



# DESIGN-GUARD

# SSDLC SECURITY

# ASSISTANT

## **PREPARED FOR**

Final Project

## **PREPARED BY**

Agry Zharfa

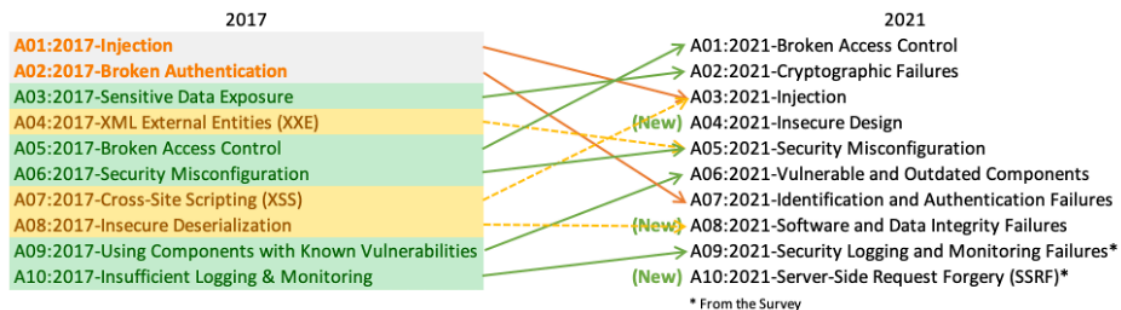
III RPLK

Rekayasa Kriptografi



# Latar Belakang

Keamanan perangkat lunak telah menjadi salah satu isu paling krusial dalam pengembangan sistem modern. Banyak serangan siber yang berhasil dieksekusi bukan karena lemahnya implementasi teknis semata, melainkan akibat desain sistem yang tidak aman sejak awal (insecure design).



Metodologi Secure Software Development Life Cycle (SSDLC) hadir untuk memastikan aspek keamanan terintegrasi dalam setiap fase pembangunan perangkat lunak. Namun, penerapan SSDLC masih menghadapi tantangan:

1. Kurangnya kesadaran tim developer terhadap praktik keamanan.
2. Sulitnya mengintegrasikan referensi standar (seperti OWASP ASVS, CWE, CISA KEV) dalam evaluasi desain.
3. Keterbatasan waktu dan sumber daya dalam melakukan threat modeling secara manual.

Di sisi lain, perkembangan Large Language Models (LLM) dan pendekatan Agentic AI memungkinkan pembuatan asisten cerdas yang dapat:

1. Memahami dokumen desain sistem.
2. Mengevaluasi potensi risiko keamanan.
3. Memberikan rekomendasi mitigasi berbasis pengetahuan terkini.

Dengan demikian, dibutuhkan sebuah sistem agentic berbasis LLM yang mampu menggabungkan RAG (Retrieval-Augmented Generation), multi-agent collaboration, serta integrasi ke sumber data keamanan (CVE, CWE, KEV, OWASP ASVS) untuk membantu analisis keamanan desain perangkat lunak.



# Tujuan Proyek

1. Mengembangkan sebuah **asisten keamanan agentic** yang mampu melakukan evaluasi risiko dalam fase desain SSDLC.
2. Mengintegrasikan **vector database** berisi dokumen keamanan standar (ASVS, CWE, KEV, dsb).
3. Menyediakan mekanisme **retrieval hybrid** dimana jika data tidak ditemukan di vector DB, sistem akan melakukan fallback ke **DuckDuckGo Web Search**.
4. Menghasilkan **laporan ancaman dan rekomendasi mitigasi** yang mudah dipahami developer.
5. Menyediakan **visualisasi threat model** menggunakan diagram interaktif (Mermaid + Zoom/Pan).



