BUILDING DATA SECURITY FROM SCRATCH

A Blueprint for Modern Organizations

ANANDRAJ,
Enterprise Security Architect.

By the end of this presentation, you'll understand

Enterprise Data Security

Overview of controls, processes, and policies.

DPDP Act 2023 Compliance

Framework for implementing compliance across the organization.

Insider Threat Protection

Strategies to protect sensitive data from internal risks.

Data-Centric Threat Modeling

Approach to conducting threat modeling focused on data.

Continuous Data Monitoring

Fundamental concepts of continuous monitoring.

Security Operations Structure

Foundational structure for organizing data security operations.

What is Data Security ?

Core Definition

The practice of protecting digital information from unauthorized access, corruption, or theft throughout its entire lifecycle.

Guiding Principle

Fundamentally enforces the CIA Triad (Confidentiality, Integrity, and Availability) to ensure data remains trustworthy and accessible.

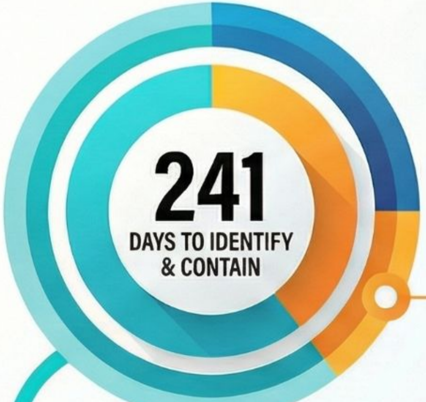
Scope of Protection

Safeguards data across all three states: at rest (storage), in transit (network), and in use (processing).

ANATOMY OF A 2025 DATA BREACH

Key statistics from the 2025 Data Breach Report, highlighting top targets, attack methods, global impact, and resolution timelines.

THE BREACH LIFECYCLE



241

DAYS TO IDENTIFY & CONTAIN

181 DAYS: AVERAGE TIME TO IDENTIFY

60 DAYS: AVERAGE TIME TO CONTAIN

FULL RECOVERY TAKES EVEN LONGER

76%

OF FULLY RECOVERED ORGANIZATIONS REPORTED THE PROCESS TOOK MORE THAN 100 DAYS

TOP TARGETS & GLOBAL HOTSPOTS

MOST COMMON TARGET: **CUSTOMER PII**

53%

OF ALL BREACHES INVOLVED COMPHOMISE OF CUSTOMER PERSONALLY IDENTIFIABLE INFORMATION (PII)

INDUSTRIES WITH HIGHEST AVERAGE DATA BREACH COSTS (2025)

- HEALTHCARE: $7.42 MILLION
- FINANCIAL: $5.38 MILLION
- INDUSTRIAL: $5.00 MILLION

GLOBAL HOTSPOTS:
HIGHEST FINANCIAL IMPACT (2025)

