**Procesbeskrivelse: Integrated Secure AI System Development Life Cycle (SAI-SDLC)**

**Version:** 2.0

**Ejer:** Account Security Manager (ASM) / Compliance-sporet

**Compliance:** ISO 9001, ISO 27001:2022, ISO 42001 (AIMS)

**Standarder:** SEI CERT Java, OWASP ASVS, FAIR Risk Management

**Revision:** ITGC, GDPR, NSIS, DNV

**1. Governance & Den Tre-strengede Styringsmodel**

For at sikre overensstemmelse med ISO 9001 (kvalitet) og ISO 42001 (AI-etik), er projektet organiseret i tre spor med forskudt kadance og uafhængig kontrol.

**1.1 Strategisk Forretningsudvikling (3 måneders horisont)**

* **Aktivitet:** Vurdering af AI-use cases jf. KOMBITs principper.
* **Risikostyring:** Gennemførsel af **FAIR-analyser** på Epics for at kvantificere økonomisk og etisk risiko ved AI-tiltag.
* **Audit-output:** Strategisk risikolog til brug for årlig revisionserklæring.

**1.2 Teknisk Arkitektur- og Kravspor (6 ugers horisont)**

* **Aktivitet:** Design af "Security by Design" løsninger.
* **Værktøjer:** Trusselsmodellering i **OWASP Threat Dragon** og kravstillelse via **OWASP ASVS** (Level 2/3).
* **AI-krav:** Definition af forklarbarhedsmetoder (**SHAP/LIME**) og registrering af datakilder i **Open Metadata**.

**1.3 Operativt Udviklingsspor (Scrum/Kanban)**

* **Aktivitet:** Implementering af sikker kode i 2-ugers sprints.
* **Security Champions:** Hvert team har en udpeget Champion (OWASP v10), der har 20% dedikeret tid til at sikre, at ISO-kravene efterleves i den daglige kodning.

**2. Den Tekniske Pipeline (DevSecOps i Azure DevOps)**

Vi anvender en automatiseret pipeline til at håndhæve vores standarder. Pipelinen fungerer som en uafhængig kontrolinstans for DNV-stikprøver.

**2.1 Sikker Kodestandard (Java)**

Vi følger **SEI CERT Oracle Coding Standard for Java** som teknisk fundament og **OWASP Secure Coding Practices** for arkitektonisk sikkerhed.

* **Håndhævelse:** **SonarQube Enterprise** er konfigureret med CERT Java-profilen som en "Hard Gate".
* **Logging:** Implementering af **OWASP Logging Cheat Sheet** principper (strukturerede logs uden PII).

**2.2 Software Supply Chain (SBOM)**

For hver build genereres en **CycloneDX SBOM**, der dokumenterer alle 3. parts biblioteker og AI-modeller. Dette danner grundlag for vores ITGC- og GDPR-dokumentation.

**2.3 AI-Compliance & XAI**

Ingen AI-model releases uden:

* **Model Card:** Dokumentation af modellens formål og bias-test (ISO 42001 A.10).
* **XAI-Validering:** Teknisk bevis (SHAP/LIME) på modellens beslutningsgrundlag (KOMBIT Princip 6).

**3. Compliance, Overvågning & Revision**

Dette spor sikrer den løbende effektivitet af vores kontroller (PDCA).

**3.1 Kontrolmiljø**

* **Månedlig Rapportering:** Account Security Manager (ASM) leverer en uafhængig sikkerhedsrapport til kunden baseret på data fra SonarQube, Azure Boards og FAIR-analyser.
* **Sikkerhedsretrospektiver:** Kvartalsvise dybdegående gennemgange faciliteret af Compliance-sporet for at justere processer og tjeklister.

**3.2 Revisions- og Certificeringsberedskab**

* **ITGC/GDPR/NSIS:** Det eksterne revisionshus auditerer de digitale "spor" i Azure DevOps (PR-godkendelser, build-logs og risiko-links).
* **DNV (ISO 27001):** Systemet er designet til stikprøvekontrol. Enhver kodelinje kan spores direkte tilbage til en trusselsmodel i Threat Dragon og et arkitekturkrav.

**4. Definition of Done (DoD) - Governance Checkliste**

For at en opgave kan markeres som "Done" i Azure Boards, skal følgende være verificeret:

1. **Sikkerhed:** SonarQube (SEI CERT Java) og CycloneDX (SCA) er "Green".
2. **Arkitektur:** Opgaven overholder den trusselsmodel, der blev lavet i 6-ugers sporet.
3. **Data:** Metadata er opdateret i Open Metadata, og PII-maskering er verificeret.
4. **AI:** Model Card er opdateret, og SHAP-analysen viser ingen ulovlig bias.
5. **Review:** Security Champion har foretaget peer-review med fokus på sikkerhed.