## Feature Security Profiler



### System Context

Tipe Aplikasi:

Web Application



Environment Deployment:

Cloud (AWS/Azure/GCP)



Tingkat Kritikalitas Bisnis:

Low (Internal tools)



Expected User Base:

<100 internal



### Core Functionality

#### Authentication & Identity Management

Select applicable authentication features:

Password Reset... x

Session Manage... x



#### Data Processing & Input

Select data processing features:

User Form Input x

File Upload x



Filetypes allowed:

Choose options





# Feature Security Profiler



## System Context

Tipe Aplikasi:

Web Application



Environment Deployment:

Cloud (AWS/Azure/GCP)



Tingkat Kritikalitas Bisnis:

Low (Internal tools)



Expected User Base:

<100 internal



## Core Functionality

### 🔒 Authentication & Identity Management

Select applicable authentication features:

Multi-Factor Aut... x



### 📊 Data Processing & Input

Select data processing features:

File Upload x

Data Import/Exp... x

User Form Input x



File types allowed:

Choose options



### 🔗 External Integration

External integrations:

Payment Proces... x





## Security Context

Data Sensitivity Level:

Public



Existing Security Controls:

Input Validation



Compliance Requirements:

None



## Security References

Additional Context and References:



Describe specific security concerns including CWE, CAPEC, and ASVS references. Example: Must address CWE-79 (XSS), CWE-89 (SQL injection), ASVS-V5.3 (Output Encoding), etc.



For better performance, include specific references in Additional Context (CWE-xx, CAPEC-xx, ASVS-Vx.x.x)



Generate Security Analysis



## SECURITY REQUIREMENTS CHECKLIST

Berdasarkan OWASP ASVS dan Proactive Controls, identifikasi MINIMAL 5 persyaratan keamanan yang paling relevan:

- **ASVS-V2.1.1 - Authentication:** Implementasi autentikasi yang kuat, termasuk penggunaan algoritma hashing password yang aman (seperti bcrypt, Argon2, atau PBKDF2) dengan panjang salt dan iterasi yang cukup. Mencegah serangan brute-force dengan menerapkan mekanisme penguncian akun setelah beberapa upaya login yang gagal. Memastikan sesi pengguna terenkripsi dan dilindungi dari serangan session hijacking. Penggunaan multi-faktor autentikasi (MFA) sangat direkomendasikan.
- **ASVS-V4.1.3 - Session Management:** Menggunakan session ID yang acak dan unik untuk setiap sesi pengguna. Menerapkan mekanisme secure cookie dengan flags `HttpOnly` dan `Secure`. Mengatur timeout session yang sesuai dengan kebijakan keamanan. Memastikan session ID tidak dapat ditebak atau dimanipulasi. Implementasi mekanisme logout yang aman untuk menghapus session pengguna secara efektif.
- **ASVS-V5.2.1 - Input Validation:** Melakukan validasi input yang ketat pada semua data yang diterima dari pengguna, termasuk validasi tipe data, panjang, format, dan karakter khusus. Menerapkan mekanisme escaping atau sanitizing untuk mencegah serangan injection (SQL injection, XSS, dll.). Menggunakan whitelist daripada blacklist untuk memvalidasi input. Menangani error handling dengan aman untuk mencegah informasi sensitif bocor.
- **ASVS-V7.1.2 - File Upload:** Membatasi tipe file yang diizinkan untuk diunggah hanya pada tipe yang aman (misalnya, gambar, dokumen). Melakukan validasi ukuran file untuk mencegah serangan denial-of-service (DoS). Memeriksa konten file untuk memastikan tidak mengandung kode berbahaya sebelum menyimpannya. Menyimpan file di luar direktori webroot untuk mencegah akses langsung. Melakukan sanitizing nama file untuk mencegah serangan path traversal.
- **ASVS-V14.2.1 - Security Logging and Monitoring:** Merekam semua aktivitas penting, termasuk login, logout, upaya login yang gagal, dan aktivitas yang mencurigakan. Menggunakan sistem logging yang terpusat dan terenkripsi. Menerapkan sistem monitoring untuk mendeteksi dan merespon ancaman secara real-time. Melakukan analisis log secara berkala untuk mengidentifikasi pola serangan dan kelemahan keamanan.