UNITED STATES: $10.22 MILLION

MIDDLE EAST: $7.29 MILLION

BENELUX: $6.24 MILLION

COMMON ATTACK VECTORS

MOST COMMON ATTACK VECTOR: **PHISHING**

16%

OF ALL DATA BREACHES WERE INITIATED BY **PHISHING**

MOST EXPENSIVE ATTACK: **MALICIOUS INSIDERS**

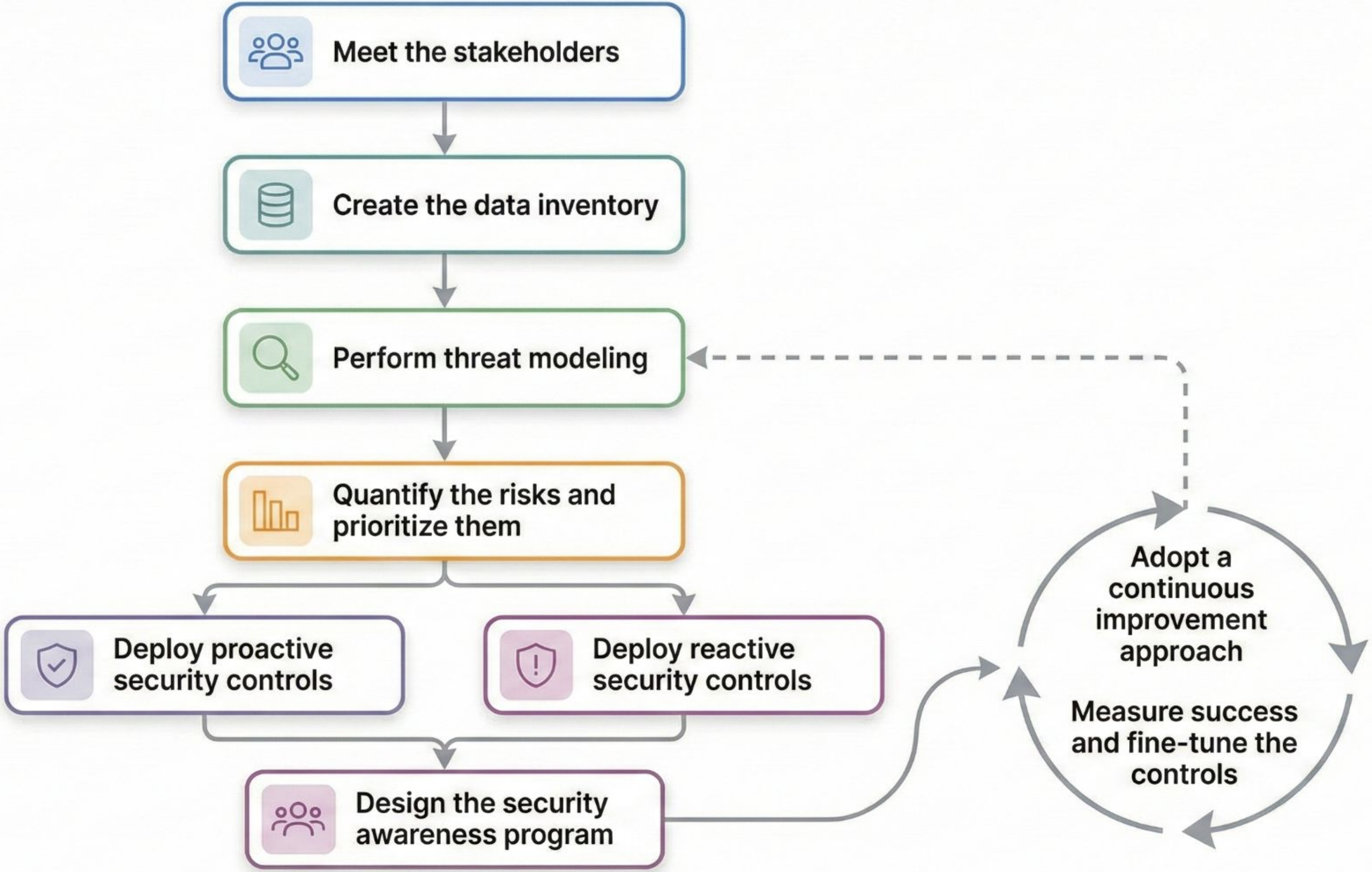
$4.92 MILLION

AVERAGE COST FOR BREACHES CAUSED BY MALICIOUS INSIDERS

NotebookLM

Data Security Roadmap

n|u



Meet the stakeholders

Create the data inventory

Perform threat modeling

Quantify the risks and prioritize them

Deploy proactive security controls

Deploy reactive security controls

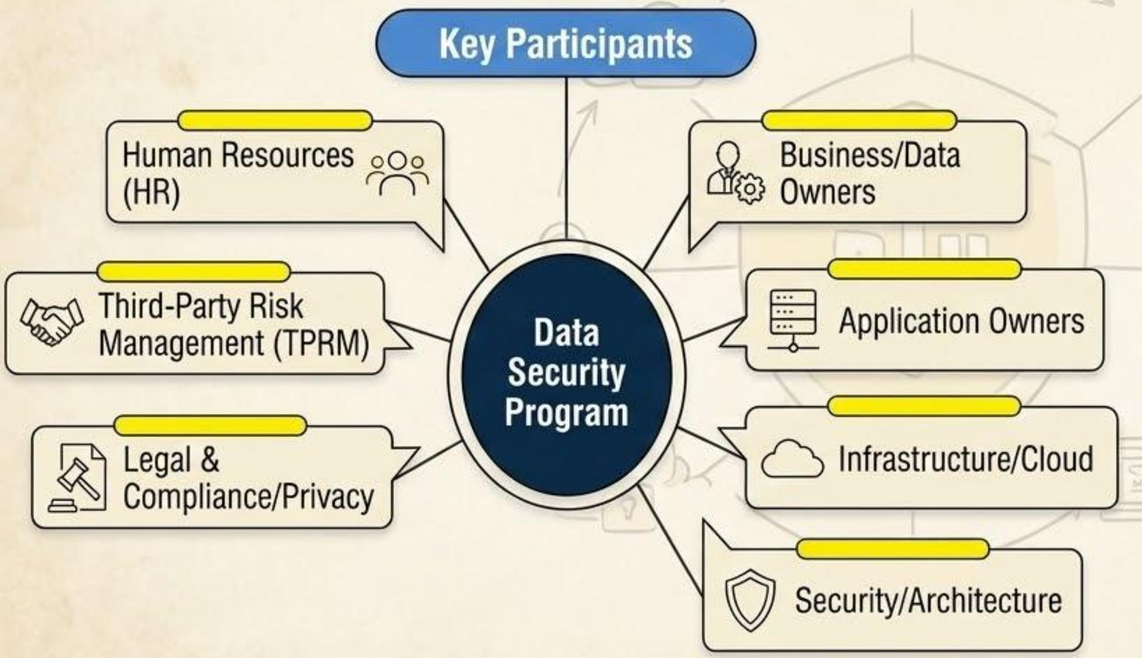
Design the security awareness program

Adopt a continuous improvement approach

Measure success and fine-tune the controls

Meet the stakeholders & Define the Mission

Stakeholder meetings are required to establish collaborative governance, translate technical risk into financial terms for informed investment decisions, and ensure accountability and alignment with regulatory compliance and business objectives.



Key Participants

Data Security Program

- Human Resources (HR)
- Third-Party Risk Management (TPRM)
- Legal & Compliance/Privacy
- Business/Data Owners
- Application Owners
- Infrastructure/Cloud
- Security/Architecture

Key Outcomes of Kick-Off

1. Agree on objectives (regulatory, business drivers, risk appetite).
2. Define scope (critical processes, systems, data, regions).
3. Assign formal data owners and custodians.
4. Approve the data security roadmap and define success KPIs.

Based on the stakeholder discussions, establish comprehensive policies to define and strengthen the organization's data security framework

n|u

CORE

