever” defaults
- Apply deletion to logs, caches, and backups too

Security Measures for Protecting Data

1. Access Control
Limit who can view or use personal data.

- Role-based access (RBAC)
- Least-privilege permissions
- Multi-factor authentication for sensitive data

Why It Matters:
Prevents unauthorized internal or external access.

2. Encryption
Protect data during storage and transmission.

- **At rest:** database encryption, encrypted backups
- **In transit:** HTTPS, TLS, secure APIs

Outcome:
Even if data is intercepted or stolen, it remains unreadable.

3. Integrity & Verification
Ensure data is not altered without authorization.

- Checksums, hashes, digital signatures
- Logs for data changes
- Versioning and audit trails

Purpose:
Guarantees accuracy and prevents tampering.

4. Backup & Recovery
Protect against accidental loss or system failures.

- Regular backups
- Off-site or cloud redundancy
- Disaster recovery planning

Why It Matters:
Losing user data violates laws, trust, and operational continuity.

5. Monitoring & Breach Handling
Detect threats early and respond fast.

- Intrusion detection, anomaly alerts
- Incident response procedures
- Mandatory disclosure timelines (GDPR, PDPL, CCPA)

Article 19A – Right to Information (Pakistan)

What the Constitution Says - Article 19A:

"Every citizen has the right to access information in all matters of public importance, subject to regulation and reasonable restrictions."

Why It Was Added

- Added through the **18th Amendment (2010)**
- Aligns Pakistan with global commitments like the ICCPR (International Covenant on Civil and Political Rights)
- Ensures transparency in government operations

What It Means for Technology

- Government systems must provide access to certain public data
- Public-sector software must support transparency and record-keeping
- Data cannot be withheld without a valid legal reason

Freedom of Information Laws in Pakistan

Federal Level

- **FOI Ordinance 2002**

  Citizens can request access to public records held by federal ministries and departments.

Provincial Laws

- **Punjab RTI Act (2013)**
- **KP RTI Act (2013)**
- **Sindh Transparency & RTI Act (2016)**
- **Balochistan FOI Act (2005)**

Each law defines how citizens can request information from provincial bodies.

Types of Data That is Accessible

- Public spending, contracts, policies, decisions
- Records that affect citizens or public interest

Not Accessible:

- Classified data, national security info, personal data of private citizens, internal memos.

Key Global Privacy Law – GDPR (Europe)

Why GDPR Matters

- One of the strictest privacy laws in the world
- Applies to **any company** handling EU citizens' data, even if located outside Europe
- Most international clients expect GDPR-aligned systems

Core Requirements

- **Clear consent** before collecting personal data
- **Right to access:** users can request their data
- **Right to delete (Right to be forgotten)**
- **Data portability:** users can download their data
- **Breach notification:** must inform regulators within 72 hours

GDPR PRINCIPLES

Fairness Transparency

Legal, Fair, and Transparent Data Processing with Clear Communication

Purpose Limitations & Data Minimization

Data Collection for Specific, Legitimate Purposes Only. Collect Only Necessary Data

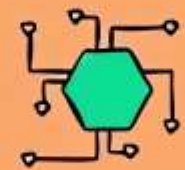
Accuracy

Maintain Accurate, Up-to-Date Personal Data with Correction Measures

Storage Limitation

Data should be retained only as long as necessary for the stated purposes.

Integrity Confidentiality



Secure Data Against Unauthorized Access and Loss

Accountability

Demonstrate GDPR Compliance and Responsible Data Management

GDPR Penalties

1. Lower Tier

Up to **€10 million** or **2% of global annual revenue**

Applies to issues like:

- Poor documentation
- Weak security measures
- Failure to notify breaches
- Not appointing a Data Protection Officer (when required)

2. Higher Tier

Up to **€20 million** or **4% of global annual revenue**

Applies to major violations like:

- Illegal data processing
- Sharing data without consent
- Ignoring user rights (access, deletion, portability)
- Using data for purposes not disclosed

10 Largest Fines for Violating GDPR

as of March 2025

Meta €1.2B	**LinkedIn** €310M
Amazon €746M	**Uber** €290M
Meta €405M	**Meta** €251M
Meta €390M	**WhatsApp** €225M
TikTok €345M	**Google LLC and Ireland** €150M

HIPAA – Health Data Privacy (USA)

HIPAA (Health Insurance Portability and Accountability Act) is the main US law protecting medical and health information.

