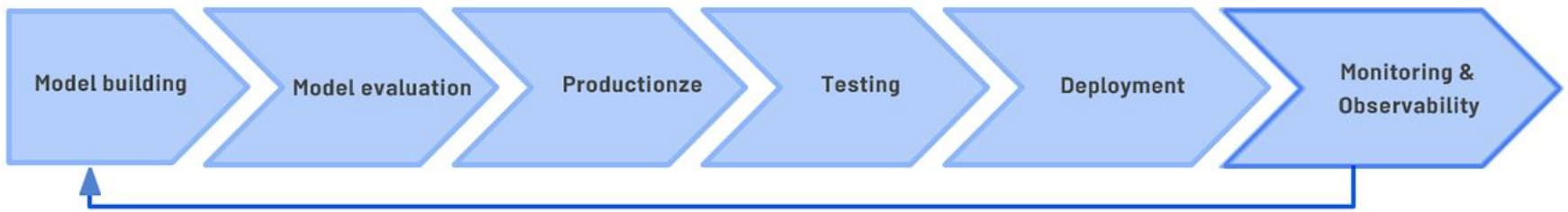
Deploying ML models in Production

Part 4: Model Monitoring

Ramin Toosi

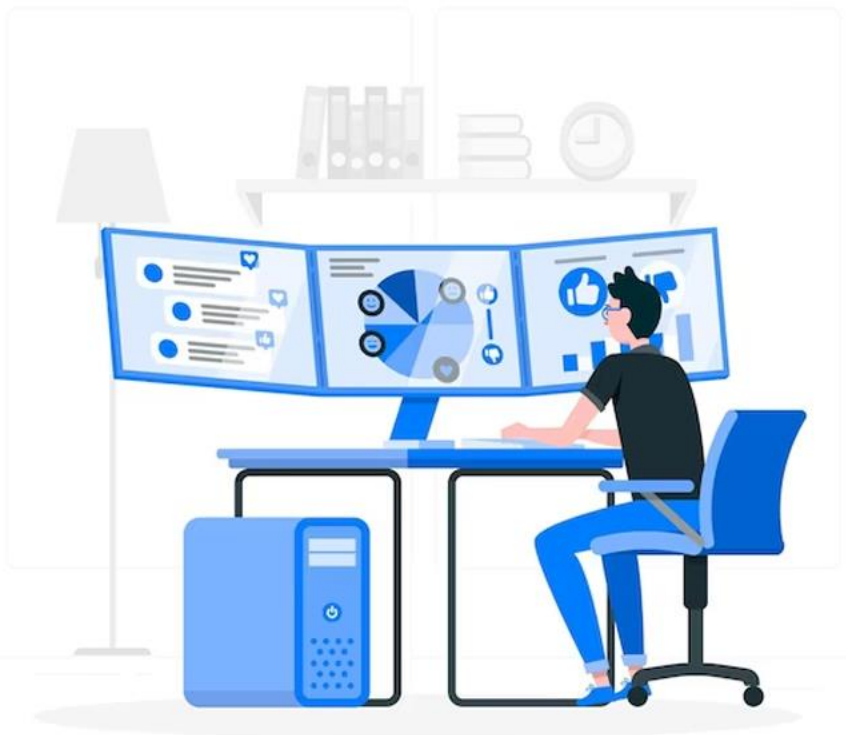
ML Infrastructure



Why Monitoring Matters

"An ounce of prevention is worth a pound of cure"