Core Policies

Data Classification Policy

- **Purpose:** Categorizes data by sensitivity for appropriate controls.
- **Key Components:** Levels (Public/Internal/Confidential/Restricted), tagging rules, handling guidelines.

Access Control Policy

- **Purpose:** Enforces least privilege and role-based access.
- **Key Components:** RBAC/ABAC, MFA, PAM, regular access reviews.

Data Encryption Policy

- **Purpose:** Protects data at rest, in transit, and in use.
- **Key Components:** AES-256 standards, key management, FIPS compliance.

OPERATIONAL

Operational Policies

Data Loss Prevention (DLP) Policy

- **Purpose:** Prevents unauthorized data exfiltration.
- **Key Components:** Endpoint/network/cloud DLP, CASB, content inspection rules.

Incident Response Plan

- **Purpose:** Manages breach detection, containment, recovery.
- **Key Components:** Escalation procedures, communication, post-incident review, dark web monitoring.

External/Vendor/Third-Party data sharing Policy

- **Purpose:** Ensures external partners meet security standards.
- **Key Components:** Risk assessments, audits, data processing agreements.

GOVERNANCE

Governance Policies

Data Retention Policy

- **Purpose:** Defines data storage duration per legal/business needs.
- **Key Components:** Schedules by data type, automated archiving, annual reviews.

Data Deletion Policy

- **Purpose:** Specifies secure data removal methods.
- **Key Components:** Triggers (retention expiry), secure wipe/shredding, verification logs.

Data Storage Policy

- **Purpose:** Outlines secure storage practices across locations.
- **Key Components:** Tiered storage, encryption, access controls, backup procedures.

