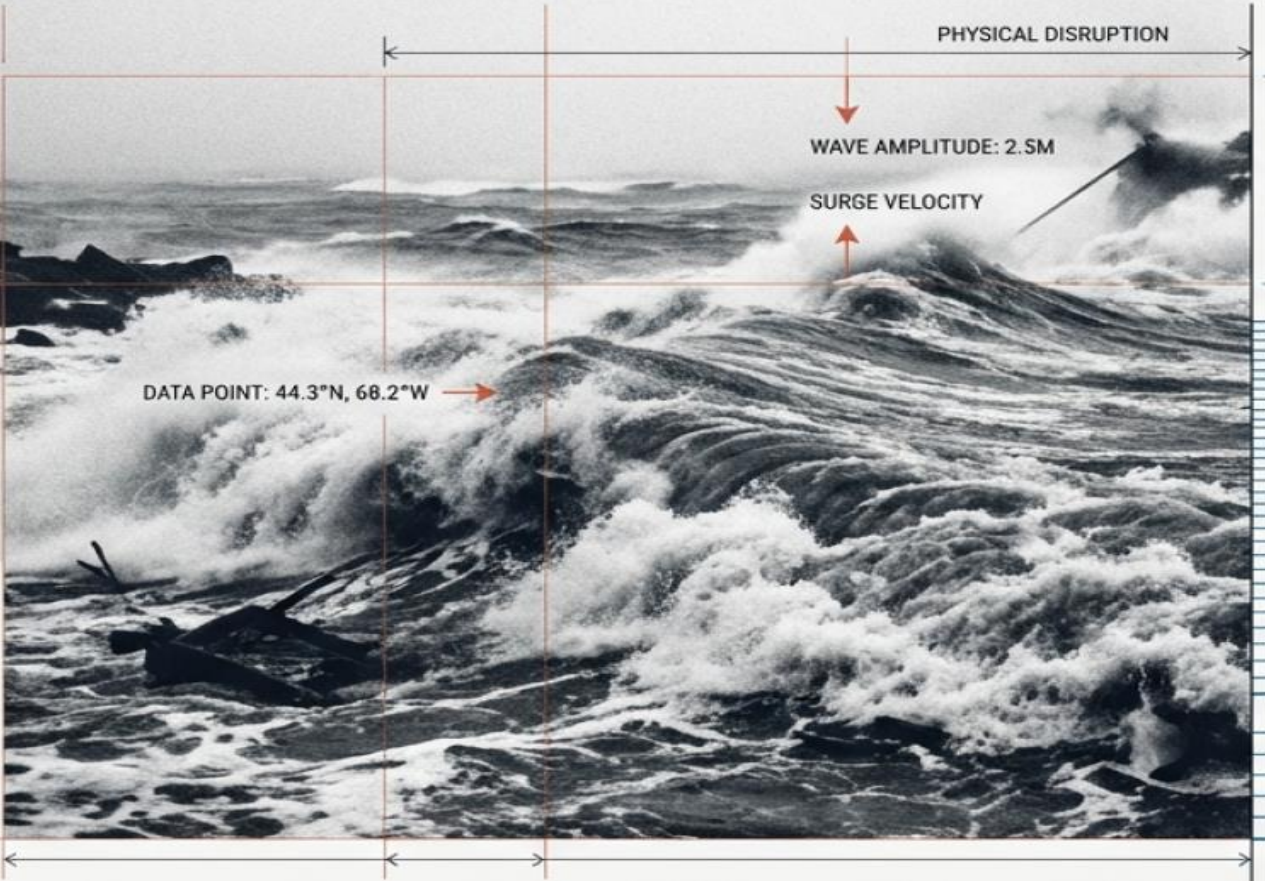
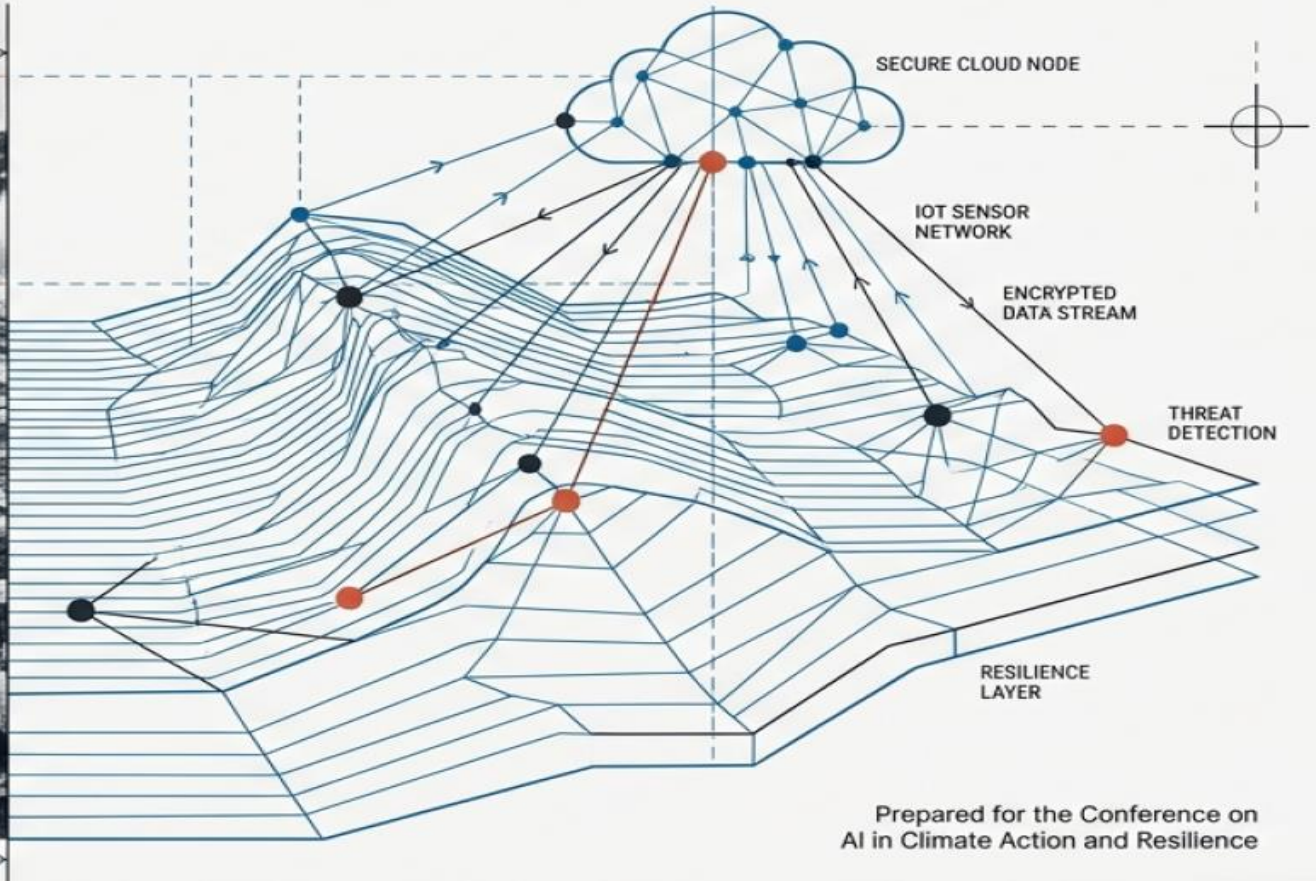
Ensuring Truth in Early Warnings