- Benjamin Franklin



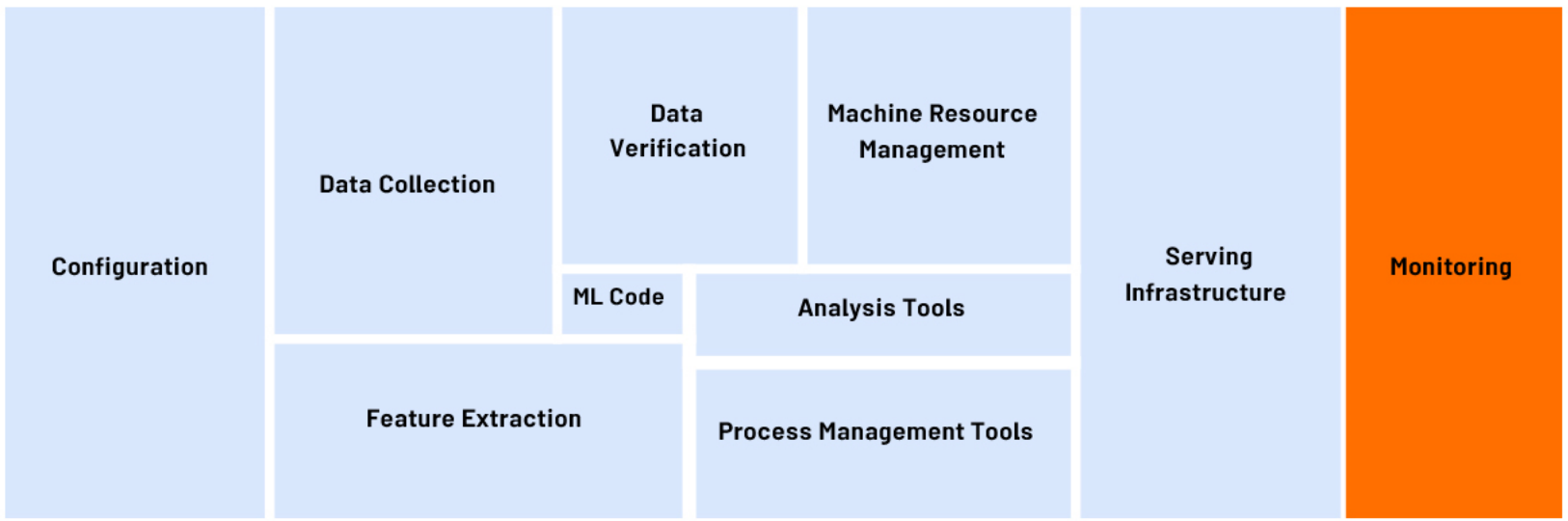
Why do you need monitoring?

- **Immediate Data Skews**
  - Training data is too old, not representative of live data
- **Model Staleness**
  - Environmental shifts
  - Consumer behaviour
  - Adversarial scenarios
- **Negative Feedback Loops**

Monitoring in ML Systems

- **ML Monitoring (functional monitoring)**
  - Predictive performance
  - Changes in serving data
  - Metrics used during training
  - Characteristics of features
- **System monitoring (non-functional monitoring)**
  - System
  - performance
  - System status
  - System reliability

Why is ML monitoring different?

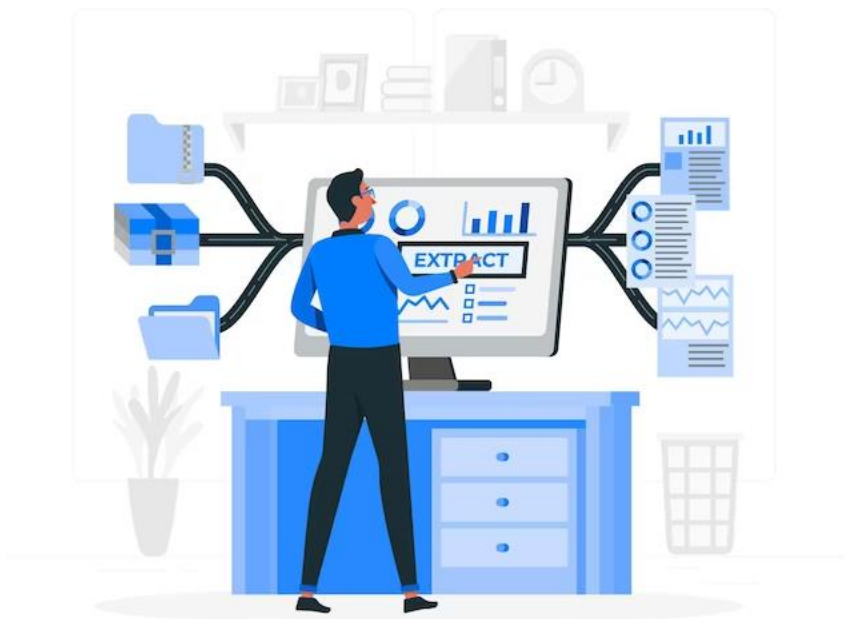


What is observability?