Employee Training Policy

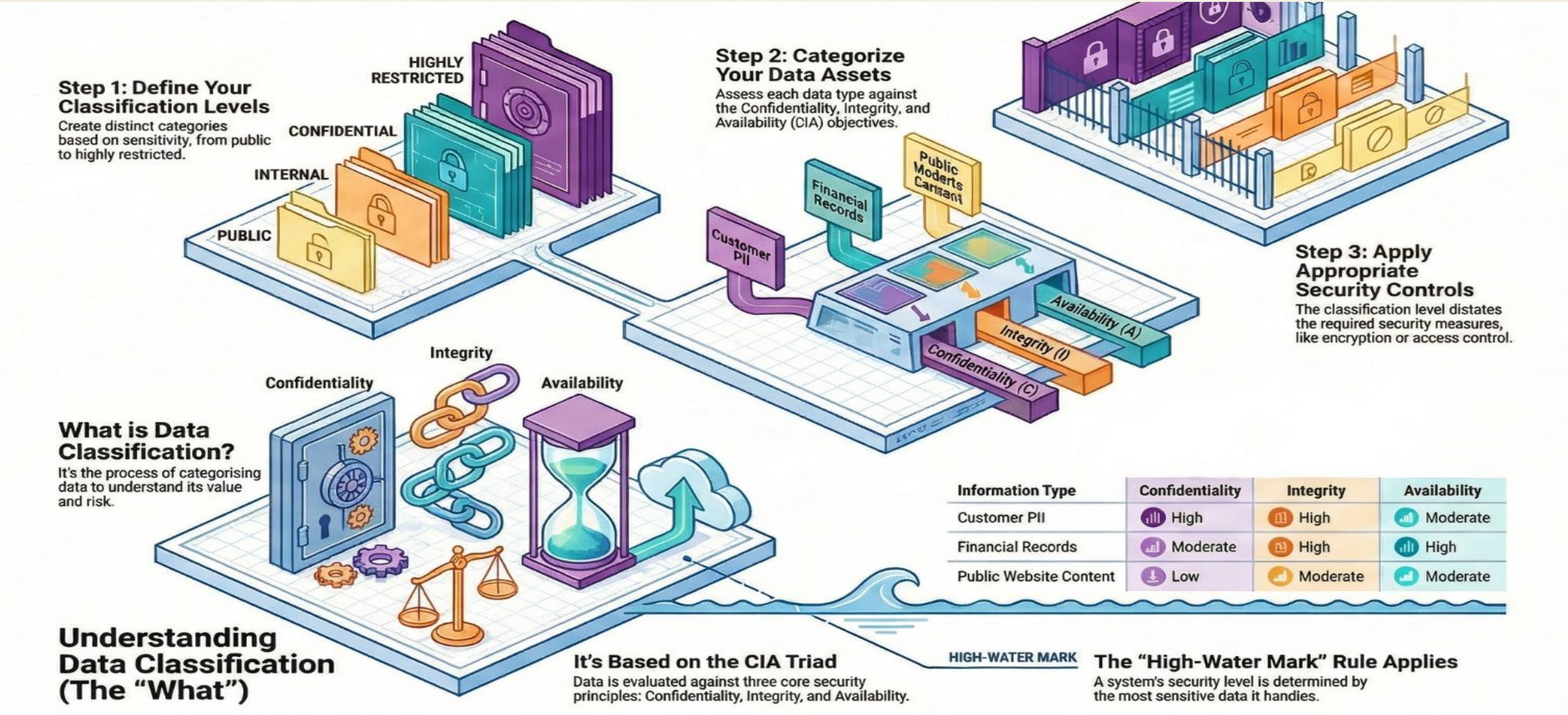
- **Purpose:** Builds security awareness across organization.
- **Key Components:** Mandatory annual training, phishing simulations, role-specific modules.

Audit and Compliance Policy

- **Purpose:** Ensures ongoing policy adherence and regulatory alignment.
- **Key Components:** Quarterly audits, framework mapping (NIST/ISO/DPDPA), remediation tracking.

Data Classification

Data classification categorizes assets (files/Data) by value and sensitivity to apply targeted protections. Using the CIA Triad to assess risk.Data classification turns DLP from pattern-matching into context-aware enforcement, enabling precise policy control, automated response, incident intelligence, and compliance visibility.



Step 1: Define Your Classification Levels

Create distinct categories based on sensitivity, from public to highly restricted.

Step 2: Categorize Your Data Assets

Assess each data type against the Confidentiality, Integrity, and Availability (CIA) objectives.

Step 3: Apply Appropriate Security Controls

The classification level distates the required security measures, like encryption or access control.

What is Data Classification?

It's the process of categorising data to understand its value and risk.

Understanding Data Classification (The "What")

It's Based on the CIA Triad

Data is evaluated against three core security principles: Confidentiality, Integrity, and Availability.

