# Security concepts and principles

### Security goals

* **Confidentiality**: keep something secret, data in communication or at rest (cryptography, authentication…)
* **Integrity:** no corruption or control hijacking, who can write (message, data hashing)
* **Availability:** system uptime, response time, free storage…
* **Privacy:** right to be left alone, personal information
* **Accountability:** login and audit trails (secure timestamping, integrity in logs…)
* **Non-repudiation:** two parties can’t deny they have interacted (trust a 3rd party and generate evidence)

### Security guiding principles

* Secure the weakest link (attacker needs only one flaw)
* Practice defense in depth (use layers of defense)
* Fail securely
* Compartmentalize (separate code into parts)
* Be reluctant to trust
* Principle of least privilege (minimum access and time necessary)
* Keep it simple (to reduce attack surface, tradeoff with usability)
* Promote privacy
* Hiding secrets is hard (hiding details is not enough, attacker can have ways of finding them)
* Use community resources

There are also some attacks described here, come back to check after and see if they were covered.

# OWASP Top 10

## Information Gathering

## An attacker wants to find an easy way of attacking, and as a developer you want to decide test scope, coverage and prioritize

What to gather?

* Application structure - identify pages, external links, and trust zones (areas with varying access permissions)
* Data flow within the application (observe how data moves between client and server, focusing on GET/POST requests, responses, and parameters)
* Infrastructure or platform (web server, applications, entry points, execution path, framework…)

Tools: web debugging proxy (example: Burp Suite): this can be used to capture and examine requests and responses, manipulate payloads, attacks… Automated scanners can’t capture everything, so manual testing is recommended.

## Injection Attacks

#### SQL injection

An attacker manipulates user input to inject malicious SQL into a query. This allows unauthorized access to data, data corruption, privilege escalation, or even destruction of the database. Also possible in XML with Xpath injection.

* + Blind SQL injection: guess database schema with binary search after checking if its vulnerable to sql injection.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.

* + Countermeasures
    - Blacklisting: filter quotes, semicolons, whitespace.. It can have problems with functional requirements and it is easy to bypass
    - Whitelisting: only allow well defined safe inputs. The problem is that RegExp (regular expressions) are hard to define for all safe values
    - Escape input: treating everything as strings and not as logic, like blacklisting, it could miss dangerous characters
    - Prepared statements & bind variables: decouple query statement and data input with a template, most robust.
      * ORM: Texto

        El contenido generado por IA puede ser incorrecto.
      * Manual: 
    - Mitigate impact
      * Avoid information leakage (don’t display errors and stack traces)
      * Limiting privileges
      * Encrypt sensitive data
      * Key management precautions (don’t store encryption key in DB)
      * Hash password

#### Session Management Attacks

HTTP is stateless, so unless you reauthenticate for every request , you can’t know for sure if the requests are from the same clients. With session management, you can assign session tokens and then validate them for every request. You can store session tokens by:

* Embedding in url: https://site.com/checkout?sessionToken= 1234. Exposes the token everywhere, basically)
* In hidden form field: <input type= “hidden” name = “sessionToken” value = “1234”> Must be included in every form manually, relies on client-side data.
* In browser cookie: setcookie: sessionToken = 1234 (these are set in headers of HTTP requests and responses). Safer, but server only sees cookie, not domain who sent the cookie.

##### Attacks

* **Session token theft**
  + Methods
    - **Network sniffing:** an attacker can see someone’s cookie with the session ID when they make requests to sites with HTTP (not HTTPS) and use it (since the server doesn’t check the domain from where it is being sent)

|  |
| --- |
| Find in settings.py  SECURE\_SSL\_REDIRECT = True (to force HTTPS and cookie settings)  SESSION\_COOKIE\_SECURE = True  SESSION\_COOKIE\_HTTPONLY = True  CSRF\_COOKIE\_SECURE = True  'django.middleware.security.SecurityMiddleware' (in MIDDLEWARE) |

* + - **Logout problem:** logging out should destroy the token in both client and server. Usually browsers delete the cookie (session expired or logout), but the server keep the session valid, so attackers who stole the token can still use it.

|  |
| --- |
| Interfaz de usuario gráfica, Texto  El contenido generado por IA puede ser incorrecto.This is wrong, in Django use  logout(request) + short session lifetimes in session.py |

* + - **Cross-site scripting**\*: attacker crafts malicious URL with script inside that returns the session cookie from the client’s browser (where the script is executed when client clicks into the malicious URL)

|  |
| --- |
| Django has built-in auto-escaping for XSS, so you should be fine. It could be unsecure if it has the “| safe” in html, which should only be used if 100% the data has been sanitized previously |