- Observability measures how well the internal states of a system can be inferred by knowing the inputs and outputs
- Observability comes from control system theory
- Observability and controllability are closely linked

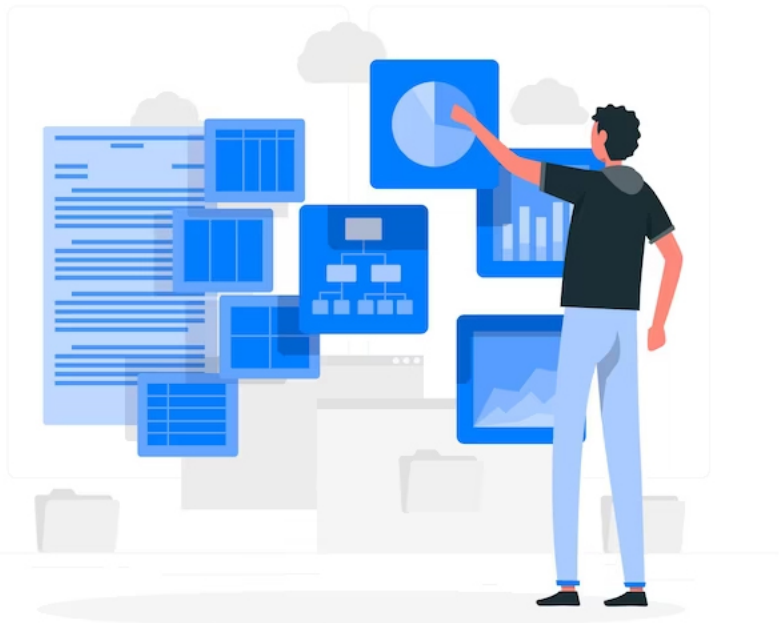
Complexity of observing modern systems

- **Modern systems can make observability difficult**
  - Cloud-based systems
  - Containerized infrastructure
  - Distributed systems
  - Microservices



Deep observability for ML

- Not only top-level metrics
- Domain knowledge is important for observability
- TensorFlow Model Analysis (TFMA)
- Both supervised and unsupervised analysis



Goals of ML observability

- **Alertable**
  - Metrics and thresholds designed to make failures obvious
- **Alertable**
  - Root cause clearly identified

Basics: Input and output monitoring

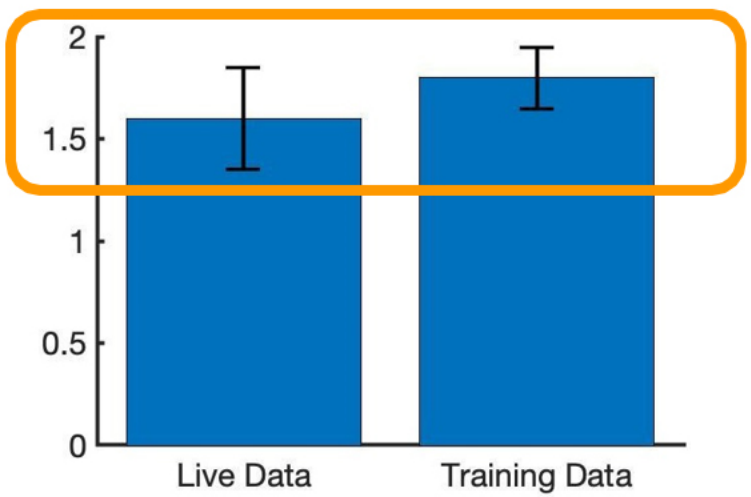
- Model input distribution
- Model prediction distribution
- Model versions
- Input/prediction correlation

