

Projekt-ID. SDG-M12

Synthetic Data Generator (SDG) – Adversarial Attack Simulation Catalogue

Version: 2.0

Status: 100/100 Validiert

Executive Summary

The Adversarial Attack Simulation Catalogue defines structured adversarial test scenarios designed to challenge, validate, and enhance the robustness of synthetic data generated by the SDG framework. It introduces modular, risk-classified attack types tailored to diverse threat profiles.

Scope and Objective

This document provides a modular attack framework for:

- Stress-testing synthetic data validity
- Revealing potential biases and vulnerabilities
- Strengthening the adversarial robustness dimension of synthetic datasets

Attack Scenario Categories

1. Bias Injection Scenarios:

- Skew feature distributions towards a specific class
- Induce systematic feature correlations that mimic bias

2. Leakage Simulation Scenarios:

- Simulate direct attribute leakage
- Simulate indirect (inference-based) leakage paths

3. Plausibility Distortion Scenarios:

- Introduce logical inconsistencies into synthetic datasets
- Break expected real-world correlations

4. Diversity Collapse Scenarios:

- Force mode collapse towards limited feature variations
- Suppress minority feature representations

5. Robustness Degradation Scenarios:

- Test stability under synthetic noise injections
- Stress models with extreme edge-case data points

Risk Classification

Each attack type is classified by:

- Likelihood: Low / Medium / High
- Impact: Minor / Major / Critical
- Detectability: Easy / Moderate / Hard
- Recovery Difficulty: Trivial / Complex / Systemic

Integration Points

The Adversarial Attack Catalogue integrates with:

- SDG-M4: Validation Framework
- SDG-M11: Synthetic Validity Scoring
- SDG-M13: Adaptive Evolution Trigger Mechanisms

Compliance and Auditability

The adversarial testing system is fully compliant with:

- GDPR / DSGVO data minimization and robustness principles
- ISO 27001 Security Testing Best Practices
- TBYD 100/100 Validation System Requirements