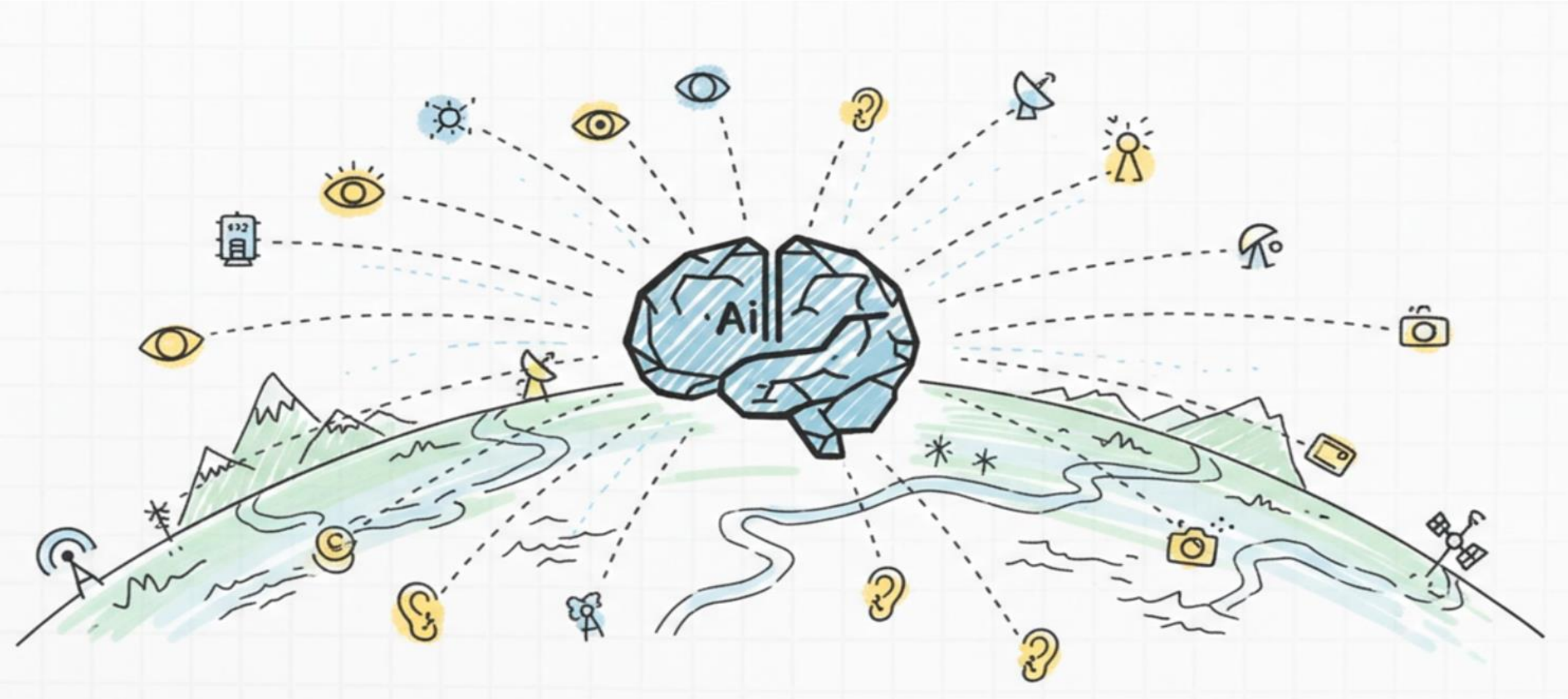
A Dynamic Security Framework for Climate Intelligence Networks & IoT





Presented by: PADSN Wijesekara (PhD (thesis submitted), B.Sc. Engineering, First-class hons.)
Lecturer, Department of Electrical and Information Engineering, Faculty of Engineering, University of Ruhuna, Sri Lanka

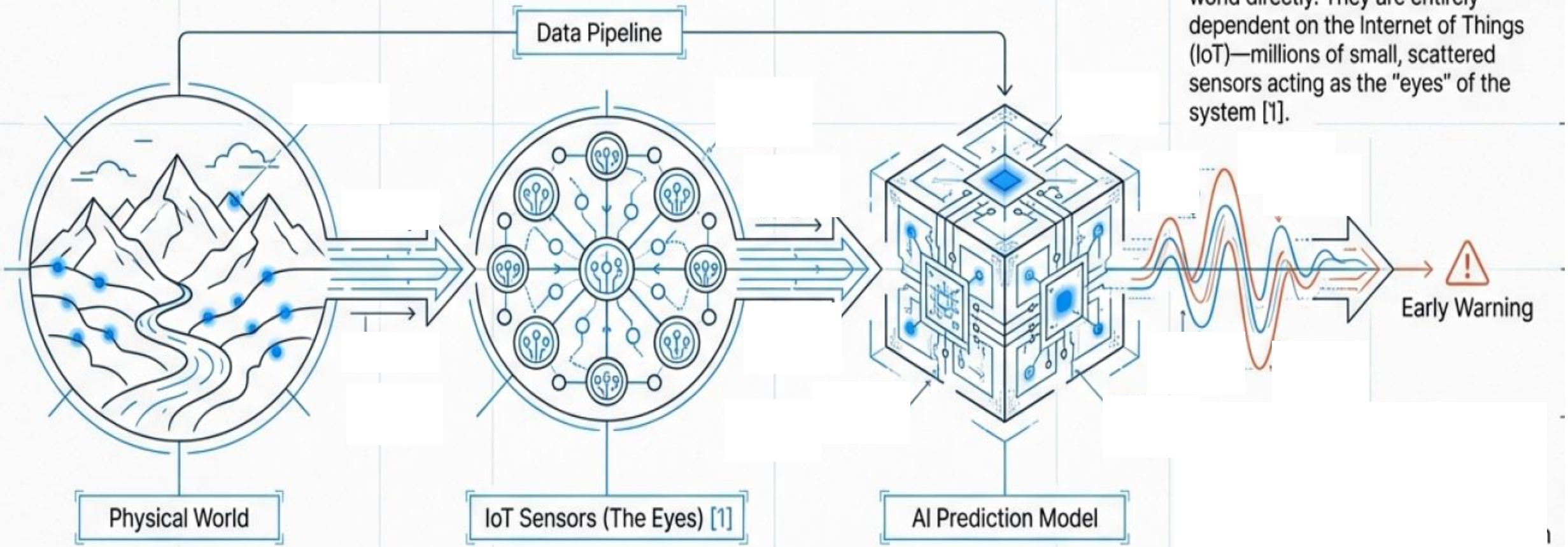
Climate intelligence in IoT



AI is Our Frontline Defense, But It Has Eyes

We currently rely on AI models to predict devastating events—when a flood will hit a valley, a landslide will bury a road, or a cyclone will strike the coast [1].

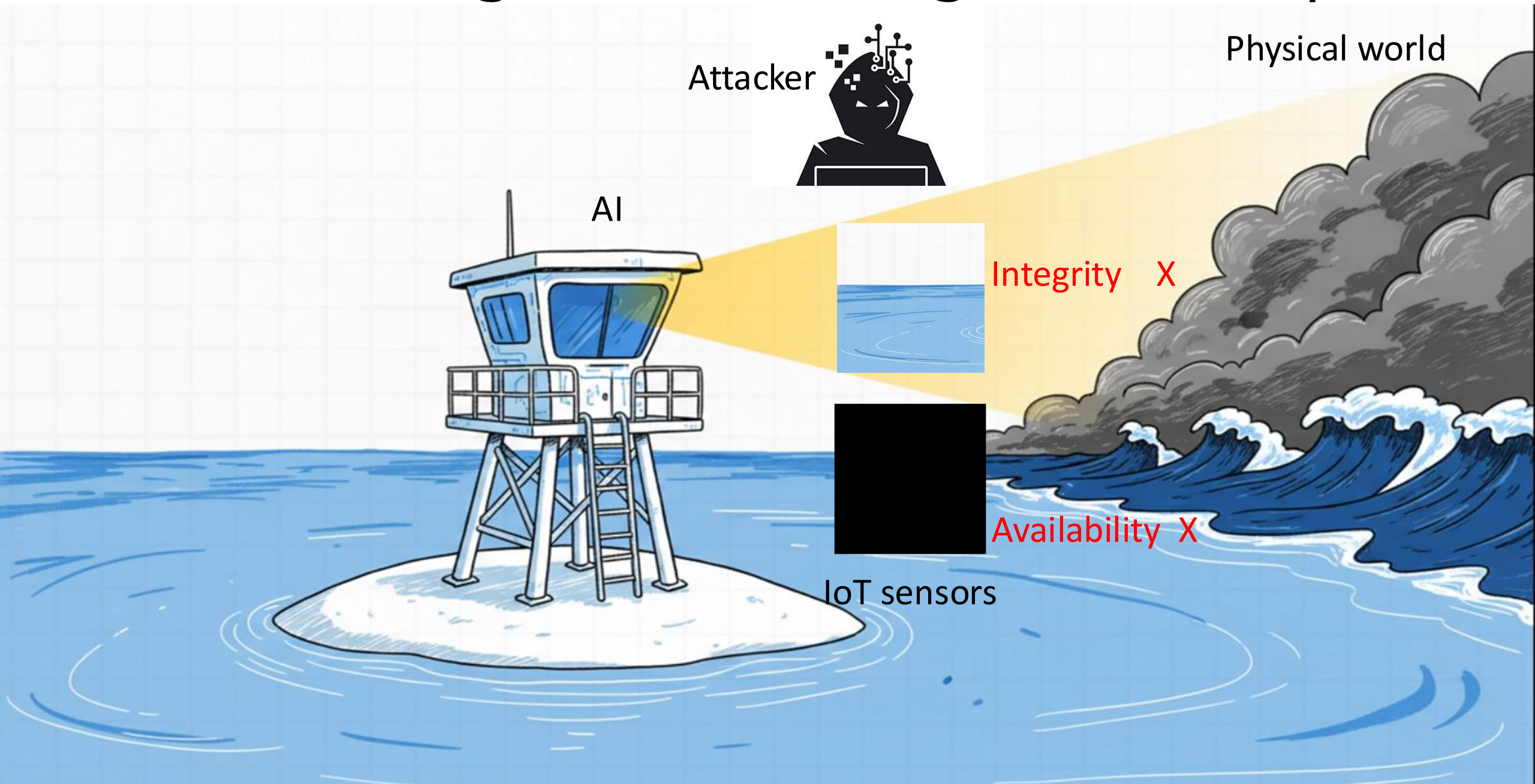
However, these models do not see the world directly. They are entirely dependent on the Internet of Things (IoT)—millions of small, scattered sensors acting as the "eyes" of the system [1].



