# SWaT-KG v0.2 Documentation

## 1. Đặt vấn đề và Ý tưởng

Trong hệ thống CPS (Cyber-Physical System) như nhà máy xử lý nước SWaT, dữ liệu phát sinh từ nhiều lớp: cảm biến vật lý, thiết bị chấp hành, bộ điều khiển (PLC), SCADA, cũng như log các kịch bản tấn công. Nếu các dữ liệu này rời rạc thì rất khó phân tích nguyên nhân – hậu quả. SWaT-KG v0.2 chuẩn hóa tất cả tri thức đó thành một graph đa lớp, nhưng theo hướng khái niệm (conceptual): giữ cấu trúc và logic chung, bỏ chi tiết số đo hay timestamp cụ thể.

## 2. Kiến trúc Schema (v0.2)

Schema gồm các lớp chính:  
- Plant / Stage: biểu diễn cấu trúc 6 giai đoạn (P1–P6).  
- Physical layer: Stage chứa unit, component, sensor, actuator. Mỗi sensor/actuator liên kết với một FeatureType.  
- Procedure & Step: mô tả quy trình vận hành qua các bước Input–Output, nhưng Input liên kết tới Setpoint khái niệm thay vì số cụ thể.  
- Control layer: PLC, SCADA, HMI, Historian, cùng mạng Level-0/1. SCADA tag ánh xạ sensor/actuator (không lưu địa chỉ register chi tiết).  
- Knowledge & Security layer:  
 • Setpoint: ngưỡng khái niệm (LL, HH, MinFlow…).  
 • Rule & Event: Rule tham chiếu Setpoint, sinh Event. Không chứa giá trị số, chỉ quan hệ khái niệm.  
 • Attack: kịch bản tấn công (SSSP, SSMP, MSSP, MSMP), gồm stage, points, intent, expected\_effects.  
 • Invariants: ràng buộc khái niệm (ví dụ: Valve CLOSED ⇒ flow ≈ 0).  
- Data Stream: chỉ lưu modes = {normal, attack}, không lưu chu kỳ lấy mẫu hay dữ liệu cụ thể.

## 3. Ví dụ minh họa

Stage P1 – Raw Water Intake  
- Sensor: LIT-101 (mực nước), FIT-101 (lưu lượng).  
- Actuator: MV-101 (van), P-101/P-102 (bơm chính và dự phòng).  
- Procedure FillCycle:  
 • Step 1: Mở MV-101.  
 • Step 2: Khi đạt Setpoint SP.P1.LIT101.LL\_OK, khởi động P-101.  
- Rule: LowLevelPumpProtect → nếu dưới setpoint LL\_OK thì dừng P-101, phát Event EV.P1.LowLevelTrip.  
  
Stage P3 – Ultrafiltration (UF)  
- Sensor: LIT-301, FIT-301, DPIT-301.  
- Actuator: P-302, MV-301..304.  
- Procedure FilterCycle:  
 • Step 1: Start UF feed khi đạt Setpoint SP.P3.LIT301.LL\_OK.  
 • Step 2: Nếu vượt Setpoint SP.P3.DPIT301.HI, mở van backwash.  
- Rule: HighDPBackwash → phát Event EV.P3.BackwashStart.

## 4. Ý nghĩa và Ứng dụng

- Chuẩn hóa tri thức khái niệm: lưu cấu trúc, quy trình, logic, kịch bản tấn công dưới dạng entity–relation.  
- Truy vấn dễ dàng bằng LLM: ví dụ “Van nào mở khi backwash?”, “Rule nào bảo vệ bơm P-101?”, “Attack nào ảnh hưởng đến LIT-301?”.  
- Phân tích nguyên nhân–kết quả: Rule, Event, Procedure, Attack, Invariant kết nối thành causal graph.  
- Hỗ trợ phát hiện & giải thích bất thường: so sánh dữ liệu thực tế với logic domain (như invariant).  
- Ứng dụng bảo mật ICS: Attack node mô tả kịch bản tấn công → dùng làm ground-truth cho IDS/ML.  
  
Khác biệt v0.2 so với v0.1:  
- Giữ lại ontology lõi, nhưng bỏ số liệu cụ thể (ngưỡng mm/bar, timestamp).  
- Thêm setpoints (ngưỡng khái niệm), invariants (ràng buộc vật lý khái quát).  
- Attack chỉ mô tả kịch bản, không chứa log chi tiết.  
- DataStream chỉ lưu chế độ, không lưu dữ liệu mẫu.