* + Solutions
    - Once user logged in (token around), communication through HTTPS
    - Remember to log out
    - Time-out session ID and delete expired session ID
    - Bind session token to client’s IP or computer
* **Session token predication attack**: some tokens can be predicted, seeing one or more tokens shouldn’t be able to predict others. Solution: use token generators from frameworks (in Django - check code for manually set session keys. Correct: request.session['key'] = value.
* **Session fixation attack:** server sends anonymous token to browser (not logged in yet), which will elevate privileges when the client logs in. If attacker overwrites the token before logging in, he will have an elevated token after logging in.
  + Methods
    - **Tampering through network**: client visits server through HTTP, attacker can intercept and alter the HTTP traffic, and he injects into response an overwritten cookie (session token).
    - **Cross-site scripting** \*: attacker crafts malicious URL with script inside that establishes the anonymous session cookie before logging in, and then when the client clicks it and logs in (elevates the session), the attacker has a logged in session.
  + Solution/Mitigation: always issue a new session token, when elevating from anonymous token to logged-in token.

|  |
| --- |
| Does request.session.flush(). Correct.  Look for manually setting session variables without regenerating sessionId (wrong) |

\*Cross-site scriptiong (XSS) – attacker can insert javascript scripts that will be run in the victim’s browser.

- **Reflected XSS**: javascript injected into request and reflected immediately in response

**- Stored XSS**: script injected into request, stored somewhere in server and reflected repeatedly

Mitigation: sanitize input data, escape (transform “<” into “&lt;” so it is displayed as text and not executed)

## Broken access control

#### Cross site request forgery (CSRF)

Browser has the sessionID cookie stored for bank.com, and when user accesses evil.com, it can make browser execute a request to bank.com (like a money transfer), and it will be sent with the right cookie. Browser automatically sends the cookies on requests, and can’t be sure if the user intentionally made the request or not.

* **Identify vulnerability**: create HTML page separate to the testing web, include something like <img src="http://mysite.com/account/del" width="0" height="0">, log into account in the web and with the session active open the HTML page and check if account has been deleted.

|  |
| --- |
| 1.Never change data in GET requests, only POST + csrf\_token  2.Check for csrf middleware |

* **Mitigation**
  + Extra authentication
  + DSRF tokens
  + SameSite cookies: browser setting that prevents cross-site cookie usage (not let evil.com access cookies for bank.com)
  + Referer-based validation: verifies the request originates from own domain
  + Validation via CSRF token: add action token as hidden field in the body of the request to genuine forms, token that will not be sent by the browser accidentally and the attacker can’t access through evil.com

#### Server Side Request Forgery (SSRF) Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico El contenido generado por IA puede ser incorrecto.

Attacker writes URL with and embedded URL in it. The second URL would usually not be accessible for the attacker, but from the server it is possible (inside the trust zone).

* **Countermeasures**: no universal fix, highly depends on requirements
  + Whitelist and DNS resolution
  + Response handling
  + Disable unused URL schemas
  + Authentication on internal services
  + Network segregation

## Security misconfiguration

#### XML External Entities (XXE)

<!ENTITY name SYSTEM "URI"> is a way of defining an external resource to be reusable in the document, but can be vulnerable to attackers reading local files, internal services, doing SSRF or DoS with inputs like this:  
Logotipo

El contenido generado por IA puede ser incorrecto. Imagen que contiene Polígono

El contenido generado por IA puede ser incorrecto.

* **Countermeasures**:
  + Disable external entity resolution
  + Use safe XML parsers – defusedxml: fromstring(request.body).

|  |
| --- |
| Check for views that do uploads or similar (request.FILES[“upload”]) and parse it as XML without filtering or using defusedxml. |

## Software and data integrity failure

#### Insecure deserialization

**Serialization**: turning an object into a data format that can be sent over the network (so python class, for example, to JSON or XML)

**Deserialization**: converting serialized string back into an object. The problem comes when the server blindly deserializes input from the user and then trusts the resulting object. Attacker modifies the serialized data and then the deserialized object will be executed, leading to SQL injection, Remote Code Execution, Authentication bypass, Data tampering, Privilege escalation.