Anemometer (wind speed), barometer (atmospheric pressure), rain gauge (rainfall), flow sensor (flow rate), soil moisture, inclinometer (ground tilt), piezometer (soil subsurface water pressure), temperature sensor, humidity sensor (air humidity), tide gauge, pressure sensor (water pressure)

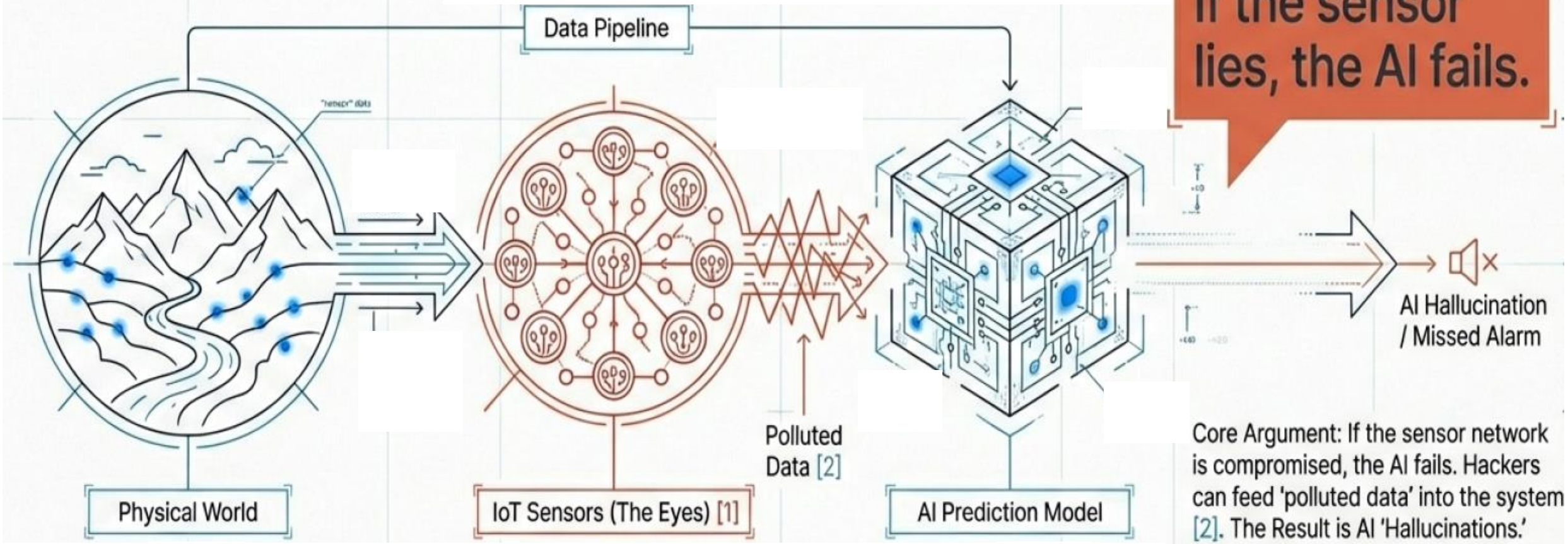
A small but critical weakness that can cause major failure

Misleading climate intelligence concept



The Achilles' Heel: Polluted Data & Hallucinations

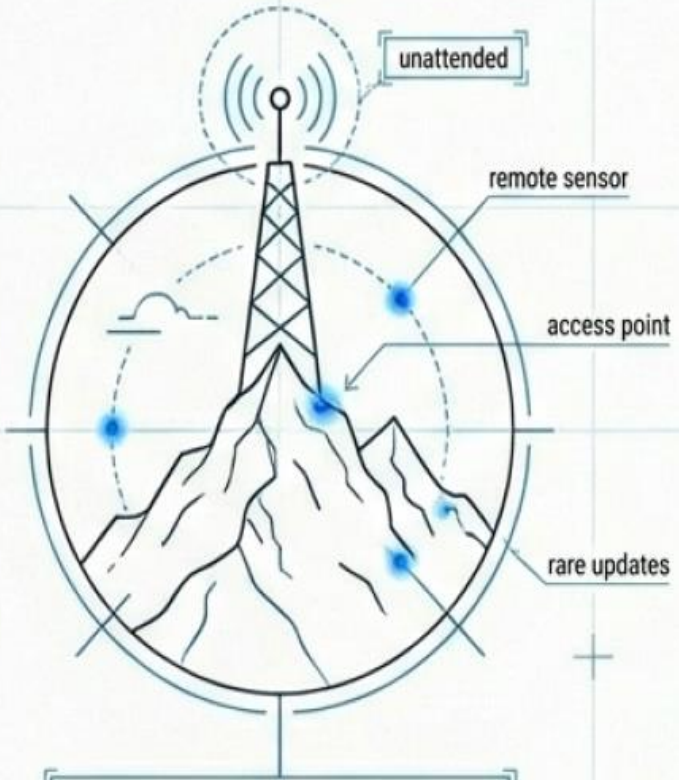
A small but critical weakness that can cause major failure



Core Argument: If the sensor network is compromised, the AI fails. Hackers can feed 'polluted data' into the system [2]. The Result is AI 'Hallucinations.' If a hacked sensor reports low water levels during a flood, the AI will not trigger the alarm [2].

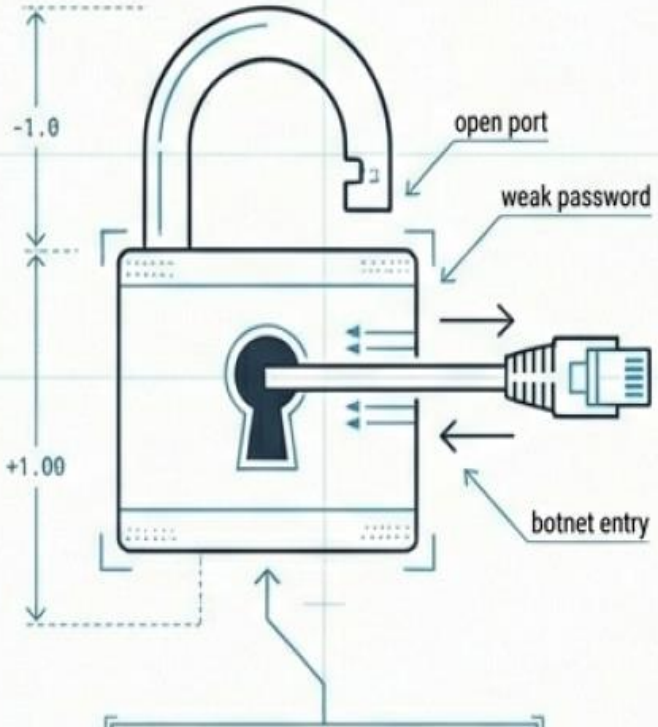
Why the Threat is Invisible to Traditional Security

Mirai - scan, check ports open?, dictionary attack for such, get control, 6 lacks, launches massive DDoS, Down domain name service (DNS)



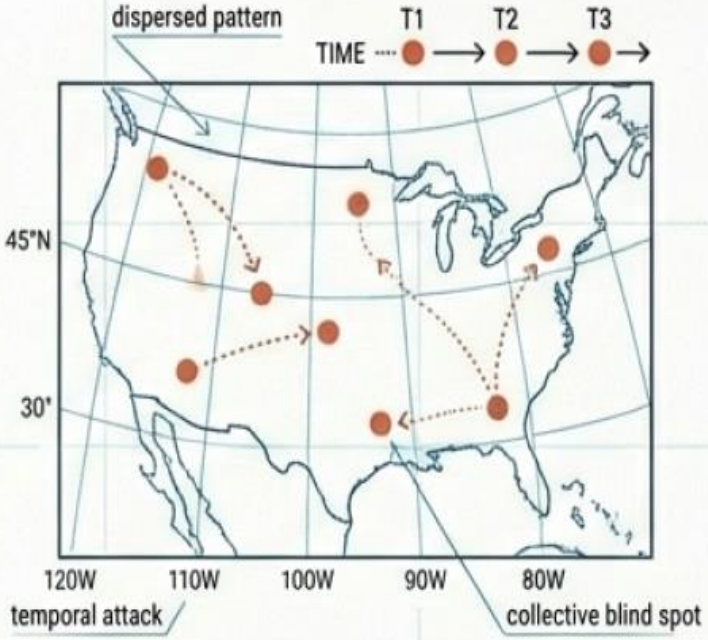
Unattended Hardware

Unlike corporate laptops, climate sensors are remote (mountainsides, riverbeds), rarely updated, and often physically accessible [4].