## THREAT SCENARIOS & ATTACK VECTORS

Berdasarkan CWE, CAPEC, dan OWASP Top 10, identifikasi MINIMAL 3 skenario ancaman yang BERBEDA dan spesifik:

### Threat 1: CWE-79 - Cross-Site Scripting (XSS)

- **Attack Vector:** CAPEC-63 - Cross-site scripting (XSS) attacks involve injecting malicious scripts into websites or web applications. An attacker could inject JavaScript code into user input fields (e.g., user profile, comments) that is then rendered on the website without proper sanitization. This allows the attacker to steal cookies, redirect users to phishing sites, or deface the website.
- **Impact:** High - Data kebocoran (cookie, informasi pengguna), pencurian sesi, pengalihan ke situs phishing, defacement website.
- **Likelihood:** Medium - Kemungkinan cukup tinggi jika validasi input tidak diimplementasikan dengan benar.
- **Affected Component:** User Form Input

### Threat 2: CWE-89 - SQL Injection

- **Attack Vector:** CAPEC-66 - SQL injection attacks involve injecting malicious SQL code into user input fields that are then used to query a database. An attacker could inject SQL code into a login form or search query to bypass authentication, retrieve sensitive data, or modify database records.
- **Impact:** High - Akses tidak sah ke data sensitif, modifikasi data, penolakan layanan (DoS).
- **Likelihood:** Medium - Kemungkinan cukup tinggi jika validasi input tidak diimplementasikan dengan benar pada fitur yang berinteraksi dengan database.
- **Affected Component:** User Form Input

### Threat 3: CWE-287 - Improper Authentication

- **Attack Vector:** CAPEC-114 - Credential stuffing attacks involve using stolen credentials from other websites to attempt to log into the target application. An attacker could obtain a list of usernames and passwords from a data breach on another website and use them to attempt to log into the application. Weak password policies also increase the likelihood of this attack.
- **Impact:** High - Akses tidak sah ke akun pengguna, pencurian data, penyalahgunaan fungsi aplikasi.
- **Likelihood:** Medium - Kemungkinan cukup tinggi jika kebijakan password lemah atau tidak ada mekanisme untuk mendeteksi dan memblokir upaya login yang gagal berulang kali.
- **Affected Component:** Username/Password Login

## SECURITY CONTROLS & COUNTERMEASURES

Berdasarkan NIST CSF dan OWASP Defense in Depth:

### Technical Controls

1. **CWE-79/ASVS-V5.2.2 - Output Encoding:** Menggunakan library encoding yang sesuai untuk encoding output yang dihasilkan dari aplikasi, khususnya pada data yang berasal dari input pengguna. Contohnya, menggunakan `htmlspecialchars()` di PHP atau library encoding yang setara di bahasa pemrograman lain untuk mencegah XSS.
2. **CWE-89/ASVS-V5.3.4 - Parameterised Queries:** Menggunakan parameterised queries atau prepared statements untuk mencegah SQL injection. Ini memastikan bahwa data pengguna tidak diinterpretasikan sebagai kode SQL, melainkan sebagai parameter dalam query. Hindari penggunaan string concatenation untuk membangun query SQL.
3. **CWE-287/ASVS-V2.2.1 - Rate Limiting:** Menerapkan rate limiting untuk membatasi jumlah upaya login yang gagal dalam jangka waktu tertentu. Ini membantu mencegah serangan brute-force dan credential stuffing. Setelah melewati batas upaya login yang gagal, akun pengguna akan diblokir sementara.

### Administrative Controls

1. **ASVS-V14.1.1 - Security Policy:** Membuat dan menerapkan kebijakan keamanan yang komprehensif yang mencakup persyaratan autentikasi, otorisasi, manajemen session, dan penanganan insiden keamanan. Kebijakan ini harus dikomunikasikan dan dipahami oleh semua pengguna dan pengembang.
2. **ASVS-V13.2.3 - Security Awareness Training:** Memberikan pelatihan keamanan secara berkala kepada pengguna dan pengembang untuk meningkatkan kesadaran akan ancaman keamanan dan cara untuk mencegahnya. Pelatihan ini harus mencakup topik seperti XSS, SQL injection, dan serangan lainnya yang relevan.