Input Monitoring

- **Do these check out?**
  - Errors: Input values fall within an allowed set/range?
  - Changes: Distributions align with what you've seen in the past?
  - Per slice, e.g., marital status (single/married/widowed/divorced)

Prediction Monitoring

- **Statistical significance**
  - **Unsupervised: Compare model prediction distributions with statistical tests**
    - **e.g., median, mean, standard deviation, min/max values**
  - **Supervised: When labels are available**



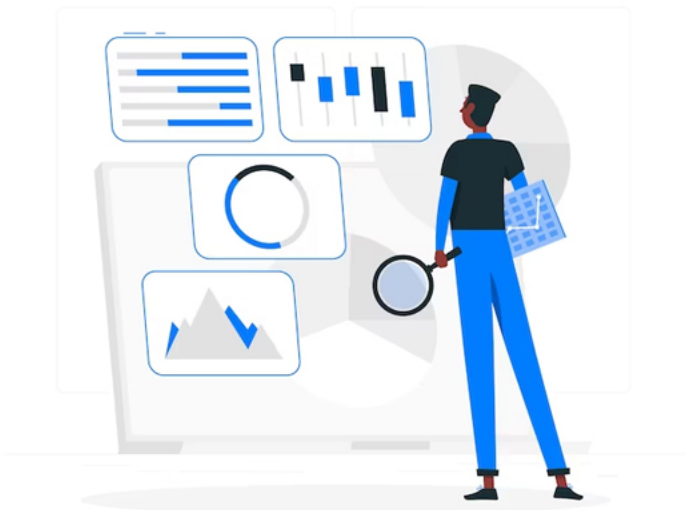
SEM

Operational Monitoring

- **ML engineering**
  - Latency
  - IO / Memory / Disk Utilisation
  - System Reliability (Uptime)
  - Auditability
- **Software engineering**
  - Receiving an HTTP request Entering / leaving a function A user logging in
  - Reading from net / writing to disk

Steps for building observability

- Start with the out-of-the-box logs, metrics and dashboards
- Add agents to collect additional logs and metrics
- Add logs-based metrics and alerting to create your own metrics and alerts
- Use aggregated sinks and workspaces to centralize your logs and monitoring



Logging

- **Log: An event log (usually just called “logs”) is an immutable, time-stamped record of discrete events that happened over time.**

Logging - Advantages

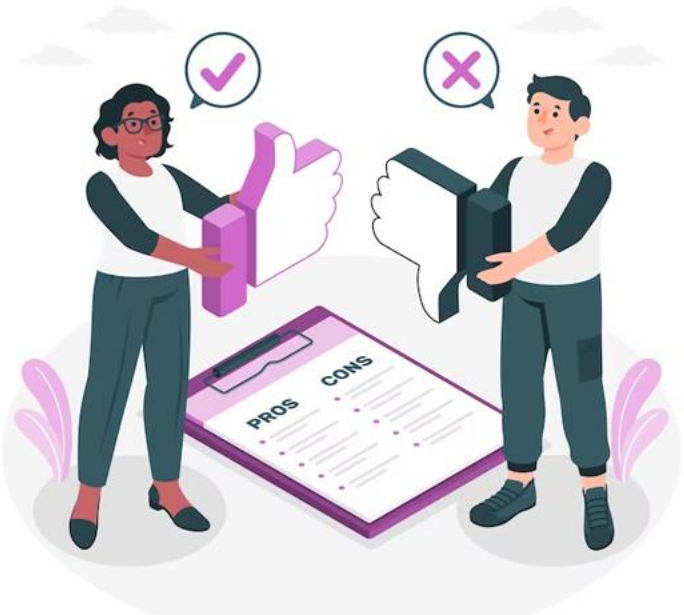
Advantages

- Easy to generate
- Great when it comes to providing valuable insight
- Focus on specific events