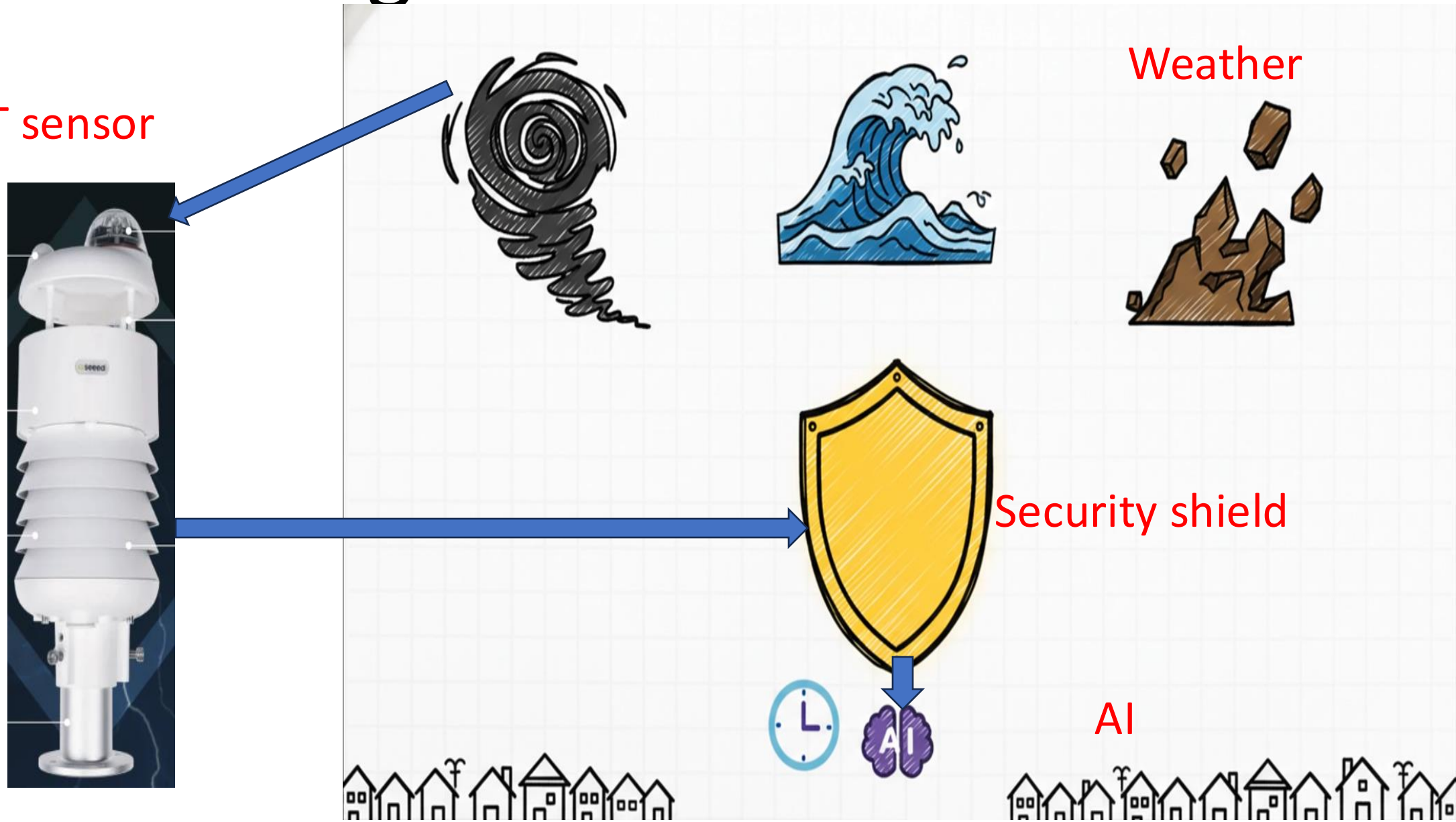
Basic Vulnerabilities

Many legacy sensors have open digital ports and weak passwords, making them easy prey for massive botnets like Mirai [2, 5].

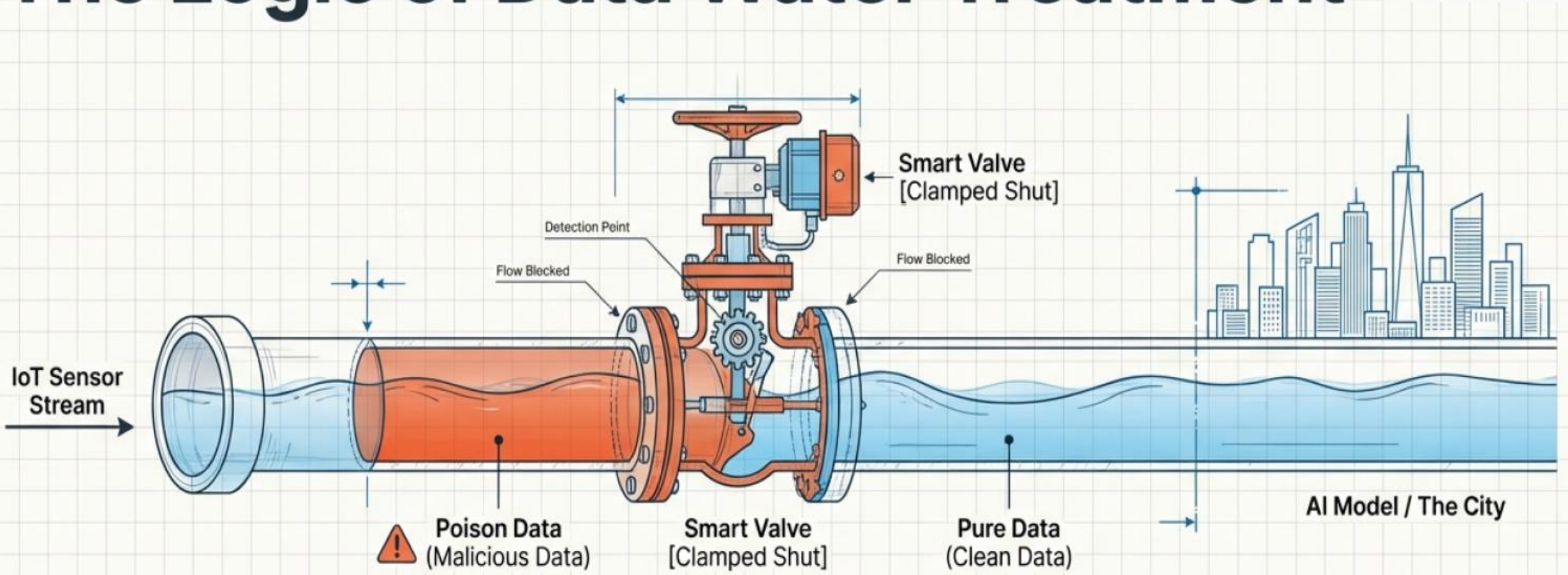


Attacks are "spatially and temporally dispersed" [6].
An attacker nudges a sensor in the north today, and one in the south tomorrow [7, 8]. Individually, they look like glitches; collectively, they blind the system [9, 10].

Climate intelligence in IoT with defense



The Logic of Data Water Treatment



The Analogy: Sensors are pipes; data is water. If “poison” (malicious data) is detected, we shut off that specific valve [27]. We ensure the city (the AI) drinks only pure water, guaranteeing the integrity of the warning system.

Strategic Shift: Building a Digital Immune System

Moving from 'Connect and Forget' to a biological defense model [1, 3].

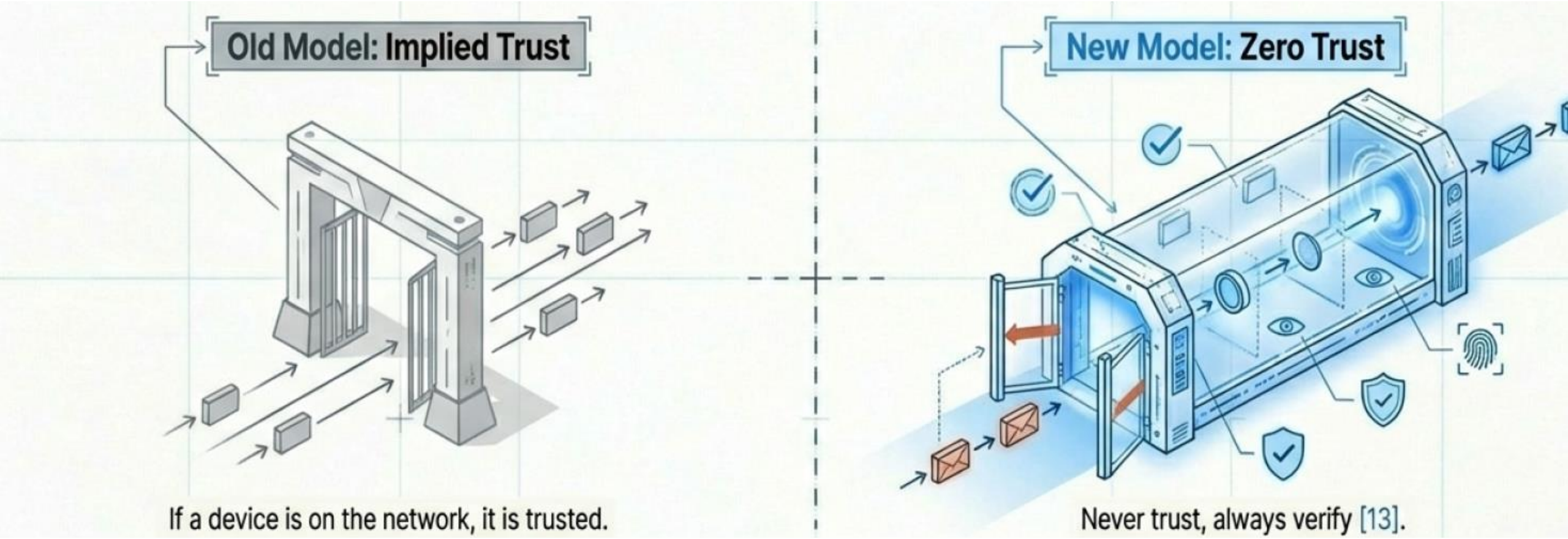
I. The Digital Guard
(Zero Trust)

