

# CICI:IPAAI:CANIS: Curated AI-ready Network telescope datasets for Internet Security

## Ricky Mok and kc claffy (UC San Diego/CAIDA)

### Summary

UCSD Network Telescope (UCSD-NT) leverages a large block of unused IPv4 address space to collect unsolicited Internet traffic. CANIS is a suite of tools and reference datasets to 1) safeguard the *data integrity* of the UCSD-NT, and 2) offer *AI-ready datasets* for researchers to train, evaluate, and validate their AI-based cybersecurity applications/models.

### Intellectual Merit

- Multi-layer monitoring capability to detect events that can induce UCSD-NT data loss
- UCSD-NT data enriched with metadata and labels
- Facilitation of training and evaluation of AI-based cybersecurity applications/models with AI-ready reference datasets

### Approach

- Develop continuous monitoring of BGP routing data, scanning campaigns, and system telemetry to ensure integrity of control and data plane between the UCSD-NT and the Internet
- Create labeling functions with IP blocklists, malware fingerprints, and other malicious engagement signals
- Develop new data format that facilitates AI applications to ingest and analyze the data