## IMPLEMENTATION CHECKLIST

- **Pre-Development:** Melakukan analisis risiko dan threat modeling untuk mengidentifikasi potensi kerentanan keamanan. Memilih framework dan library yang aman dan teruji. Menentukan standar coding yang aman.
- **During Development:** Menerapkan teknik secure coding practices, termasuk validasi input, output encoding, dan parameterised queries. Melakukan code review secara berkala untuk mendeteksi potensi kerentanan. Menggunakan static dan dynamic application security testing (SAST/DAST) tools.
- **Testing Phase:** Melakukan pengujian keamanan yang komprehensif, termasuk pengujian penetrasi, pengujian keamanan fungsional, dan pengujian keamanan unit. Menggunakan tools pengujian keamanan otomatis dan manual.
- **Deployment:** Menggunakan infrastruktur cloud yang aman dan terkonfigurasi dengan benar. Menerapkan firewall dan sistem deteksi intrusi (IDS/IPS). Memantau aktivitas sistem secara berkala.
- **Maintenance:** Melakukan pembaruan dan patching secara rutin untuk memperbaiki kerentanan keamanan yang ditemukan. Memantau log keamanan untuk mendeteksi aktivitas yang mencurigakan. Melakukan audit keamanan secara berkala.

## RISK ASSESSMENT SUMMARY

- **Critical Risk:** CWE-89 (SQL Injection) - Akses tidak sah ke data sensitif dan potensi kerusakan database.
- **Recommended Priority:** 1. CWE-89, 2. CWE-79, 3. CWE-287
- **Quick Wins:** Implementasi rate limiting untuk membatasi upaya login yang gagal dan penggunaan parameterised queries untuk mencegah SQL injection.

CWE References ⓘ

10


CAPEC References ⓘ

3


ASVS References ⓘ

10

>  Knowledge Base References

 Export Analysis

 New Analysis

 All required references present



# Struktur Sistem

Secara garis besar, sistem ini terdiri dari beberapa komponen utama..

## 1. Vector Database

Security Requirements DB → berisi dokumen OWASP ASVS, Proactive Controls.

Threat Patterns DB → berisi CWE, CAPEC, KEV, MITRE ATT&CK.

## 2. Hybrid Retrieval

Internal RAG → query ke vector DB.

Fallback Web Search → jika tidak ada hasil, gunakan DuckDuckGo (via LangChain connector).



# Integrasi dan Output

Sistem ini dapat diperluas dengan berbagai integrasi:

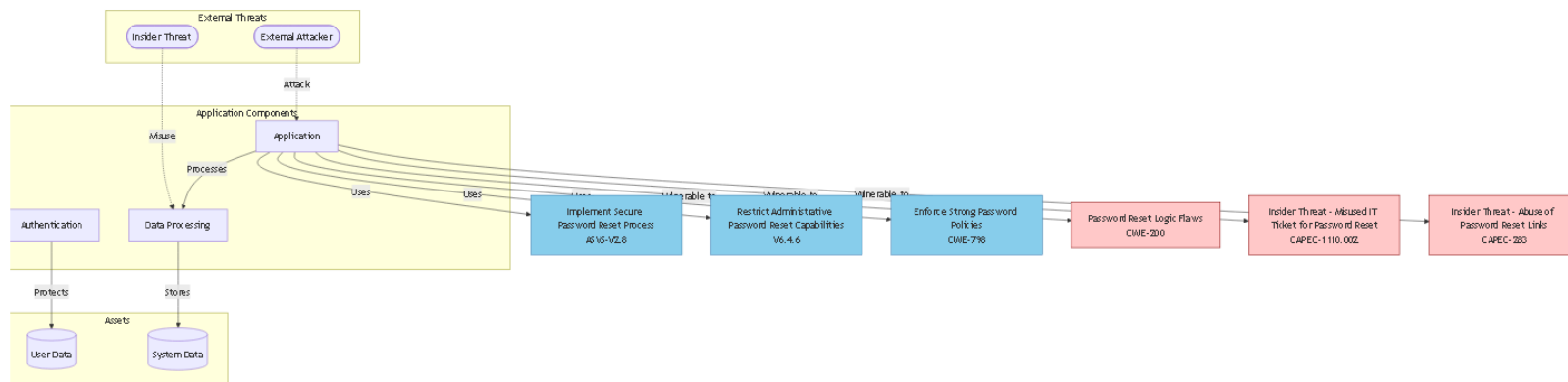
1. LLM API: Groq (fast inference), Gemini (reasoning multimodal), Ollama (local model).
2. Threat Intelligence: CISA KEV, NVD CVE
3. Search Connector: DuckDuckGo Web Search (fallback real-time retrieval).