II. The AI Detective
(ShadowScan)

III. The Network Analyst
(ADEPT/APEX)

The goal is to actively detect and isolate threats so we can act on the truth.

Pillar I: The Digital Guard — Zero Trust Architecture

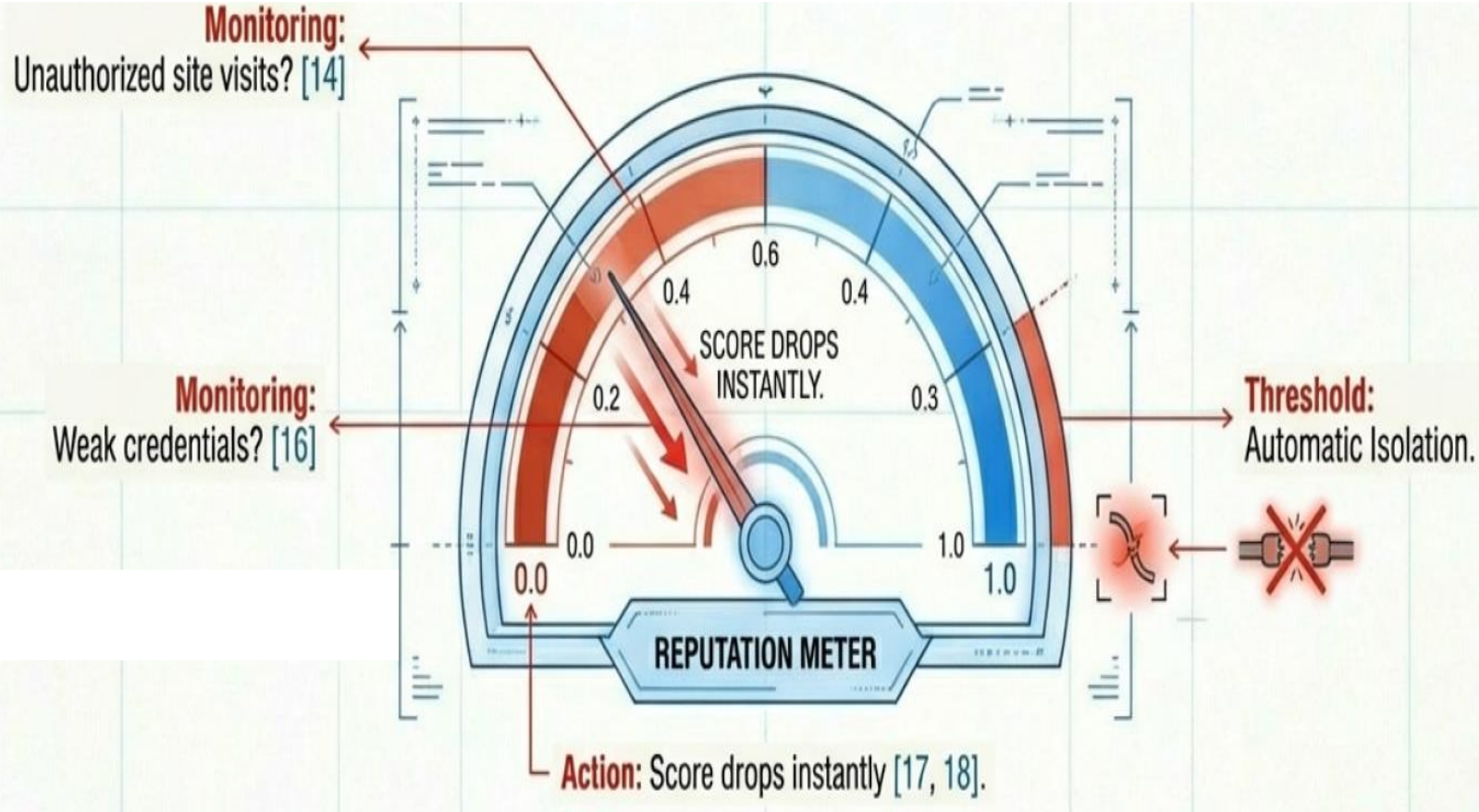


- Always authenticated
- Least privilege access
- Micro-segmentation

We propose a Zero Trust Architecture [11, 12]. Implied trust is removed. Every device must constantly prove its integrity to remain part of the data stream.

The Mechanism: Dynamic Trust Scores

$$\text{Trust Score} = \alpha\text{ML} + \beta\text{EA} + \gamma\text{CR} + \delta\text{ST}$$



Every sensor is assigned a dynamic score from 0 to 1 [15]. When the score falls below a threshold, the AI prediction model automatically ignores that sensor's readings [14]. One 'bad apple' is isolated before it spoils the prediction.