Logging Disadvantages

Disadvantages

- Excessive logging can impact system performance
- Aggregation operations on logs can be expensive (i.e., treat logs-based alerts with caution)
- Setting up & maintaining tooling carries with it a significant operational cost



Logging in Machine Learning

Key areas

- Use logs to keep track of the model inputs and predictions

Input red flags

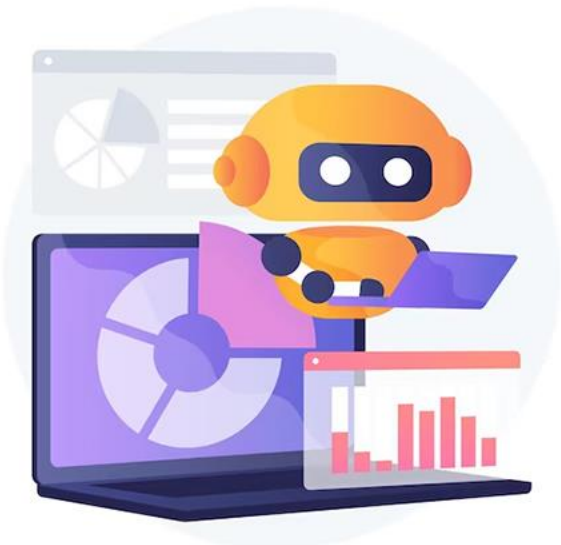
- A feature becoming unavailable
- Notable shifts in the distributions
- Patterns specific to your model

Storing log data for analysis

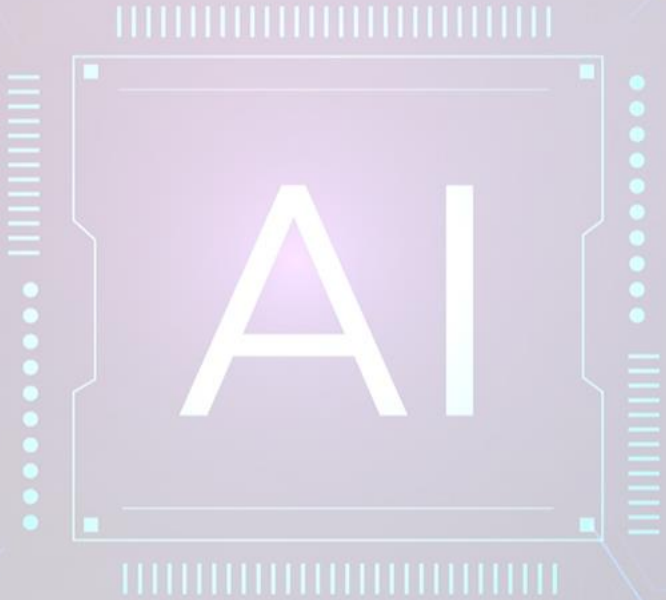
- Basic log storage is often unstructured
- Parsing and storing log data in a query-able format enables analysis
  - Extracting values to generate distributions and statistics
  - Associating events with timestamps
  - Identifying the systems
- Enables automated reporting, dashboards, and alerting

New Training Data

- Prediction requests form new training datasets
- For supervised learning, labels are required
  - Direct labeling
  - Manual labeling
  - Active learning
  - Weak supervision



Secure & Private AI



General Data Protection Regulation (GDPR)

- Regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA)
- Give control to individuals over their data
- Companies should protect the data of employees and consumers
- When the data processing is based on consent, the data subject has the right to revoke their consent at any time

Security and Privacy Harms from ML Models

Informational Harms

- Relate to unintended or unanticipated leakage of information

Behavioural Harms

- Relate to manipulating the behavior of the model itself, impacting the predictions or outcomes of the model

Data Anonymization

- **Recital 26 of GDPR defines Data Anonymization**
- True data anonymization is:
  - Irreversible
  - Done in such a way that it is impossible to identify the person
  - Impossible to derive insights or discrete information, even by the party responsible for anonymization
- GDPR does not apply to data that has been anonymized