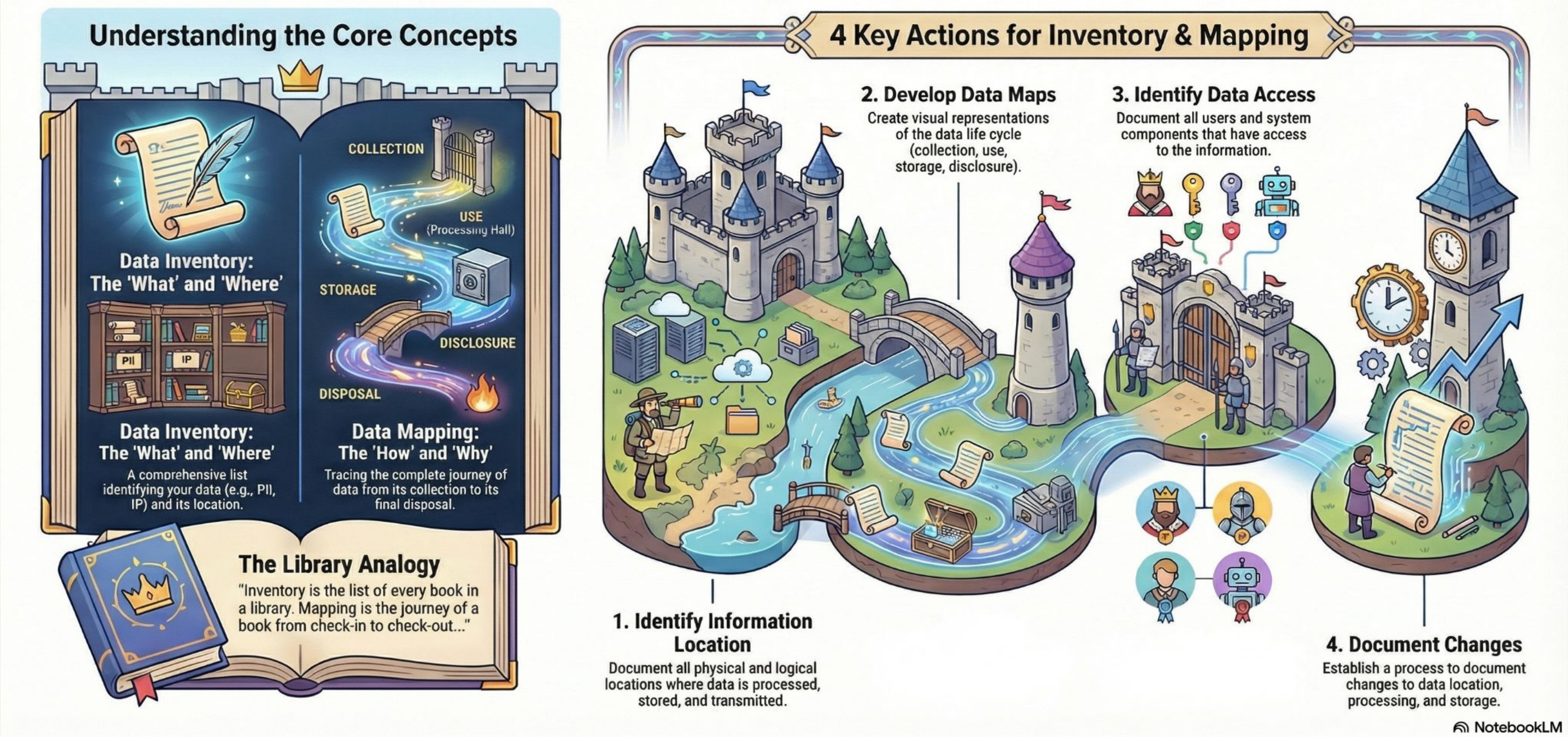
Information Type	Confidentiality	Integrity	Availability
Customer PII	High	High	Moderate
Financial Records	Moderate	High	High
Public Website Content	Low	Moderate	Moderate

The "High-Water Mark" Rule Applies

A system's security level is determined by the most sensitive data it handles.

Data Inventory & Mapping

You cannot protect what you don’t know exists. Data inventory and mapping are mandatory first steps to visualize data flows, identify critical assets, and enable targeted protection. This requires performing **data discovery** using appropriate data discovery tools.