**5. Visualisering af Processen i Azure DevOps**

| Niveau | Kadance | Ansvarlig | Primært Værktøj |
| --- | --- | --- | --- |
| **Epics** | 3 måneder | Forretningsudvikling | FAIR-analyse |
| **Features** | 6 uger | Teknisk Arkitektur | Threat Dragon / ASVS |
| **Stories** | 2 uger | Scrum Team / SC | SonarQube / CERT Java |
| **Overvågning** | Løbende | Compliance / ASM | Dashboards / Revision |

| **SDLC Fase** | **Aktivitet** | **ISO 9001 (Kvalitet)** | **ISO 27001 (Sikkerhed)** | **ISO 42001 (AI)** |
| --- | --- | --- | --- | --- |
| **Krav** | Backlog Refinement | Krav-sporing (Work Items) | FAIR Risikovurdering | AI Use-case vurdering |
| **Design** | Design Spike | Design Review | **Threat Dragon** | Etisk screening |
| **Kodning** | Pull Request | Peer-review | **SonarQube** scan | Bias-tjek af data |
| **Test** | QA / Verifikation | Test-dækning | **OWASP ZAP** scan | **SHAP/LIME** validering |
| **Release** | Definition of Done | Godkendelses-log | **CycloneDX** SBOM | **Model Card** publicering |

| **Fase** | **Azure DevOps Værktøj** | **Sikkerhedsaktivitet** | **Metodik (Sources)** |
| --- | --- | --- | --- |
| **Plan** | Azure Boards | Mini-FAIR vurdering af nye features. | **FAIR-lite** |
| **Design** | Azure Repos / Wiki | Threat Modeling & Arkitektur-review. | **Threat Dragon & Secure by Design** |
| **Code** | Azure Repos / Sonar | Statisk analyse & Secrets scanning. | **SonarQube & Gitleaks** |
| **Build** | Azure Pipelines | Supply Chain scanning (SBOM). | **CycloneDX** |
| **Test** | Azure Pipelines | Dynamisk test af kørende applikation. | **OWASP ZAP** |
| **Release** | Azure Artifacts | Validering af S-BOM og K8s manifester. | **K8s Scan / Checkov** |

| **Annex A Kontrol-gruppe** | **Specifik Kontrol (Eksempler)** | **Placering i jeres SDLC** | **Implementering / Værktøj** |
| --- | --- | --- | --- |
| **A.2: AI Policy** | A.2.1 AI Policy | **Forretningsspor** | Overordnet AI-strategi mappet til KOMBIT principper. |
| **A.3: Organization** | A.3.1 Roles & Responsibilities | **Governance** | Definition af **ASM**, **Security Champions** og AI-Board. |
| **A.4: Resources** | A.4.2 Data Resources | **Teknisk spor (6 uger)** | Sikring af datakvalitet og adgang via **Open Metadata**. |
| **A.5: Impact Assessment** | A.5.1 AI Impact Assessment | **Forretningsspor (3 mdr)** | **FAIR-analyse** og etisk screening af nye Epics. |
| **A.6: AI Life Cycle** | A.6.1 Life Cycle Management | **Hele SDLC** | Selve jeres SAI-SDLC manual (det dokument vi skrev). |
| **A.7: Design & Dev.** | A.7.2 Security & Privacy | **Udvikling (Sprints)** | **SonarQube**, **ASVS** og **SEI CERT Java**. |
| **A.8: Data for AI** | A.8.2 Data Provenance | **Teknisk spor / Pipeline** | **CycloneDX (ML-BOM)** og metadata-tags på træningsdata. |
| **A.9: External Parties** | A.9.1 Supplier Relationships | **Compliance-spor** | Scanning af 3. parts AI-biblioteker via **CycloneDX**. |
| **A.10: Transparency** | A.10.2 Documentation | **Release-fase (DoD)** | Obligatorisk generering af **Model Cards**. |
| **A.10: Explainability** | A.10.3 Explainability | **Udvikling / Test** | Teknisk validering via **SHAP / LIME**. |
| **A.11: Monitoring** | A.11.1 Continuous Monitoring | **Compliance / Drift** | Overvågning af **Model Drift** og KPI'er (PDCA). |

#### 1. A.4.1 Relevant Expertise (Human Resources)

* **Udfordring:** Hvordan dokumenterer I, at dem der arbejder med AI, faktisk ved hvad de laver?
* **Løsning:** Implementer en **træningslog** i jeres Compliance-spor. Dokumentér at jeres **Security Champions** har gennemgået træning i OWASP AI Exchange og ISO 42001.

#### 2. A.11.2 Event Logging for AI Systems

* **Udfordring:** Dette går ud over standard system-logs. Det kræver logning af modellens beslutninger og de input, der førte til dem (uden at bryde GDPR).
* **Løsning:** Udvid jeres **Java Logging Standard** til at inkludere "Inference Logs", som gemmer input-features og modellens konfidensscore i et beskyttet audit-spor.