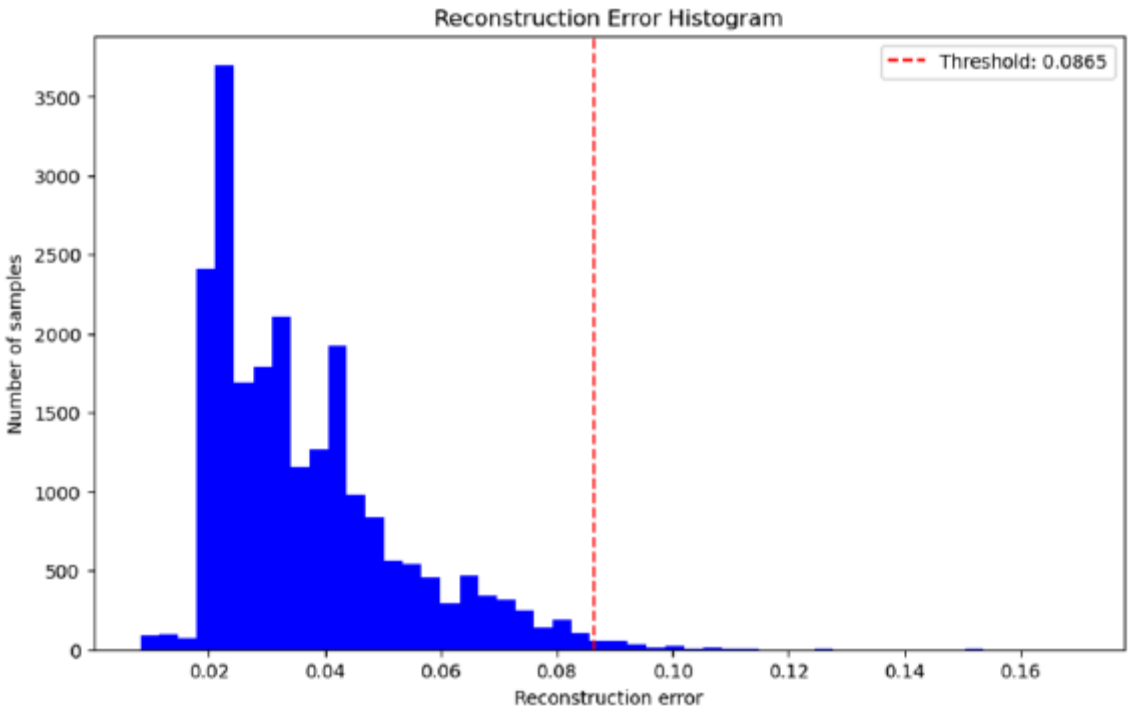
ML – Adaptive threat detection

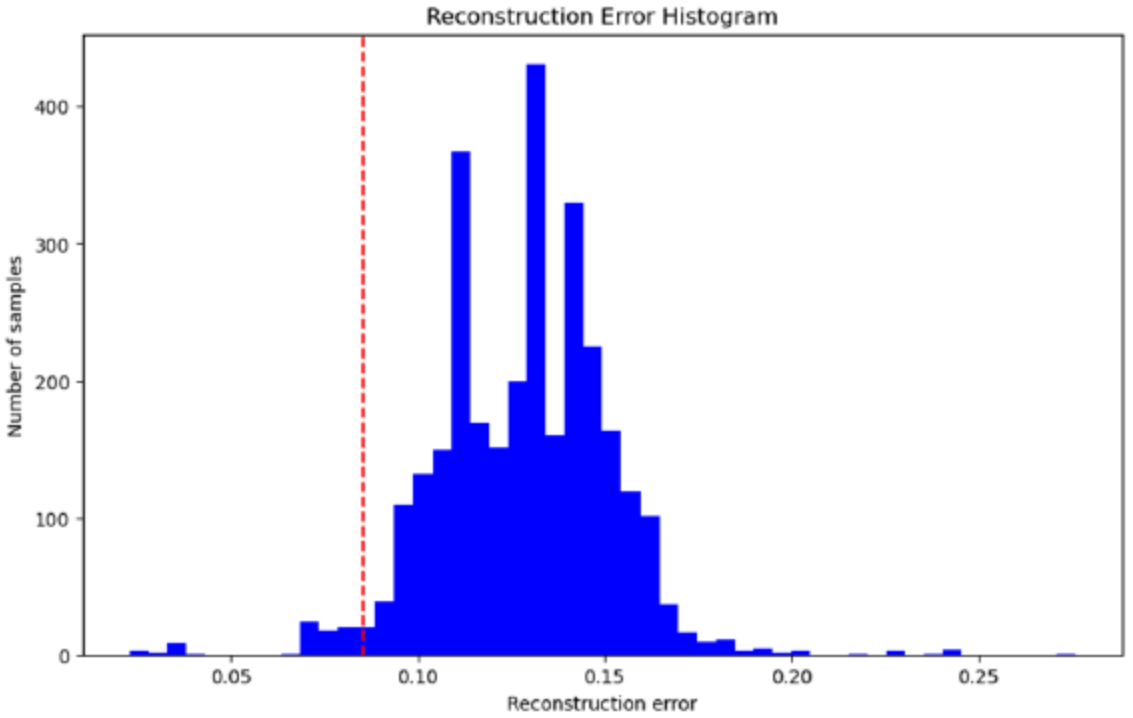
EA – Ease of access

CR – Cyber risk

ST – Security level (how much protected)

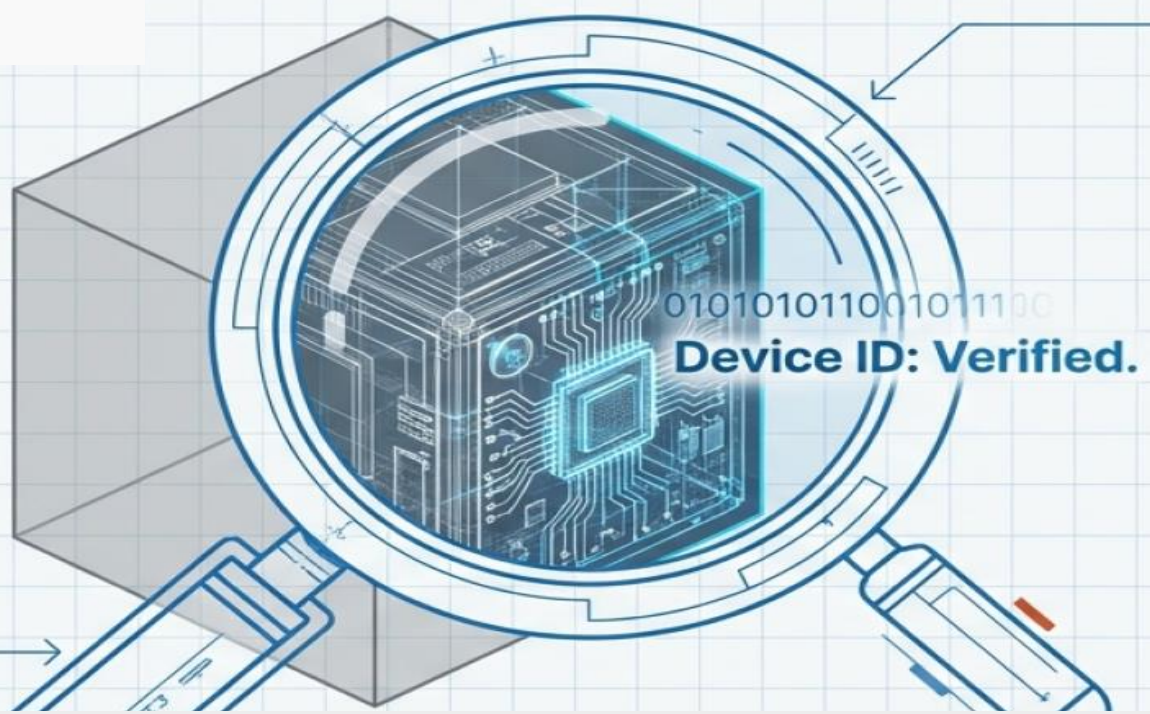
Results – Dynamic trust scores





Pillar II: The AI Detective — ShadowScan

The Challenge: In climate emergencies, new sensors are deployed rapidly [19]. Traditional security blocks them because they aren't on the list.

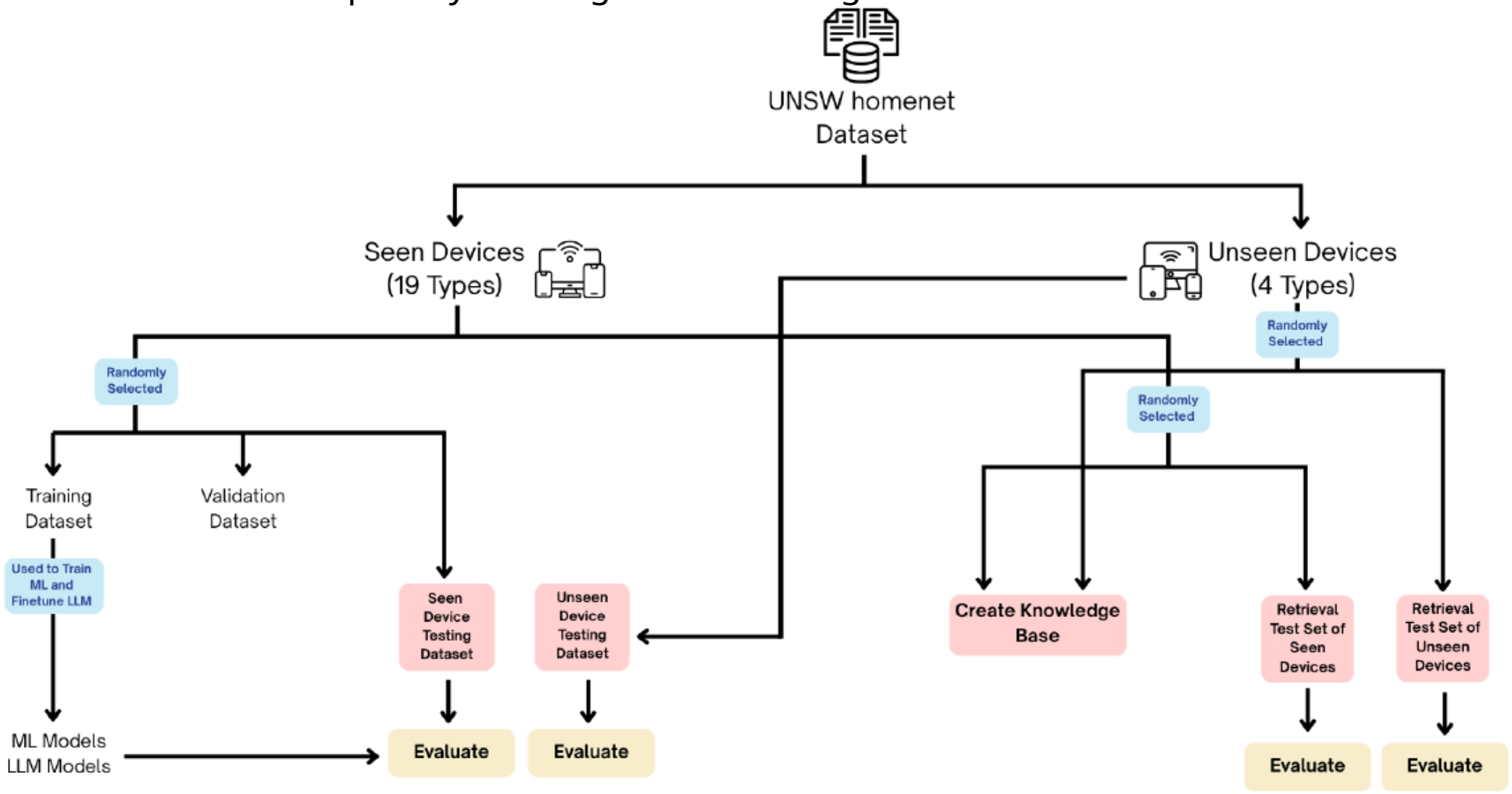


The Solution: ShadowScan utilizes utilizes Large Language Models (LLMs)—the tech behind ChatGPT—to recognize device fingerprints instantly [20].

