# Project 3 Proposal

# Project Title

IoT Attack & Benign Network Traffic Analysis

# Group Members

Geoff Wilson

Bella Oshomuvwe

Endy Mateo

Andrea Brown

Randy Esser

# Project description

Our project follows the Data Engineering track to perform ETL functions on a public dataset published by the University of California Irvine. RT-IoT2022 contains network traffic analysis details of traffic patterns for twelve (12) known benign and attack patterns (i.e., network traffic metadata collected during three normal IoT behaviours and nine known attack behaviours).

We have implemented ETL functions to clean, fine-tune, and focus the data on areas of interest and for use in intrusion detection systems. Our project utilizes two Python libraries not previously covered in class, [psycopg2](https://pypi.org/project/psycopg2/) and [suricataparser](https://pypi.org/project/suricataparser/), to produce rules for Suricata, the most widely deployed open-source intrusion detection system.

The data is stored in PostgreSQL via SQLAlchemy. Two separate tables store the (a) raw network traffic details and (b) the traffic signature categories. There are 12 categories of traffic, including 3 benign and 9 attack patterns.

We’ve included an optional function to allow the user to select specific segments of the dataset to produce focused Suricata rulesets.

About the dataset (from UCI):

“The RT-IoT2022, a proprietary dataset derived from a real-time IoT infrastructure, is introduced as a comprehensive resource integrating a diverse range of IoT devices and sophisticated network attack methodologies. This dataset encompasses both normal and adversarial network behaviors, providing a general representation of real-world scenarios. Incorporating data from IoT devices such as ThingSpeak-LED, Wipro-Bulb, and MQTT-Temp, as well as simulated attack scenarios involving Brute-Force SSH attacks, DDoS attacks using Hping and Slowloris, and Nmap patterns, RT-IoT2022 offers a detailed perspective on the complex nature of network traffic. The bidirectional attributes of network traffic are meticulously captured using the Zeek network monitoring tool and the Flowmeter plugin. Researchers can leverage the RT-IoT2022 dataset to advance the capabilities of Intrusion Detection Systems (IDS), fostering the development of robust and adaptive security solutions for real-time IoT networks.”

# Project GitHub repository

<https://github.com/Matendy12/Project-3/tree/main>

# Data set Link

<https://archive.ics.uci.edu/dataset/942/rt-iot2022>

[Jupyter notebooks](https://github.com/Matendy12/Project-3/tree/main/code)

ERD