* **Countermeasures**:
  + Don’t deserialize untrusted input
  + Use data-only formats like JSON
  + Avoid pickle, yaml.load() on input (python libraries to serialize and deserialize data)
  + Validate input manually after deserialization
  + Use digital signatures (to check if content has been tampered with)
  + Run in low-privilege environments

|  |
| --- |
| Search for picke and yaml loads or raw use of deserialized data in code |

## Identification and authentication failure

#### Broken authentication

Authentication: process of verifying who you are by

* **Something you know**: passwords & security questions. Easy to implement and understand, but easy to crack and forget
* **Something you have**: bankID device, phone (SMS). Hard to crack, but can be broken, stolen and forged (cloned) and the strength only depends on how difficult it is to forge.
* **Something you are**: biometrics. Hard to crack and to steal, but not always accurate, privacy issues and hard to replace

Crack a password

* Vulnerable password storage: basic but vulnerable, all users compromised if hacker obtains password file.
  + Countermeasure: Hashing – compare hash of password when sent
* Dictionary attack: use words from dictionary and compute the hashes (online or offline)
  + Countermeasure:
    - **Salting** – include additional info in hash, hash password concatenated with salt (a random number) & store salt in password file. Good for online dictionary attack, but ineffective against online.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.\*Hashing salt does not make sense, because then you can’t reverse the salt and then can’t verify the password properly

* + - **Password Pepper**: like salt but all passwords have the same pepper value, which is stored in an encrypted form in another secure place. Defends better against dictionary attacks and also against offline attacks better.

Other password security techniques

* **Filtering**: guarantee strong password by setting a minimum length, requiring mixed characters and measuring the strength
* **Limiting logins**: allow 3-4 logins and lock account, inconvenient to forgetful user.
* **Aging password**: require to change passwords every once in a while (users will do workarounds and become more insecure)
* **Last login / Protective monitorin**g: notify users of suspicious login, educate users to pay attention and report possible attacks
* **One-time password**: send one-time password through SMS
* **Two-factor/two-channel authentication**: combine different ways of authentication
* **Password recovery** with URL tokens, PINs, Offline methods and security questions

CAPTCHA and reCAPTCHA: Completely Automated Public Turing Test to Tell Computers and Humans Apart. Commonly used to block bots. Machine learning is catching up.

## Security logging and monitoring failures

#### Insufficient logging and monitoring

* Not logged auditable events like logins and transactions
* Unclear/inadecuate/no log messages from warnings and errors
* Logs not monitored for suspicious activity
* Logs only stored locally
* Not effective alerting thresholds
* Unable to detect, escalate or alert for attacks in real time

## Insecure design

Clickjacking: attacker can place a layer with a URL that he wants the user to click that seems transparent by setting HTML iframe and opacity.

* Countermeasures
  + X-Frame-Options: deny – completely disables the loading of the page in a frame
  + X-Frame-Options: sameorigin only embed from the same server
  + X-Frame-Options: allow-from – embed only from whitelist

+ OWASP web testing guide

# Cryptography

Secure communication: establish a shared secret key (through face-to-face, handshake algorithms…) and then transmit the data using the shared key. Cryptography is the basis for many security mechanisms, but not the solution to all security problems.

Kerckhoff’s principle: the encryption algorithm is open, the only secret is the key, that should be chosen at random and kept secret

#### Step 1. Establish shared secret key

* **Public key encryption**: system where you don’t use the same key to encrypt and to decrypt. To encrypt, you use the public key of the receptor, and then only the receptor can decrypt it with his secret key. Mathematically speaking, this happens with a trapdoor one-way function, that is easy to go forward, but not to invert. Y = F(PK,x) easy; F^-1 (SK,y) = x only possible with SK
* **Digital signature**: bind document to author, in digital world it will depend on the content of the document.  
   Interfaz de usuario gráfica, Texto, Aplicación

  El contenido generado por IA puede ser incorrecto. Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

  El contenido generado por IA puede ser incorrecto.
* **Certificate Authority**: third party that ensures that when Bob sends his public key to Alice, it is a verified and real key
* **SSL/TLS handshake**

Interfaz de usuario gráfica, Aplicación, Word

El contenido generado por IA puede ser incorrecto.TLS 1.2 is insecure, use TLS 1.3  
  
RSA – prime factorization, standardized, simple, effective, widely used

ECDSA – limited support (y^2 = x^3 +ax + b), higher complexity, faster, shorter keys and limited support. Uses points on elliptic curves

#### Step 2. Transmit data using a shared secret key

* **Confidentiality**
  + **Substitution cipher:** letter matrix, each letter defined by letter row and letter column

Imagen que contiene Escala de tiempo

El contenido generado por IA puede ser incorrecto.

* + **Shift cipher**: associate letter with number, choose one and shift every letter by that position with wraparound. Only 26 possible keys, insecure.

Gráfico de cajas y bigotes

El contenido generado por IA puede ser incorrecto. Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

* + **The Vigenere cipher**: key is a string, shift each character by the amaunt dictated by the character of the key. Much more keyspace 26^n (n length of key), but also insecure, since you can crack it this way: Guess the length of the key (it is correct if the letters obtained from the same character of the key follow the normal frequency of letters in English) and then guess each character of the key by substituting the most used characters in the text with the most used characters in english.