Understanding the Core Concepts

Data Inventory: The 'What' and 'Where'

Data Inventory: The 'What' and 'Where'

A comprehensive list identifying your data (e.g., PII, IP) and its location.

Data Mapping: The 'How' and 'Why'

Tracing the complete journey of data from its collection to its final disposal.

The Library Analogy

"Inventory is the list of every book in a library. Mapping is the journey of a book from check-in to check-out..."

4 Key Actions for Inventory & Mapping

1. Identify Information Location

Document all physical and logical locations where data is processed, stored, and transmitted.

2. Develop Data Maps

Create visual representations of the data life cycle (collection, use, storage, disclosure).

3. Identify Data Access

Document all users and system components that have access to the information.

4. Document Changes

Establish a process to document changes to data location, processing, and storage.

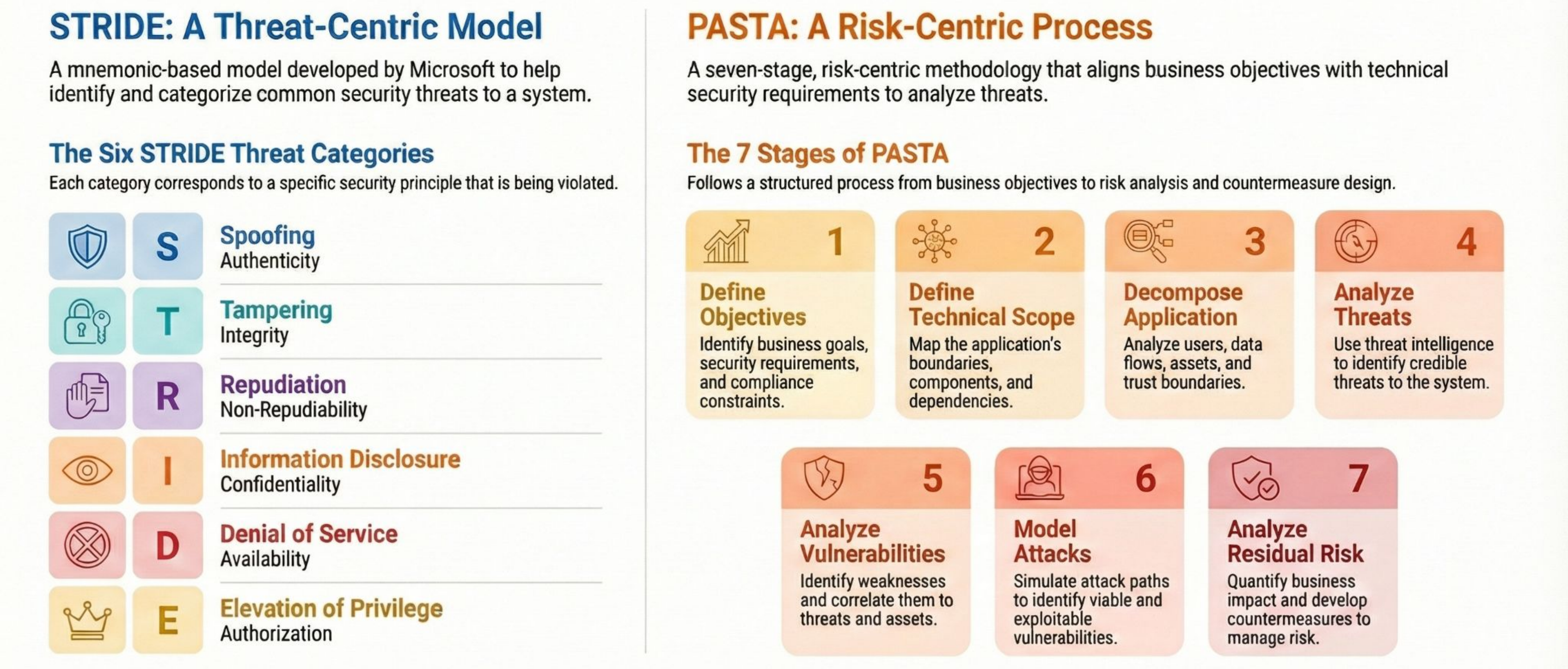
Data Inventory Example

Based on the developed data classification framework, create a comprehensive data inventory by mapping all key data attributes, including access ownership, storage location, storage method, and classification level.