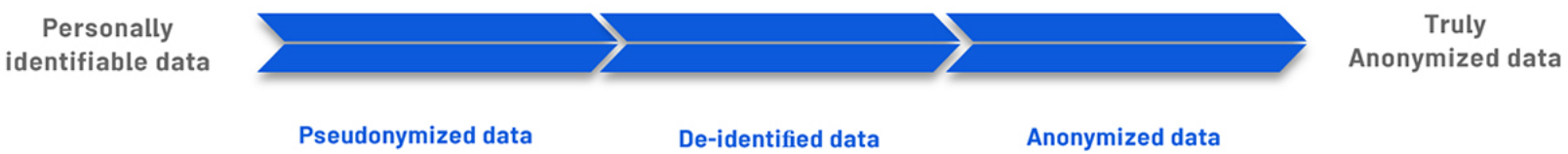
Pseudonymisation

- GDPR Article 4(5) defines pseudonymisation as:
  - "... the processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information"
- The data is anonymized by switching the identifiers (like email or name) with an alias or pseudonym

Pseudonymisation vs Anonymization

Information	Pseudonymized	Anonymized
Chelsea	Puryfrn	*******
Kumar	Xhzne	*******
Zaed	Mnrq	*******
John	Wbua	*******
Doe	Qbr	*******
Alex	Nyrk	*******

Spectrum of Privacy Preservation

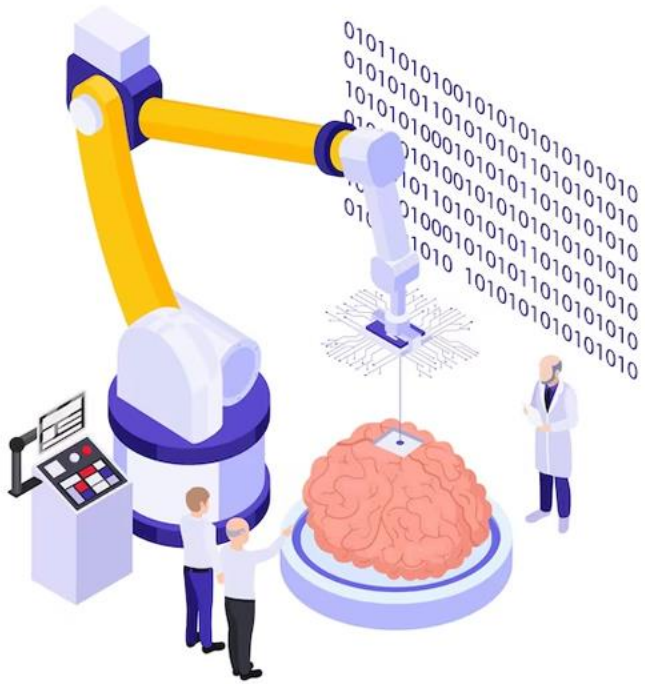


What Data Should be Anonymized?

- Any data that reveals the identity of a person, referred to as identifiers
- Identifiers applies to any natural or legal person, living or dead, including their dependents, ascendants, and descendants
- Included are other related persons, direct or through interaction
- For example: Family names, patronyms, first names, maiden names, aliases, address, phone, bank account details, credit cards, IDs like SSN

What is the Right to Be Forgotten?

- "The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay"