(encrypt HOLA with key LUNA : (H)7+(L)11 = 18 = S; 14(O)+20 (U) = 34 mod 26 = 8 = I …  
  
Texto

El contenido generado por IA puede ser incorrecto.

* + **Interfaz de usuario gráfica, Texto

    El contenido generado por IA puede ser incorrecto.One time pad**: use random key of the same size of the plaintext used only once, with the same logic as Vigenère. Perfectly secure if used correctly, but key must be as long as plaintext and be used once. It has no integrity protection, if something is change in the cipher, it’s changed in the plain too.  
    Basically, with the XOR, the attacker can know the differences between the two original texts, and by applying some mechanisms on the similar parts they can identify both messages

Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.

* + **Stream cipher**: use short key as seed to generate a pseudo random key as long as the message to make OTP encryption
  + **Block cipher**: cut text into blocks and encrypt each using short keys. In DES (Data Encryption Standard) you have each block 64 bits, k=56 bits and 16 rounds, and its reversible round function. Applies a 48-bit key to the rightmost 32 bits to produce a 32-bit output. Then, rightmost 32 bits are swapped with leftmost 32 bits
    - ECB (Electronic code book) is wrong, it can disclose plaintext information (same blocks represented same way)
    - CBC (Cipher Block Chaining) with random initialization vector, each block XORs with the previous before ciphering.
* **Integrity**
  + **Interfaz de usuario gráfica, Texto, Aplicación

    El contenido generado por IA puede ser incorrecto.MAC** (Message Authentication Code): with message and shared secret key a MAC function can be applied to generate a tag sent with the message. The receiver then can recalculate the MAC with the same key and compare the result (if it is the same, its integrity hasn’t been modified)
  + **HMAC** - uses hash functions like SHA-256 or SHA-3
  + **Interfaz de usuario gráfica

    El contenido generado por IA puede ser incorrecto.ECBC** – Enhanced CBC mode that adds an integrity verification system.
* **Combination**
  + MAC then Encrypt (TLS)
  + Encrypt and MAC (SSH)
  + Encrypt then MAC (IPsec)

# Authorization, Authentication, Multi-Level, SSO…

## Authorization and MultiLevel Security

Access control: determines which entities (users, programs, machines) can access specific resources (files, memory, services). Initially physical in nature, access control evolved to handle multiple programs and users interacting with a shared system.

### Levels

* Application: complex and context-sensitive rules
* Middleware: enforce consistency (databases ensuring transaction integrity)
* Operating system: manage file and resource access
* Hardware: uses CPU and memory management units to restrict memory access

### Models

* **Discretionary access control (DAC)**: owner of a resource decides how it can be shared (choose to give read, write, or other access to users. Does not distinguish between user and process, so an attacker can make a trojan horse (process executing malicious programs) that can exploit the authorization of the user (weak control over information flow)
  + Access control matrix

Tabla

El contenido generado por IA puede ser incorrecto.

* + Authorization table

Interfaz de usuario gráfica, Aplicación, Tabla

El contenido generado por IA puede ser incorrecto.Tabla

El contenido generado por IA puede ser incorrecto.

* + Access control list (ACL) – used in modern OS

Tabla

El contenido generado por IA puede ser incorrecto.Diagrama

El contenido generado por IA puede ser incorrecto.

* + Capabilities

Tabla

El contenido generado por IA puede ser incorrecto.Diagrama

El contenido generado por IA puede ser incorrecto.

* **Mandatory access control (MAC)**: system enforces a security policy independent of the user’s action (in DAC users could take their own permissions), basis of regulations by central authority. The access class is assigned to each object and subject. Cumbersome administration.
  + Bell-LaPadula model – no read up and no write down. This prevents the Trojan horse attempts to copy secret data into an unclassified file (confidentiality)

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

* + Biba model: no write up and no read down. No improper modification of high-integrity objects and not contaminated due to reading and using unreliable data.  
    Interfaz de usuario gráfica, Aplicación, Teams

    El contenido generado por IA puede ser incorrecto.
  + Combination: objects and subjects have to be assigned to two access classes (one for confidentiality and one for integrity).
* **Role-based access control (RBAC**): method to manage user permissions in complex systems based on roles instead of individual entities. Easy authorization management and it maps to real-world role hierarchy. Coarse-grain access control
* **Attribute-based access control (ABAC)**: access decisions based on attributes like department, owner of assets, time, location, device type… and a policy/rule engine evaluates whether access is allowed based on these attributes. Fine-grain access control

## Authentication and Single sign-on

Without SSO: user must login separately to each domain and it stores a separate cookie in the browser, and due to same-origin policy, cookies from one domain can’t be read or reused by another domain.