Data Asset Name	Description	Owner / Dept.	Storage Location	Classification Level	Handling Policy
Website Content	Public-facing marketing text and images	Marketing	Content Mgmt System (CMS)	Public	No restrictions.
Employee Directory	List of names, internal emails, and extensions	HR / IT	Intranet Portal	Internal	Employees only. Do not share externally.
Customer Database	Client names, addresses, and purchase history	Sales	CRM (e.g., Salesforce)	Confidential	Access limited to Sales/Support. Encrypt in transit.
Source Code	Proprietary algorithms and application code	Engineering	Private Git Repo	Restricted	Need-to-know basis only. MFA required.
Payroll Records	Salaries, SSNs, Tax IDs, Banking details	HR / Finance	HRIS System	Restricted	Strictly limited. Audit logs enabled.
Support Tickets	Customer complaints and technical issues	Support	Helpdesk Software	Confidential	Support staff only. PII redaction required.

Threat Modeling

Threat modeling is a **proactive** security exercise performed during the design phase to **identify and prioritize risks** before a single line of code is written. By systematically analyzing the application's architecture through the eyes of an attacker, teams can uncover structural flaws and implement defenses when they are cheapest to fix. It effectively shifts security left, transforming it from a reactive patch into a core design feature, ensuring the system is secure by design rather than by accident.



STRIDE: A Threat-Centric Model

A mnemonic-based model developed by Microsoft to help identify and categorize common security threats to a system.

The Six STRIDE Threat Categories

Each category corresponds to a specific security principle that is being violated.

S	Spoofing	Authenticity
T	Tampering	Integrity
R	Repudiation	Non-Repudiability
I	Information Disclosure	Confidentiality
D	Denial of Service	Availability
E	Elevation of Privilege	Authorization

PASTA: A Risk-Centric Process

A seven-stage, risk-centric methodology that aligns business objectives with technical security requirements to analyze threats.

The 7 Stages of PASTA

Follows a structured process from business objectives to risk analysis and countermeasure design.

1 **Define Objectives** Identify business goals, security requirements, and compliance constraints.

2 **Define Technical Scope** Map the application's boundaries, components, and dependencies.

3 **Decompose Application** Analyze users, data flows, assets, and trust boundaries.

4 **Analyze Threats** Use threat intelligence to identify credible threats to the system.

5 **Analyze Vulnerabilities** Identify weaknesses and correlate them to threats and assets.

6 **Model Attacks** Simulate attack paths to identify viable and exploitable vulnerabilities.

7 **Analyze Residual Risk** Quantify business impact and develop countermeasures to manage risk.

Map the Security controls

Attack Vector / Security control	Email Security	Browser isolation	IAM with RBAC / ABAC/PAM / IGA & MFA	UEBA	Data Labeling / Tagging	DevSecOps	Third-Party Risk Management process	SOC	Dark and Deep web Monitoring	MDM	DLP
Phishing	🟩	🟥	🟥					🟥			
Third-Party/Supply Chain			🟥	🟥	🟥		🟥	🟥			
Malicious Insider			🟥	🟥						🟩	🟩
Compromised Credentials			🟥	🟥				🟥	🟥		
Vulnerability Exploitation		🟥		🟥				🟥			
Denial-of-Service (DDoS)						🟥		🟥			
Physical Theft/Security Issue										🟩	🟩

Security Control available	🟩
Security Control is not available	🟥

Security Controls - Data Security focused

Device/Platform	Category	Proactive/Reactive	Key Function
CASB	Cloud Security	Both	Discover & control SaaS usage, prevent shadow IT
CSPM	Cloud Security	Proactive	Detect cloud misconfiguration
DSPM	Cloud Security	Proactive	Detect data storage misconfiguration
Cloud Data Encryption	Cloud Security	Proactive	Encrypt data in transit
Network DLP	Data Protection	Proactive	Block exfiltration at network
Endpoint DLP	Data Protection	Proactive	Control USB, cloud uploads
Cloud DLP	Data Protection	Proactive	Protect SaaS apps
MDM	Endpoint Security	Proactive	Device management & remote wipe
Endpoint Encryption	Endpoint Security	Proactive	Encrypt data at rest
IAM Platform	Identity & Access	Proactive	Centralized identity management
PAM	Identity & Access	Proactive	Vault credentials, enforce approvals
MFA	Identity & Access	Proactive	Enforce multi-factor authentication
UEBA	Identity & Access	Reactive	Detect anomalous behavior
SIEM	Monitoring & Detection	Reactive	Centralize logs & detect incidents
Dark Web Monitoring	Threat Intelligence	Reactive	Monitor for stolen credentials
Brand Monitoring	Threat Intelligence	Reactive	Monitor for brand mentions
Secure Email Gateway (SEG)	Email Security	Both	Email protection + threat detection with DLP
Vulnerability Scanner	Vulnerability Mgmt	Proactive	Scan for vulnerabilities
BAS - Breach and attack simulation	Vulnerability Mgmt	Proactive	Automate attack simulation
UEM	Device Management	Proactive	Unified endpoint management
Case Management	Incident Response	Reactive	Track incidents
Forensics Platform	Incident Response	Reactive	Post-incident forensics