It applies to:

- Hospitals, clinics, labs
- Health insurers
- Any software handling patient data (EHR, telemedicine, fitness & wellness apps)

What HIPAA Protects

Protected Health Information (PHI):

- Medical history
- Diagnoses, lab reports
- Prescriptions
- Biometrics
- Any data identifying a patient + their health condition

7 Steps to HIPAA Compliance

These Seven elements are the basic requirements that all effective compliance programs must address in order to adhere to the HHS Office for Civil Rights' (OCR) strict HIPAA enforcement tactics.

1 Implement Policies & Procedures

Create clear, documented rules and guidelines for handling patient health information (PHI). These guidelines tell your employee how to protect this information.

2 Appoint Compliance Officer

Choose someone responsible for ensuring that your organization follows HIPAA rules. This person helps oversee the compliance committee, which makes sure everyone in your organization follows the rules.

3 Conduct Effective Training

Make sure your employees know about HIPAA and how to keep patient information safe. Train them so they understand the rules and why they're important.

4 Develop Lines of Communication

Establish ways for employees to ask questions or report concerns about HIPAA compliance. Encourage open communication within your organization.

5 Regularly monitor & Audit

Regularly check and review how well your organization is following HIPAA rules. This helps you catch any problems early and fix them.

6 Enforce Standards

Make sure everyone in your organization knows the consequences of not following HIPAA rules. Publicize the penalties or punishments for violating these rules.

7 Swift Corrective Action

If you find that someone in your organization didn't follow HIPAA rules, take action to fix the problem and prevent it from happening again. Address any violations as soon as you discover them.

Other Important Global Privacy Laws

1. CCPA – California Consumer Privacy Act (USA)

- Gives users the right to know what data is collected
- Users can opt out of data selling
- Requires clear disclosure of data practices
- Penalties for mishandling personal data

Relevance: Any product targeting US users must follow CCPA-style rules.

2. UAE PDPL – Personal Data Protection Law (2021)

- Applies to all companies operating in or targeting UAE
- Requires consent for data processing
- Mandates secure storage and breach reporting
- Data transfers outside UAE need safeguards

Relevance: Many Pakistani dev firms serve UAE clients.

3. Saudi Arabia PDPL (2021)

- Very strict on consent and cross-border data transfers
- Sensitive data (biometric, health, financial) has extra rules
- Requires explicit user rights: access, correction, deletion
- Heavy penalties for data mishandling

Relevance: Growing KSA tech market makes this essential.

What Is Data Mining?

Data mining is the process of **discovering patterns, trends, and relationships** in large datasets using statistical and computational techniques.

Why It Matters in Tech

- Powers recommendation systems (Netflix, YouTube)
- Enables targeted advertising and personalization
- Supports fraud detection, risk scoring, sentiment analysis
- Used heavily in e-commerce, finance, healthcare, and social networks

Key Techniques

- **Classification:** Predicting categories (spam vs. not spam)
- **Clustering:** Grouping similar users/products
- **Association Rules:** "People who buy X also buy Y"
- **Prediction:** Forecasting future behavior

Secondary Use of Data

Using collected data for a **different purpose** than the one originally stated or expected by the user.

Common Examples

- A fitness app sharing health trends with advertisers
- Google using search history to personalize ads
- Meta sharing data for political micro-targeting
- Mobile apps selling precise GPS movement data to data brokers

Why It Is a Problem

- Users did not consent to the new purpose
- Increases risk of misuse and profiling
- Violates privacy laws (GDPR, CCPA, PDPL)
- Weakens user trust and can damage brand reputation