With SSO: user logs into a service provider that redirects to a central identity provider, user enters credentials once and the central IdP sends back confirmation. This allows user to access multiple services without having to log in again

SSO trends:

* Moving from SOAP/XML to HTTP/JSON
* Adoption of social login (Google, Facebook)
* Standard protocols
  + OpenID Connect for authentication: ID token, logging in (SSO), Who are you?
  + OAuth 2.0 for authorization: access token, accessing APIs/data, What can you do?

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto. Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.

## Control hijacking attacks

The attacker’s goal is to take over the target machine and execute arbitrary code by hijacking application control flow. Compromise confidentiality, integrity and availability. Targets mainly C/C++ code.

Attackers use buffer overflows to hijack control, the attacker fills a buffer and overwrites nearby memory, changing the program’s behavior.

These attacks can go undetected if done early since the bug is unknown to the vendor at the beginning.

How does buffer overflow work: in the memory layout you have the Stack (function return addresses and variables) and the heap (objects and function pounters), which are working at runtime, so they are the attacker’s focus

Escala de tiempo

El contenido generado por IA puede ser incorrecto. An attacker could overwrite a variable that determines access control (not even a need to execute code)

An attacker could overwrite the return address, pointing to malicious code they injected that then will be executed

An attacker could read past a buffer and get sensitive information (heartbleed bug)

* Countermeasures:
  + Always use safe functions: UNSAFE FUNCS (strcpy, strcat, gets, strncpy…)
  + Leverage defenses in compilers
  + Check length when read/write buffer
  + Use tools to audit source code
  + Rewrite software in type-safe language
  + Type-safe languages help, since the programmer is no longer in charge of defining the size of the variable.

# Threat modeling and STRIDE

Threat modeling: a way of imagining the vast vulnerability landscape of a system and ways to attack it. Look at a system from an adversary’s perspective to anticipate attack goals.

**Why?** To understand and document a system’s threat environment, discover possible weaknesses early, how to best spend the security budget…

**When?** During the design phase after establishing security requirements and attack surface analysis, and before implementation. Continuous refinement, early and frequent analysis and aligned to the organization’s development practices.

* **McGraw**’s model emphasizes that security is not one step, but many interwoven actions throughout development, and threat modeling aligns with requirements (abuse cases) and risk analysis .
* In **Agile**, threat modeling should be integrated into each agile cycle. Project inception (high-level threats), requirements planning (threats with higher impact), sprint planning (where are the threats), sprint execution (develop, update and complete), final release planning (complete models)

**Who?** Security experts, developers and engineers, stakeholders and team members, and attackers.

* Anti-Pattern: Hero Threat modeler - don’t rely on “that one smart person” to do the threat modeling, everyone should contribute.
* Pattern: Varied viewpoints - diverse team with appropriate subject matter experts and cross-functional collaboration.

**How?** There is no single best or correct way of performing threat modelling, it is a question of trade-offs and what we want to achieve by doing it.

* Anti-Pattern: Perfect representation – better to create multiple TM representations, there is no single ideal view.

#### Threat agents

By identifying likely attackers, we can prioritize the most realistic threats and apply appropriate countermeasures

* The spooks: government intelligence services. High skills and salary.
* Cyber Warriors: military/government offensive units. High skills and salary.
* The crooks: cybercriminals. Profit-driven, efficient. Jobless, with a lot of motivation.
* The geeks: curious hackers, tinkerers. May exploit for fun or curiosity
* The terrorists: politically motivated, destructive. Unpredictable, potentially targeting infrastructure. No skills but highly motivated.
* CEO criminals: sophisticated fraud rings. Deeply strategic, targeted attacks, insiders
* The swamp: trolls with some skill and some motivation to make some fun
* The insiders: people who want to expose something, not a lot of skills or money. Mostly people who are fired.

#### Software centric threat models

Models that focus on the software being built or a system being deployed. Pattern: Systematic approach – achieve thoroughness and reproducibility by applying security and privacy knowledge in a structured manner

1. Identify critical assets
2. Decompose the system to be assessed
3. Identify possible points of attack
4. Identify threats
5. Categorise and prioritise the threats
6. Mitigate

#### STRIDE

* Spoofing: pretending to be something/someone you’re not (fake webs, emails, CSRF…)
* Tampering: modifying something you’re not supposed to modify (forms, URLs, files, databases…)
* Repudiation: claiming you didn’t do something regardless whether you did it or not (have not received, attacking logs, use someone else’s account…)
* Information disclosure: exposing info to people who are not authorized to seeit (steal files/database contents, eavesdrop network data, system/api information…)
* Denial of Service: attacks designed to prevent a system from providing service (network flooding, crashing software, making systems slow, filling storage…)
* Elevation of Privilege: program or user technically able to do things that they’re not supposed to do (XSS, buffer overflow, injection attacks, modify access control, social engineering…)

Notations

* Misuse cases: extends UML use cases, high-level negative scenarios, easy to grasp by stakeholders

Diagrama

El contenido generado por IA puede ser incorrecto.