Risk Prioritization

The FAIR Model: Decomposing Risk into Quantifiable Factors

$$\text{Risk} = \text{Loss Event Frequency (LEF)} \times \text{Loss Magnitude (LM)}$$

FAIR quantifies risk by analyzing how often a loss occurs and its probable financial cost.

Understanding Loss Event Frequency (LEF)

Threat Event Frequency (how often a threat acts)

Vulnerability (the probability of success)

This is how often a threat is likely to cause a loss. It's derived from Threat Event Frequency and Vulnerability.

Understanding Loss Magnitude (LM)

Primary Loss

Secondary Loss (e.g., reputation damage)

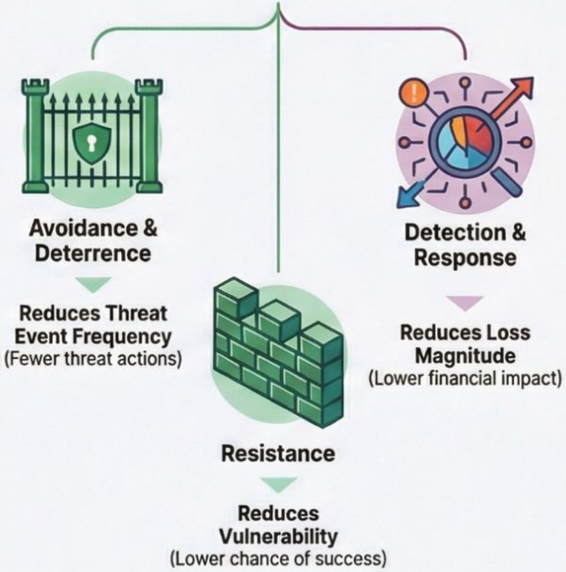
This is the probable financial impact of a loss event, including direct Primary Loss and indirect Secondary Loss.

How Security Controls Impact Quantified Risk

Controls are mapped to specific risk factors to measure their effectiveness.

The FAIR-CAM (Controls Analytics Model) shows how different control types directly reduce parts of the risk equation.

Four Ways Controls Reduce Risk



Avoidance & Deterrence

Reduces Threat Event Frequency (Fewer threat actions)

Resistance

Reduces Vulnerability (Lower chance of success)

Detection & Response

Reduces Loss Magnitude (Lower financial impact)

From Quantification to Prioritization

Quantified risk enables data-driven prioritization.

By expressing different risks in financial terms, you can objectively compare and rank them.

High Financial Impact Threat
Medium Financial Impact Threat
Low Financial Impact Threat

Focus resources on threats with the highest financial impact.

This aligns cybersecurity investments with business objectives and justifies the cost of new controls.

What Next ?

1. Design Plan & Policies

Design the **architectural plan** and **policies** required to deploy the necessary security controls.

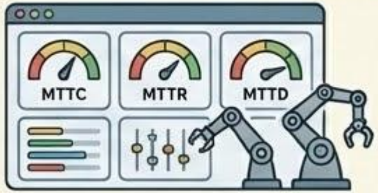
2. Develop Timelines

Develop **tentative timelines** for each project based on the risk-prioritization outcomes.

3. Establish Security Ops Team

Establish a **security operations team** to manage day-to-day activities related to the implemented controls.

4. Automate Operations



Automate security operations to reduce MTTC, MTTR, and MTTD.

5. Conduct Continuous Audits

Conduct continuous audits of security controls (e.g., CART, CTEM framework) to continuously identify security gaps.

THANK YOU!

Any Questions?

Use the QR code to get in touch with me.