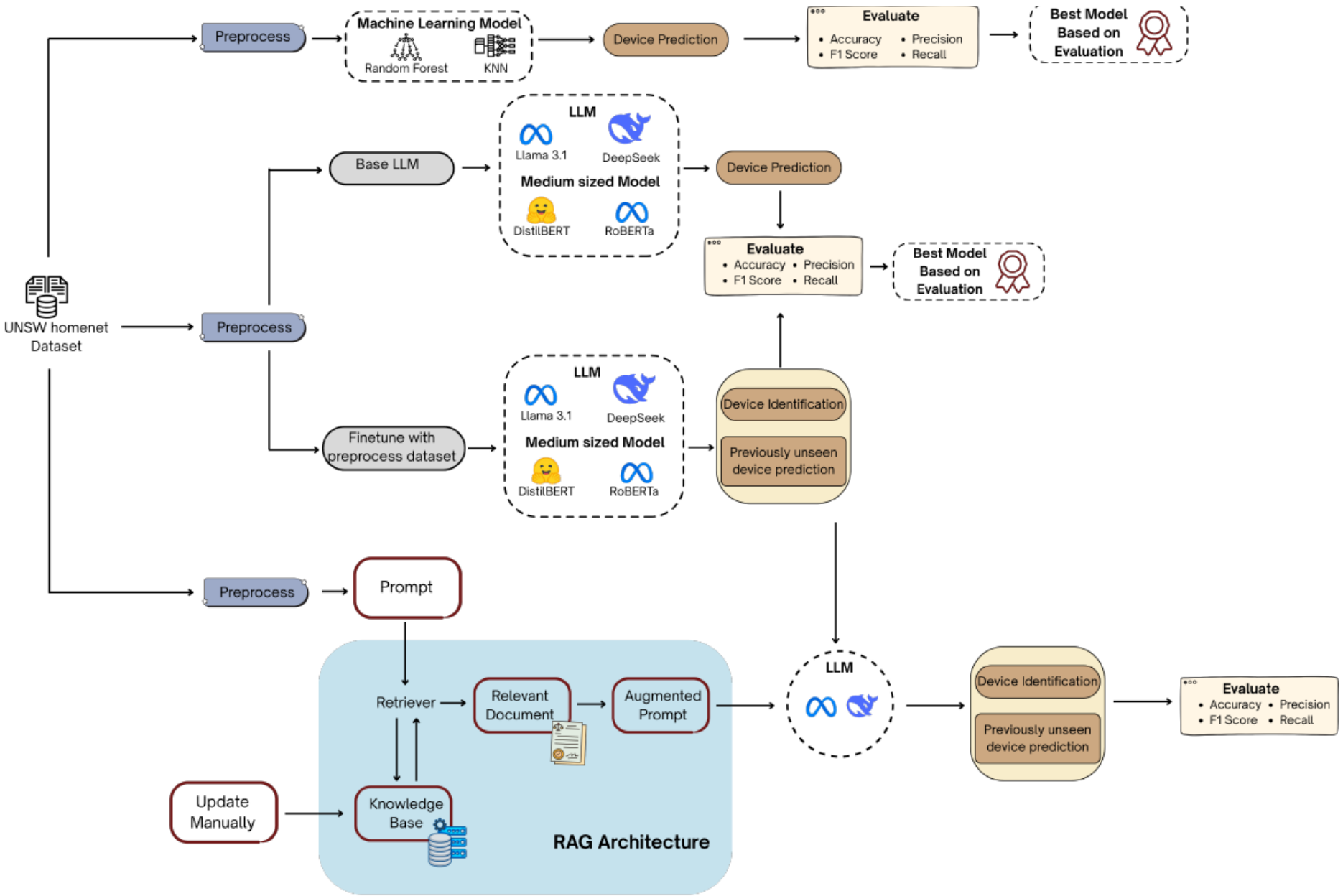
- A fine-tuned LLM approach for IoT device fingerprinting, enabling accurate identification based on network traffic
- A zero-shot prompt engineering method based on LLM with RAG to detect previously unseen IoT devices without requiring model retraining.

Shadow Scan

- Zero shot has the capability to recognize something it has never seen before.



Shadow Scan


UNSW homenet Dataset
Preprocess
Machine Learning Model
Random Forest
KNN
Device Prediction
Evaluate
• Accuracy • Precision
• F1 Score • Recall
Best Model Based on Evaluation
Preprocess
Base LLM
LLM
Llama 3.1
DeepSeek
Medium sized Model
DistilBERT
RoBERTa
Device Prediction
Evaluate
• Accuracy • Precision
• F1 Score • Recall
Best Model Based on Evaluation
Preprocess
Finetune with preprocess dataset
LLM
Llama 3.1
DeepSeek
Medium sized Model
DistilBERT
RoBERTa
Device Identification
Previously unseen device prediction
Preprocess
Prompt
Retriever
Relevant Document
Augmented Prompt
Update Manually
Knowledge Base
RAG Architecture
LLM
Device Identification
Previously unseen device prediction
Evaluate
• Accuracy • Precision
• F1 Score • Recall

Shadow Scan

TABLE V
EVALUATION RESULTS OF THE THREE BASE LLMS

	Accuracy	Precision	Recall	F1 Score
Llama	6%	0.01%	5%	1%
DistilBERT	4.8%	0.3%	4.8%	0.4%
RoBERTa	5.6%	0.3%	5.6%	0.6%

TABLE VI
EVALUATION RESULTS OF THE FOUR FINETUNED LLMS

Model	Accuracy (%)		Precision (%)		Recall (%)		F1 Score (%)	
	Seen	Unseen	Seen	Unseen	Seen	Unseen	Seen	Unseen
Llama	87	10	86	10	87	10	87	10
DistilBERT	89	0	90	0	90	0	90	0
RoBERTa	90	0	90	0	90	0	90	0