* Attack trees: possible ways of achieving an attack goal, tree with AND/OR nodes, more technical than misuse cases. Attributes: cost, detectability, difficulty, impact, penalty, profit, probability, special skill, time  
  Diagrama

  El contenido generado por IA puede ser incorrecto.
* Bow-tie diagram: model a single unwanted event at a time, different causes/threats to unwanted events, different consequences once the event has happened, preventive/reactive controls, tradition from safety

Diagrama

El contenido generado por IA puede ser incorrecto. Diagrama

El contenido generado por IA puede ser incorrecto.

* Data Flow Diagram: understand the system, dataflow between subsystems, find attack surface and critical components, define trust/privilege boundaries.

Diagrama

El contenido generado por IA puede ser incorrecto. Diagrama

El contenido generado por IA puede ser incorrecto.

# Risk Management during development

## Risk Management Framework

Continuous and iterative process to identify, rank, track and understand software security risk as it changes over time

1. Understanding the business context: business goals (circumstances to care about and risk scales), business assets (what are you trying to protect) and stakeholders (who cares – users, regulators, attackers…)
   1. Risk dimensions and scales: likelihood and impact/consequence

Interfaz de usuario gráfica, Sitio web

El contenido generado por IA puede ser incorrecto.Tabla

El contenido generado por IA puede ser incorrecto.

1. Identify business risks: threaten directly one or more of a customer’s business goals. Usually related to data, time, money, reputation and legal. Example (too difficult to use, unavailable, failure at checks, too expensive…)
2. Identify technical risks and link them with business risks: technical risks are threats and attacks that may bring negative impacts on business (how it happens). Inputs to identify: documents, user feedback, testing and threat intelligence. Tools: misuse cases, attack trees, DFDs… (Here is where threat modeling comes into scene, that is how you identify technical risks) Examples (web crashes, network jammed by DOS attack, SQL injection…

Diagrama

El contenido generado por IA puede ser incorrecto.

1. Synthesize and prioritize risks:

Tabla

El contenido generado por IA puede ser incorrecto.Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

Tabla

El contenido generado por IA puede ser incorrecto.Tabla

El contenido generado por IA puede ser incorrecto.

1. Define the risk mitigation strategy: reducing the likelihood, the severity of impacts and derive security requirements
   1. Security requirements: statement of needed security functionality that ensures one of many security properties of software being satisfied. What you require, not how to achieve it (understandability, clarity, cohesion, testability.

Tabla

El contenido generado por IA puede ser incorrecto.

1. Carry out fixes and validate: implementation of mitigation strategies and risk-based testing (focused on security requirements)

## Risk quantification

In security we assume a hostile opponent who can cause failure at the least convenient time and in the most damaging way possible. Breaches are inevitable because defenders have to be right 100% of the time, whereas attackers only have to be right once.

CVSS (Common vulnerability scoring system): standardized way of measuring the technical severity of a vulnerability. NOT a direct risk value by itself.

* Base: constant over time and across user environments
* Threat: characteristics of a vulnerability that change over time
* Environmental: unique to a user’s environment
* Supplemental: do not modify the final score, gives additional insight.

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

|  |
| --- |
| High-level overview |

# Static Analysis and Tools for Security

* Weaknesses and vulnerabilities
  + **Weaknesses**: errors in software implementation, code, design or architecture that if left unaddressed could result in vulnerabilities
  + **Vulnerabilities**: mistake/weakness in software that can be directly used by an attacker to gain access to a system or network
  + **Exploit**: piece of software containing attack vectors that could be directly used to take advantage of a vulnerability
* What and why static code analysis
  + **What is static analysis**: passive scanning of application code without executing it. A source code security analyzer examines source code to detect and report weaknesses that can lead to security vulnerabilities (results in a report form, integrated in IDEs and CI/CD)
  + **Why**: because it can catch security defects early in the software development life cycle (already in the implementation, significant for code review)
* **Static Application Security Tools**
* How does the program analyzer work
  + Source code: Source code parsed into Abstract syntax tree transformed into Control Flow Graph and then data flow analysis over it
  + Byte code
  + Binarial
* Techniques for static code analysis
  + Model construction for source code analysis (lexical analysis, semantic analysis, control flow analysis)
  + Security knowledge
  + **Pattern matching** (report security issues if patterns are found in the code)
  + **Control flow analysis** (construct graph for if, while, for, switch, exceptions and cover all through analysis)
  + **Data flow analysis** (analyse data based on control flow graph)
    - Taint analysis (source, exit and sanitization points). If tainted data (untrusted) reaches a dangerous sink, a security issue may occur

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

El contenido generado por IA puede ser incorrecto.