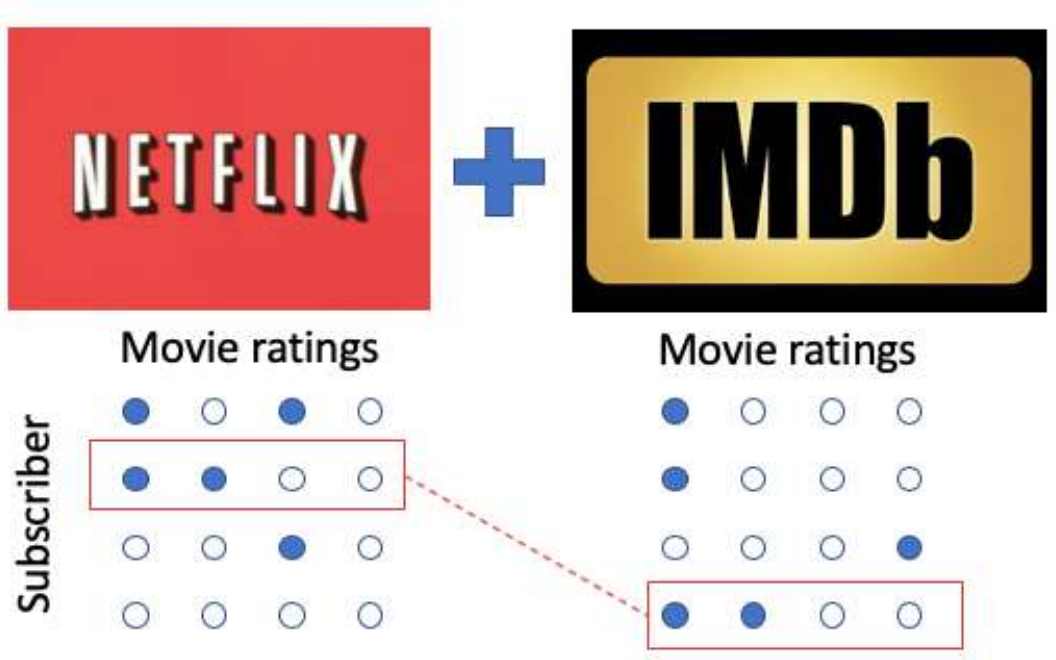
Re-Identification Risks

The process of taking **supposedly anonymous data** and matching it with other information to identify specific individuals.