## Output Sistem

1. Laporan Evaluasi Keamanan (teks & JSON).
2. Diagram Threat Model Interaktif

### Threat Model Visualization

#### Interactive Threat Model



### Key Threats Identified

#### High Risk Threats: 11

- Password Reset Logic Flaws (CWE-200)
- Weak Session ID Generation (CWE-863)
- Weak Password Policy (CWE-798)

#### Medium Risk Threats: 16

- Insider Threat - Misused IT Ticket for Password Reset (CAPEC-1110.002)
- Insider Threat - Abuse of Password Reset Links (CAPEC-283)
- Lack of Password Complexity Enforcement (CWE-798)

- 3. Rekomendasi Mitigasi yang berbasis standar (OWASP ASVS, CWE, dsb).
- 4. Trace Agentic Workflow → transparansi reasoning setiap agent.

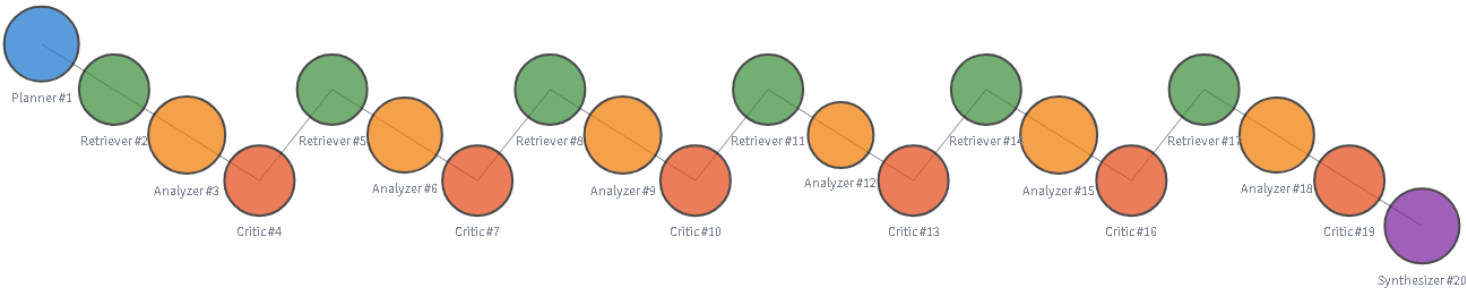
Agent Actions Trace

|    | id       | role        | action_type  | reasoning  | confidence | duration | timestamp                |
|----|----------|-------------|--|--|------------|----------|--------------------------|
| 0  | 0eac3adf | planner     | security_analysis_planning                                   | Decomposed security analysis into specialized domain tasks                 | 0.900000   | 7.43s    | 2025-01-24T08:00:00.000Z |
| 1  | 38fd6f7d | retriever   | information_retrieval_Authentication and Session Management  | Retrieved 11 relevant documents for Authentication and Session Management  | 0.800000   | 4.50s    | 2025-01-24T08:00:00.000Z |
| 2  | 61b24866 | analyzer    | security_analysis_Authentication and Session Management      | Analyzed security risks for Authentication and Session Management domain   | 0.950000   | 6.06s    | 2025-01-24T08:00:00.000Z |
| 3  | d1093b46 | critic      | analysis_review_Authentication and Session Management        | Reviewed analysis quality for Authentication and Session Management        | 0.800000   | 3.51s    | 2025-01-24T08:00:00.000Z |
| 4  | 98443a18 | retriever   | information_retrieval_Authentication and Session Management  | Retrieved 35 relevant documents for Authentication and Session Management  | 0.800000   | 3.06s    | 2025-01-24T08:00:00.000Z |
| 5  | f0cc0d4a | analyzer    | security_analysis_Authentication and Session Management      | Analyzed security risks for Authentication and Session Management domain   | 0.900000   | 6.54s    | 2025-01-24T08:00:00.000Z |
| 6  | 51239877 | critic      | analysis_review_Authentication and Session Management        | Reviewed analysis quality for Authentication and Session Management        | 0.800000   | 4.36s    | 2025-01-24T08:00:00.000Z |
| 7  | 0bebeb34 | retriever   | information_retrieval_Authentication and Session Management  | Retrieved 22 relevant documents for Authentication and Session Management  | 0.800000   | 2.81s    | 2025-01-24T08:00:00.000Z |
| 8  | f4f3c28f | analyzer    | security_analysis_Authentication and Session Management      | Analyzed security risks for Authentication and Session Management domain   | 0.950000   | 66.17s   | 2025-01-24T08:00:00.000Z |
| 9  | f4e9d628 | critic      | analysis_review_Authentication and Session Management        | Reviewed analysis quality for Authentication and Session Management        | 0.800000   | 4.41s    | 2025-01-24T08:00:00.000Z |
| 10 | 41ca9bb3 | retriever   | information_retrieval_Input Validation and Data Sanitization | Retrieved 46 relevant documents for Input Validation and Data Sanitization | 0.800000   | 3.28s    | 2025-01-24T08:00:00.000Z |