#### 3. A.5.3 Stakeholder Communication

* **Udfordring:** Hvordan informeres borgere eller sagsbehandlere om, at en AI er involveret?
* **Løsning:** Dette skal adresseres i jeres **Forretningsspor**. Der skal ligge et krav om "Brugerinformation" i jeres Epics, som sikrer, at UI-designet inkluderer oplysning om AI-brug (KOMBIT Princip 14).

#### 4. A.11.4 Response & Remediation

* **Udfordring:** Hvad gør I, hvis modellen pludselig begynder at udvise bias i produktionen?
* **Løsning:** Opret en **"AI Incident Response Plan"**. Den skal beskrive, hvordan man hurtigt "ruller tilbage" til en tidligere version af modellen eller deaktiverer AI-funktionen i K8s, hvis en risiko materialiserer sig.

### 1. Specialiseret Work Item Tracking (WIT) for AI

For at gøre AI-sikkerhed synlig og uomgængelig, bør I udnytte ADO's mulighed for at tilpasse arbejdsprocesser.

* **Custom Work Item Types:** Opret en ny type Work Item kaldet **"AI Governance Task"**. Dette gør det muligt at køre rapporter specifikt på AI-modenhed.
* **Zero-Tolerance Policy (Quality Gate):** Implementer en regel i jeres **Release Pipeline**, der automatisk fejler eller stopper en release, hvis der findes åbne Work Items med tagget #AI-Security eller #High-Risk.
* **AI-Track Tags:** Brug specifikke tags til at styre det "særlige spor":
  + #XAI (for SHAP/LIME opgaver)
  + #ModelCard (for dokumentation)
  + #TrainingData (for dataproveniens)

### 2. Det særlige AI-spor: Fra Træningsdata til Explainability

Her er de fire komponenter, du nævner, omsat til konkrete procedurer i jeres ADO-backlog:

#### A. Håndtering af Træningsdata (Data Governance)

* **WI-krav:** Hver model skal have en tilknyttet opgave, der dokumenterer datakilderne i **Open Metadata**.
* **Handling:** Verificering af, at træningsdata er anonymiseret jf. jeres OIO-standarder, før de fodres til modellen.

#### B. Model Cards (Gennemsigtighed)

* **WI-krav:** En "Definition of Done" for en AI-feature er produktionen af et **Model Card**.
* **Indhold:** Modellens begrænsninger, tiltænkte brug, bias-tests og ydeevne (jf. **KOMBIT Princip 6 og 14**).

#### C. SHAP & LIME (Explainable AI - XAI)

* **WI-krav:** For hver model, der påvirker en borgersagsbehandling, skal der implementeres en **XAI-komponent**.
* **Handling:** Udvikleren skal dokumentere, at SHAP eller LIME værdierne er tilgængelige for sagsbehandleren, så de kan se, *hvilke* faktorer (fx alder, indkomst, historik) der har vægtet højest i en anbefaling. Dette understøtter **KOMBIT Princip 9 (Menneskelig kontrol)**.

### 3. Eksempel: FAIR-risikovurdering i en Design Spike

Som du efterspurgte, er her et bud på, hvordan en **FAIR-analyse** integreres i dette flow.

**Scenarie:** *"Risiko for at AI-modellen diskriminerer en bestemt gruppe borgere pga. bias i træningsdata."*

1. **Loss Event Frequency (Hvor tit sker det?):**
   * **Threat Event Frequency:** Hvor ofte kører modellen? (Høj - hver gang en sagsbehandler åbner en sag).
   * **Vulnerability:** Har vi testet for bias? Hvis WI for bias-test i ADO er "Håndteret", er sårbarheden lav.
2. **Loss Magnitude (Hvad koster det?):**
   * **Primary Loss:** Tid brugt på at omstøde forkerte afgørelser.
   * **Secondary Loss:** Juridiske bøder (GDPR/EU AI Act) og massivt tillidstab hos kommunerne (tabt omdømme jf. **KOMBIT Princip 14**).
3. **Resultat:** FAIR-analysen giver et kronebeløb på risikoen. Hvis beløbet er højere end omkostningen ved at implementere **SHAP/LIME**, så bliver XAI-opgaven i ADO prioriteret som en "Blocker".

### 4. Sammenhængen: Din nye procesbeskrivelse

Ved at kombinere disse elementer kan du præsentere kunden for følgende flow:

"Vores AI-udvikling kører i et dedikeret **Governance-spor** i Azure DevOps. Vi sikrer overholdelse af **KOMBITs principper** ved at:

1. **Spærre for releases**, hvis der er åbne sikkerheds- eller bias-opgaver.
2. **Kvantificere risici** via FAIR i alle Design Spikes, så vi handler på fakta frem for mavefornemmelser.
3. **Garantisikre gennemsigtighed** ved at kræve Model Cards og teknisk forklarlighed (SHAP/LIME) som en del af vores 'Definition of Done'.
4. **Audit-logge alt**, så hver eneste models træningsdata og sikkerhedstjek er sporbare direkte i Azure Boards."