* **Performance metrics**: to measure how good a SAST tool perform
  + **Soundness**: sound if never overapproximates the set of bugs in a given program
  + **Completeness**: complete if never underapproximates the set of bugs in a given program

# Penetration testing

* Steps for a successful web application test
  + Understand the application you are testing
  + Quickly check for low hanging fruit (automated scanners or tools for the clearer things)
  + Dive deep into what makes sense
* Don’t get stuck in the mindset of only checking for vulnerabilities, or relying on tools. They can provide a false sense of security if not fully understood. Easy to make stuff without knowing how it works.
* Pen testing: type of security assessment where someone simulates an attack on a system before real attackers do.
  + Don’t hack things without permission, don’t go out of scope, do “ethical” testing
  + (1) Reconnaissance: check public information, try to decompile the code, look at network traffic. Learn a lot just by observing how the application work and then start to dive into the potential weaknesses
  + Try to step on the good side, there are plenty of regulation stuff that changes
  + (2) Download code and attack against local installation, passively listening to traffic, broadcast beacons without interfering with others and report findings to the organizations
    - BAD: active burp scans against servers, test wormhole attacks, broadcast beacons that interfere and sell exploits on the dark web

# Diagrama El contenido generado por IA puede ser incorrecto.Secure coding with LLMs

* Coding assistants can:
  + Automate repetitive tasks
  + Reduce cognitive load
  + Accelerate navigation of unfamiliar languages/frameworks
  + Maintain “flow state” during development
  + Democratize coding expertise
* Potential risk categories
  + Sensitive data leaks
  + Suggesting vulnerable code
  + Overlooking security
  + Training data from sources with vulnerable code
  + Missing security context in unfamiliar data domains and has limited awareness of environmental security requirements
  + Models themselves being vulnerable to attack and manipulation
  + Developer overreliance on AI (trust too much)
* Key developer practices
  + Choose AI assistant wisely
  + Apply secure prompt engineering
    - Explicit security requirements in prompts, framework-specific security guidance, context setting and example-driven prompting
  + Add real time vulnerability detection tools
  + Embed security in development workflow
  + Human-in-the-loop validation
* Key organizational practices
  + Implement AI-aware security toolchains
  + Develop clear security standards for AI-generated code
  + Provide security training for AI tool users
  + Establish accountability frameworks for AI contributions
  + Monitor and iterate on security processes

# Privacy by design

Data privacy: right to privacy and right to decide when someone can use your own personal data, interdisciplinary. This is regulated by the General Data Protection Regulation (GDPR) + other relevant laws depending on domains. Law does not distinguish between what is processed by IT/ manually. The point is not to avoid using personal data, but to use it correctly in the correct context.

**Challenge:** to secure high quality privacy assessments, you need relevant domain expertise, else the lack of understanding will result in false safety and lack of compliance with the Personal Data Act.

Important definitions:

* Personal Data: any information related to a natural person, one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
* Processing: operation or set of operations performed on personal data, whether or not by automated means (collection, storage, alteration, use, destruction…)
* Special categories of personal data: ethnic origin, political opinions, religious beliefs, genetic data, health, sex life or orientation shall be prohibited to process, except under very specific situations
* Roles:
  + Controller: natural or legal person, authority or body which determines the purposes of the processing of personal data and the means to be used. (A hospital decides to collect and manage patient health records. The hospital decides what information is needed and how it will be used)
  + Processor: natural or legal person, authority or body which processes personal data on behalf of the controller (the hospital hires an IT company to host and maintain the patient database. That company stores and secures the data but does not decide what data is collected or how it is used)
  + Recipient: natural or legal person, authority or body to which the personal data are disclosed, whether a third party or not (the hospital sends patient data to a national health authority as part of a legal obligation).

## Data Protection Impact Assessment

Risk assessment tool required under GDPR to help organisations identify and minimize the privacy risks of their data processing activities. Formal document that many business avoid or delay because they think it’s complex, but it does not need to be long or complicated. It is mandatory when data processing is likely to result in a high risk.

**Challenge**: GDPR is cryptic for those who develop digital products, since it is hard to translate GDPR requirements into process and solution.