Recitals 65 and 66 and in Article 17 of the GDPR

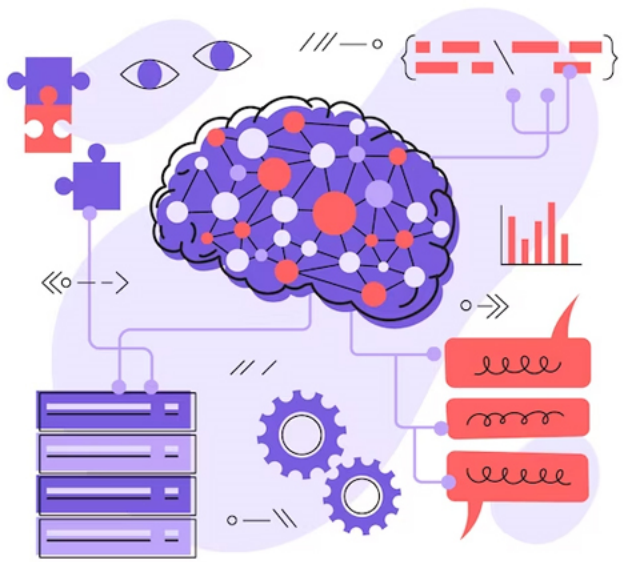
Right to Rectification

- "The data subject shall have the right to obtain from the controller without undue delay the rectification of inaccurate personal data concerning him or her. Taking into account the purposes of the processing, the data subject shall have the right to have incomplete personal data completed, including by means of providing a supplementary statement."

Chapter 3, Art. 16 GDPR

Other Rights of the Data Subject

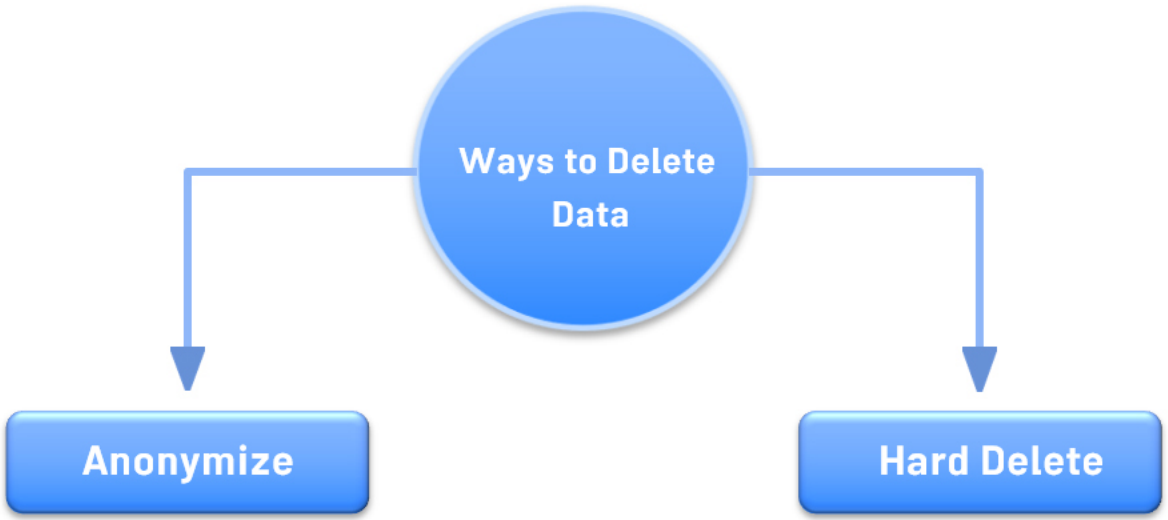
- **Chapter 3 defines a number of other rights of the data subject, including:**
  - Art. 15 GDPR – Right of access by the data subject
  - Art. 18 GDPR – Right to restriction of processing
  - Art. 20 GDPR – Right to data portability
  - Art. 21 GDPR – Right to object



Implementing Right To Be Forgotten: Tracking Data

- For a valid erasure claim
  - Company needs to identify all of the information
  - related to the content requested to be removed
  - All of the associated metadata must also be erased
    - Eg., Derived data, logs etc.

Forgetting Digital Memories



Issues with Hard Delete

- Deleting records from a database can cause havoc
- User data is often referenced in multiple tables
- Deletion breaks the connections, which can be difficult in large, complex databases
- Can break foreign keys
- Anonymization keeps the records, and only anonymizes the fields containing PII

Challenges in Implementing Right to Be Forgotten

Challenges

- Identifying if data privacy is violated
- Organisational changes for enforcing GDPR
- Deleting personal data from multiple back-ups