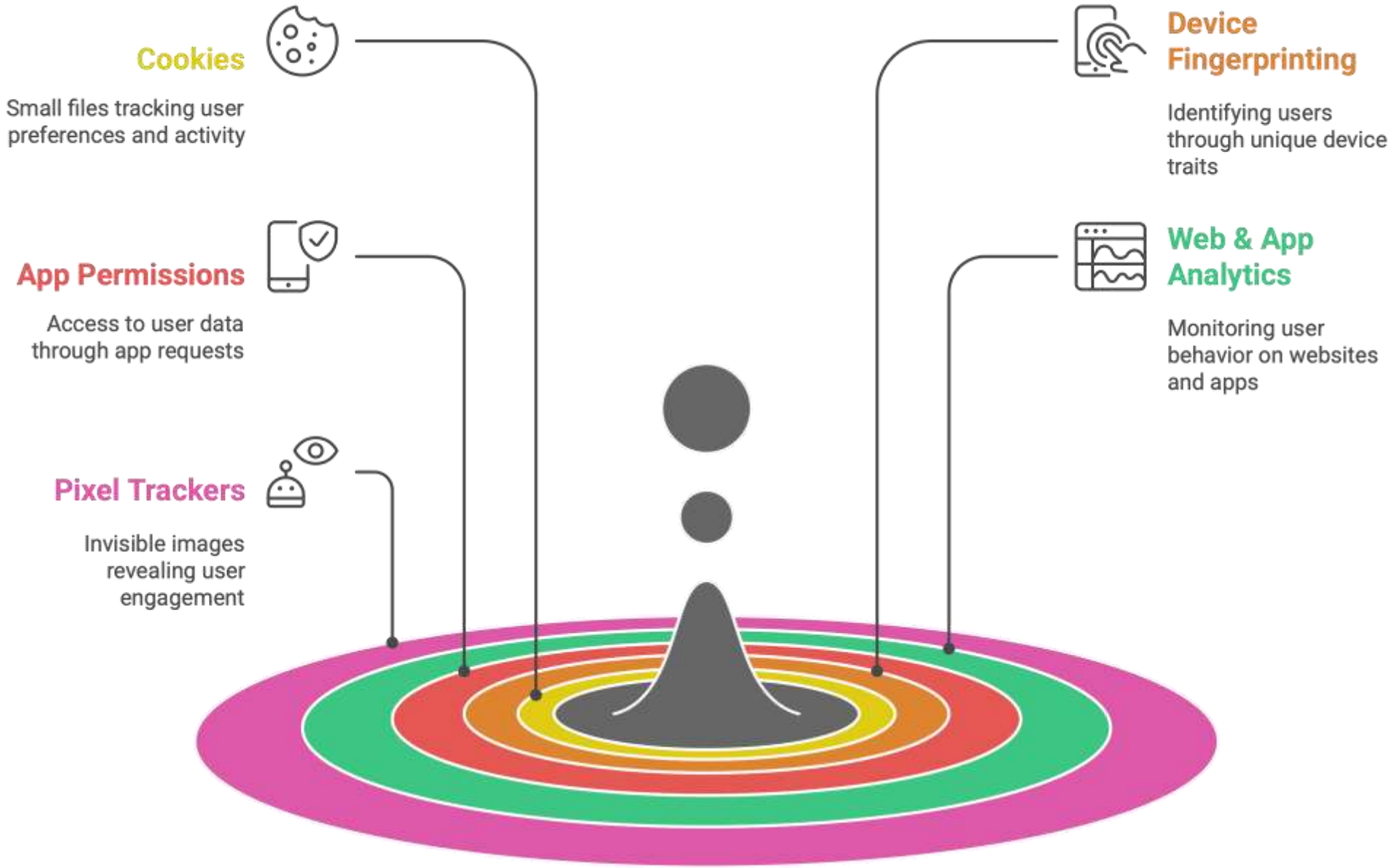
Why It Happens

- Anonymous datasets often contain patterns that can be linked
- Combining multiple datasets increases identifiability
- Modern AI and data mining make pattern matching easier

In the 2007 Netflix Prize case, researchers **Arvind Narayanan and Vitaly Shmatikov** demonstrated that they could de-anonymize the "anonymous" Netflix movie ratings by linking them to publicly available ratings on IMDb. By correlating a small number of shared ratings, they were able to identify specific users with high confidence, showing that anonymized data can be re-identified by combining it with other public datasets.



Tracking Technologies in Modern Systems



Location Data, IoT Tracking & Smart Systems

1. Location Tracking

Modern devices constantly generate location data through:

- GPS
- Wi-Fi networks
- Cell towers
- Bluetooth beacons

Risks:
Movement patterns reveal home, workplace, routines, and personal habits.

2. IoT Devices

- Smart devices collect continuous data:
- Cars (speed, braking, crash logs, routes)
- Wearables (heart rate, steps, sleep, health indicators)
- Smart home systems (cameras, sensors, voice assistants)

Risks:
Low security, weak encryption, and cloud dependence expose sensitive information.