#### Processing of personal data (iterative process):

* **Describe the processing of personal data**
  + Identify the purpose: (subprocess) or service that uses personal data that should have business focus and should be split into smaller and more specific during the development cycle. Needs domain expertise to be properly defined

Texto

El contenido generado por IA puede ser incorrecto.

* + Identify the scope of the processing
  + Define sources, recipients and responsibilities
  + Should be understandable for all parties and gradually elaborated over several iterations.
  + Do we need a complete DPI? Validate High-risk:Texto

    El contenido generado por IA puede ser incorrecto.

Imagen que contiene Texto

El contenido generado por IA puede ser incorrecto.

* **Assess compliance with the privacy principles**: inherent part of the digital product, identify and implement the necessary technical measures to comply with each of the principles.
  + Lawfulness, fairness and transparency: compliance with privacy principles as an inherent part of the digital product. Ensure real co-determination, real transparency and predictable treatment.
    - Consent, performance of contract, compliance with legal obligation, vital interests, public interest…
  + Purpose limitation
    - Identify the most appropriate purposes that focus on end-user-oriented services
    - Verify that personal data processed is not used in any context other than what is necessary and lawful
    - Role-based access control is a good basic measure to ensure purpose limitation
    - Verify that all processing included are necessary and lawful
  + Data minimization: verify that only information that is necessary for the purpose is collected
  + Accuracy: personal data processed is correct and the result of processing is correct (also in a year or two)
  + Storage limitation: consider storage periods, storage time, exceptions, ask for analysis and statistics…
  + Information security: consider measures to ensure confidentiality, integrity and availability

Texto

El contenido generado por IA puede ser incorrecto.Texto

El contenido generado por IA puede ser incorrecto.Texto

El contenido generado por IA puede ser incorrecto.

* Assess the risk to the data subjects’ rights and freedoms: rights of data subjects are described in GDPR, including right of access, to protest, to portability… For the product: how the rights of the data subjects are fulfilled and which rights are relevant will vary.

Gráfico

El contenido generado por IA puede ser incorrecto.

* Authorization

Escala de tiempo

El contenido generado por IA puede ser incorrecto.

Accountability: data controller and data processors are responsible for ensuring that necessary privacy assessments are carried out and documented.

Interference with privacy should match necessity and proportionality

Personal data as testdata: the use of production data for testing is still processing of personal data. When necessary, justify why and ensure good practices

Who? Product Teams should be responsible for description of processing operations, assess compliance and risk assessment. In some cases they need to involve lawyers. In the lead: the product/service owner.

# Microservice security

Monolithic architecture has limited scalability, has single point of failure and you need to rebuild a lot in order to change something.

Microservice architecture: software design approach where an application is structured as a collection of small independent services that communicate with each other through well-defined APIs.

* Advantages
  + Highly maintainable and testable
  + Independently deployable
  + Organized around business capabilities
* Challenges
  + Trust between services: some services deployed might be malicious, or communication might be insecure
  + Large attack area
  + Testing
  + Low visibility: cloud infrastructure tends to be opaque and disparate, so it is hard to establish a secure communication on an opaque infrastructure.
  + Polyglot architecture: knowing or using several languages, so you need the right security expertise at every framework in the stack.
* Countermeasures:
  + Rate throttling: to defend against DoS attacks, throttle traffic flow based on configuration sending the feedback on time to the senders to warn them not to send more packets
  + Authentication and authorization:
    - At API gateway – enforce verifiable client identification at entry points, control access by providing authorization policies, throttling request traffic
    - From API to microservices – external entity identity propagation
    - Between microservices – mTLS (each microservice has to carry a private/public key pair and use that pair to authenticate to the recipient microservices), token-based (caller microservice obtains a signed token by invoking a special security token service using its own service ID and pwd)
    - At microservices – service level authorization: gives each microservice more control to enforce access control policies
* HSM Bootstrapping: hardware security module to defend against attacks targeting hardware, guarantee confidentiality and integrity of execution environments
* Patterns
  + Circuit breaker: service failure protection and handle so it does not propagate, real-time monitoring and alerting, default behavior when failure
  + Command Query Responsibility Segregation: separates read and update operation to optimize performance, scalability and security
  + Strangler: transform, co-exist and eliminate, use when migrating.
  + Phantom token: combination of opaque and JWT tokens. Client retrieves opaque token (random string), client forwards token in requests to API, reverse proxy looks up JWT token and reverse token replaces the opaque with JTW in the actual request to the microservice
  + Sidecar proxy: attached to parent application, co-locate a cohesive set of tasks with the primary application but place them inside their process or container