| 11 | 601c0294 | analyzer    | security_analysis_Input Validation and Data Sanitization     | Analyzed security risks for Input Validation and Data Sanitization domain  | 0.700000   | 8.50s    | 2025-01-24T08:00:00.000Z |
| 12 | 313269b3 | critic      | analysis_review_Input Validation and Data Sanitization       | Reviewed analysis quality for Input Validation and Data Sanitization       | 0.800000   | 3.77s    | 2025-01-24T08:00:00.000Z |
| 13 | 7cda6cfd | retriever   | information_retrieval_Input Validation and Data Sanitization | Retrieved 27 relevant documents for Input Validation and Data Sanitization | 0.800000   | 2.97s    | 2025-01-24T08:00:00.000Z |
| 14 | 66af73ae | analyzer    | security_analysis_Input Validation and Data Sanitization     | Analyzed security risks for Input Validation and Data Sanitization domain  | 0.950000   | 7.76s    | 2025-01-24T08:00:00.000Z |
| 15 | d41018a3 | critic      | analysis_review_Input Validation and Data Sanitization       | Reviewed analysis quality for Input Validation and Data Sanitization       | 0.800000   | 5.06s    | 2025-01-24T08:00:00.000Z |
| 16 | 5eef8e02 | retriever   | information_retrieval_Input Validation and Data Sanitization | Retrieved 16 relevant documents for Input Validation and Data Sanitization | 0.800000   | 2.81s    | 2025-01-24T08:00:00.000Z |
| 17 | a1b09344 | analyzer    | security_analysis_Input Validation and Data Sanitization     | Analyzed security risks for Input Validation and Data Sanitization domain  | 0.900000   | 7.57s    | 2025-01-24T08:00:00.000Z |
| 18 | 05cf260e | critic      | analysis_review_Input Validation and Data Sanitization       | Reviewed analysis quality for Input Validation and Data Sanitization       | 0.800000   | 4.08s    | 2025-01-24T08:00:00.000Z |
| 19 | 41088f7c | synthesizer | final_report_synthesis                                       | Synthesized final report from 2 domain analyses                            | 0.900000   | 69.88s   | 2025-01-24T08:00:00.000Z |


Threat Model Agent Confidence Workflow Graph

📷 🔍 + 📄 🗑️ 🔄


Agentic Workflow (Interactive)



# Evaluasi

 **Optimized Results**

|                  |                       |            |             |
|------------------|-----------------------|------------|-------------|
| Avg Faithfulness | Avg Context Precision | Avg CWE F1 | Avg ASVS F1 |
| 0.990            | 0.955                 | 0.983      | 0.987       |
| ↑ +0.395         | ↑ +0.788              | ↑ +0.664   | ↑ +0.447    |
| Avg Relevancy    | Avg Context Recall    |            |             |
| 0.919            | 0.452                 |            |             |
| ↑ +0.261         | ↑ +0.156              |            |             |

 **Optimized Results**

|                  |                       |            |             |
|------------------|-----------------------|------------|-------------|
| Avg Faithfulness | Avg Context Precision | Avg CWE F1 | Avg ASVS F1 |
| 0.913            | 0.955                 | 0.985      | 0.988       |
| ↑ +0.318         | ↑ +0.788              | ↑ +0.665   | ↑ +0.448    |
| Avg Relevancy    | Avg Context Recall    |            |             |
| 0.911            | 0.452                 |            |             |
| ↑ +0.253         | ↑ +0.156              |            |             |

🏆 Excellent Performance! Overall: 0.867