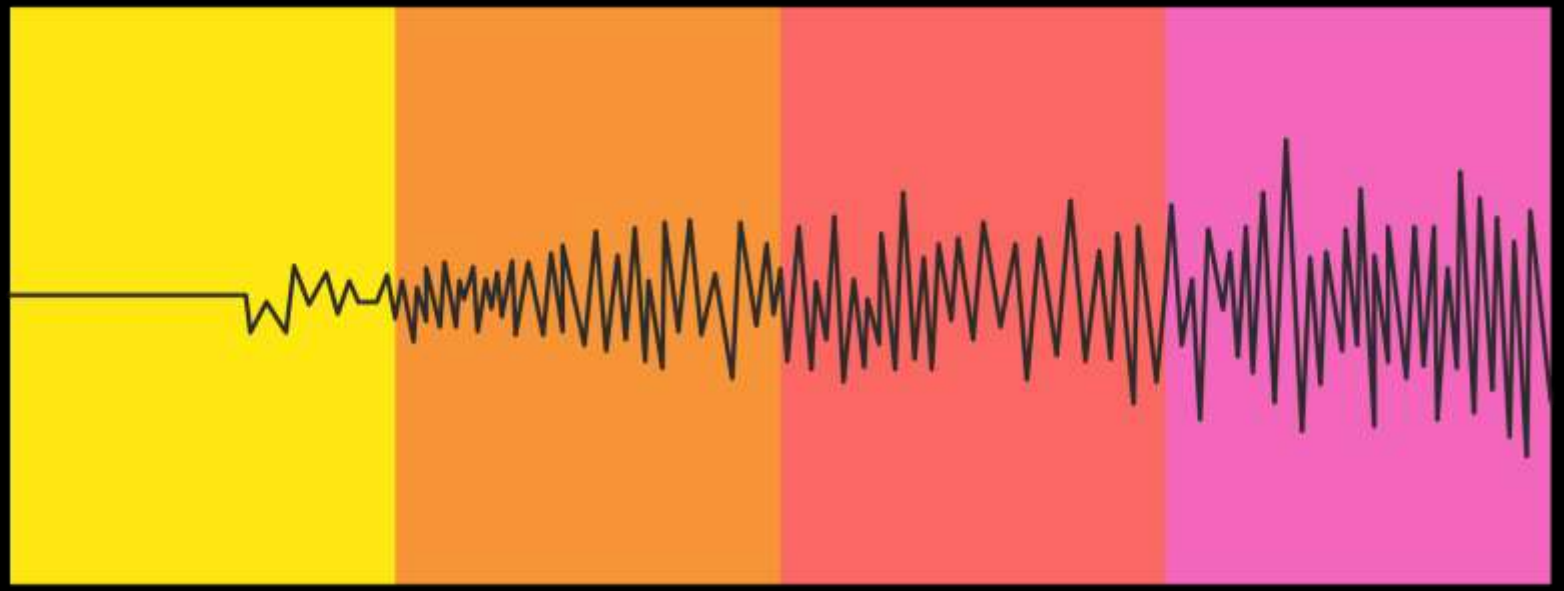
3. Smart City & Public Surveillance

- Automated license-plate readers
- CCTV with facial recognition
- Traffic, parking, and toll systems
- Safe City camera networks

Risks:
Centralized databases become attractive targets and raise serious privacy concerns.

AI-Based Identification & Modern Surveillance

Basic ⟷ Highly Invasive



Facial Recognition	Voice Recognition	Biometric Analysis	Behavior Profiling
Identifies using facial features	Identifies using speech patterns	Identifies using unique biological traits	Predicts behavior and personal traits
CCTV cameras Smartphones Public surveillance systems Social media photos	Tone Accent Speech patterns	Fingerprints Iris scans Gait (how someone walks) Typing patterns	Interests Habits Political views Shopping behavior Health conditions

Privacy by Design (For Developers)

1. Build Privacy Into the Architecture
Protect data at every layer: backend, frontend, APIs, databases
Include privacy in the initial system design, not as an afterthought

Example: Design the database schema with minimal personal fields.

2. Default to Privacy
Collect the minimum required data
Set conservative defaults for permissions, visibility, and tracking

Example: Location access off unless the user explicitly enables it.

3. Be Transparent
Tell users what data is collected and why
Make privacy settings easy to find and understand

Example: Clear prompts for camera, mic, or contact access.

4. Limit Access Internally
Use role-based access control
Ensure only the necessary team members can see or modify data

Example: Developers should not have unrestricted access to production data.

5. Secure Data Throughout Its Lifecycle
Encrypt data in storage and during transfer
Monitor for misuse
Ensure proper deletion and retention policies

Example: Automatic deletion of logs older than 90 days.

6. Regular Testing
Privacy impact assessments
Penetration tests
Misconfiguration checks

Example: Check if logs accidentally store sensitive info.

Secure Data Lifecyle

Data Collection

Collect only what is necessary and clearly inform users.

Data Sharing

Share data safely and only with trusted, compliant parties.

Data Storage

Store data securely with encryption and controlled access.

Data Retention

Keep data only as long as needed, then remove it on schedule.

Data Use

Use data only for the purpose it was collected.

Data Deletion

Delete data securely and ensure it is removed from all systems.

Thank You