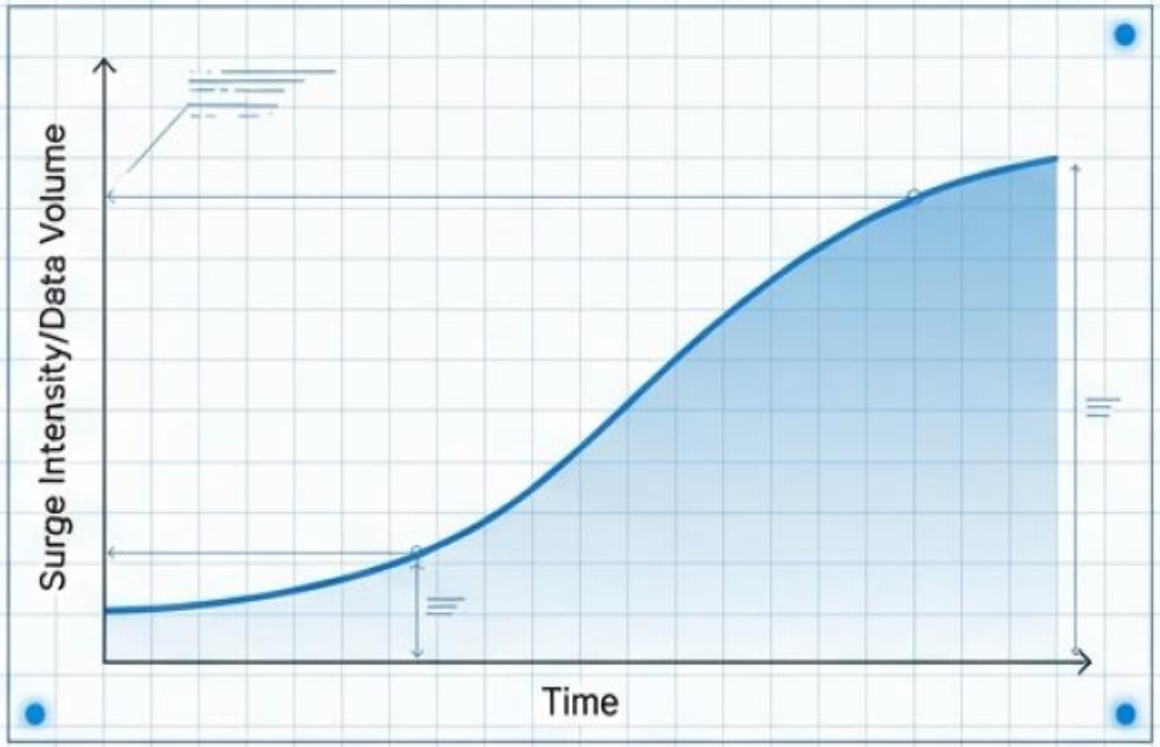
TABLE VII
EVALUATION RESULTS OF THE FINETUNED LLM + RAG

	Accuracy	Precision	Recall	F1 Score
Seen	94%	95%	94%	94%
Unseen	81%	100%	81%	89%

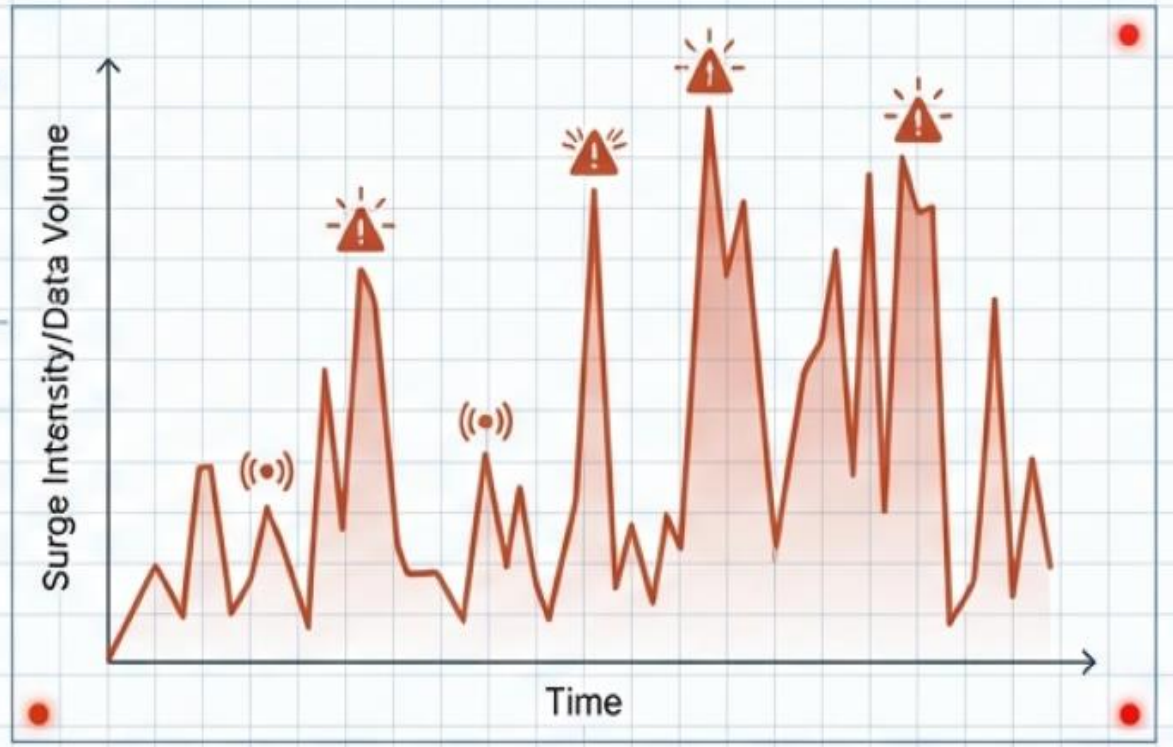
TABLE VIII
OVERALL RESULTS

		Evaluation Metrics Results (%)							
Model	Approach	Accuracy		Precision		Recall		F1 Score	
		Seen	Unseen	Seen	Unseen	Seen	Unseen	Seen	Unseen
Random Forest	-	84	0	84	0	84	0	84	0
KNN	-	73.3	0	73.4	0	73.2	0	73.3	0
Llama	Base LLM	-	6	-	0.01	-	5	-	1
	Finetuned LLM	87	10	86	10	87	10	87	10
DistilBERT	Base LLM	-	4.8	-	0.3	-	4.8	-	0.4
	Finetuned LLM	89	0	90	0	90	0	90	0
RoBERTa	Base LLM	-	5.6	-	0.3	-	5.6	-	0.6
	Finetuned LLM	90	0	90	0	90	0	90	0
RAG + LLM (LLaMA)		94	81	95	100	94	81	94	89

Pillar 3: ADEPT - Detection and Identification of Correlated Attack Stages in IoT Networks



Natural Weather Surge (Flood)



Malicious Digital Surge (Attack) [24, 25]

These frameworks act as detectives for the regional network [6, 23]. They use pattern-mining to "connect the dots" across space and time [9, 24], identifying illogical data spikes that indicate an attack rather than a weather event.

Pillar 3: ADEPT

- Adept processes IoT traffic locally at the gateways to detect anomalous activities with respect to the normal profiles of the IoT devices. Alerts are generated only for those traffic flows that do not match against the profile and, subsequently, they are sent to the security manager. In comparison to sending all the network traffic to a central entity.
- Frequent itemset mining (FIM)—at the security manager to exploit the spatial and temporal characteristics of alerts received from different gateways and efficiently extract patterns corresponding to attack stages.
- Makes use of both alert-level and pattern-level information and classifies the malicious activities to probable stages of attacks using machine learning techniques.

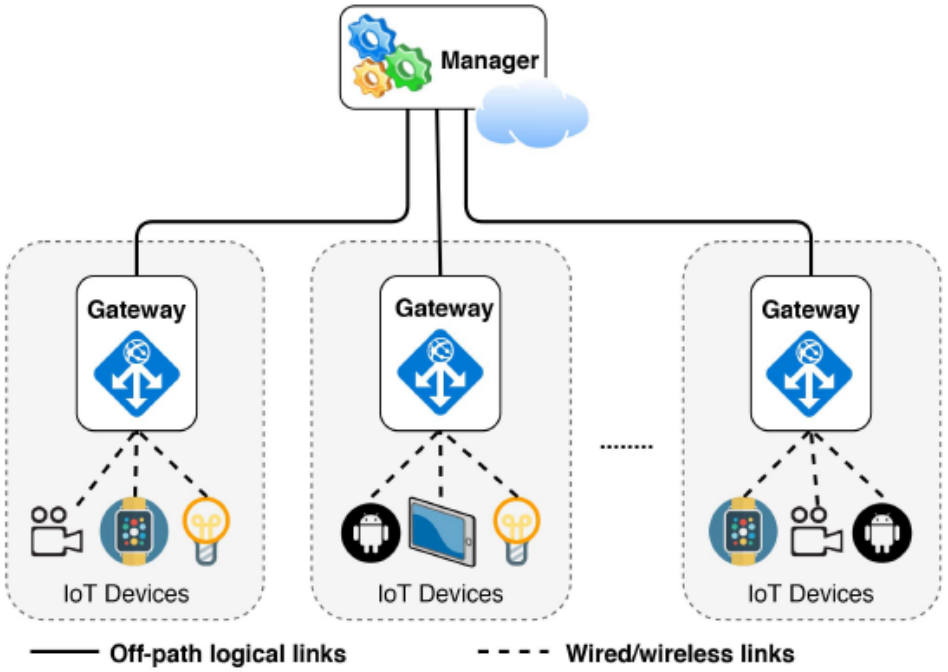


Fig. 1. Hierarchically distributed network architecture of Adept.

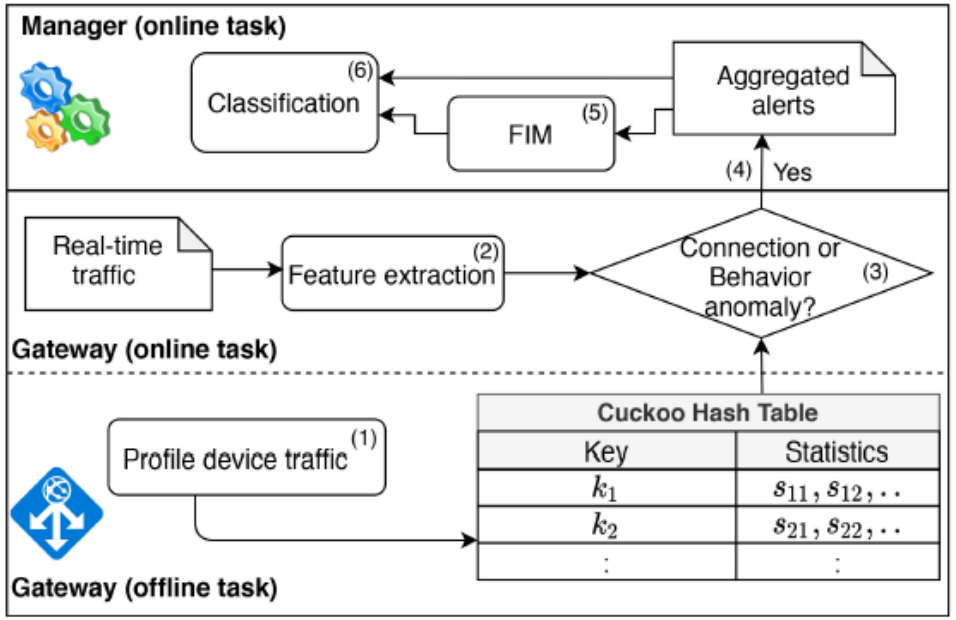


Fig. 2. Functional block diagram of Adept.

Pillar 3: ADEPT

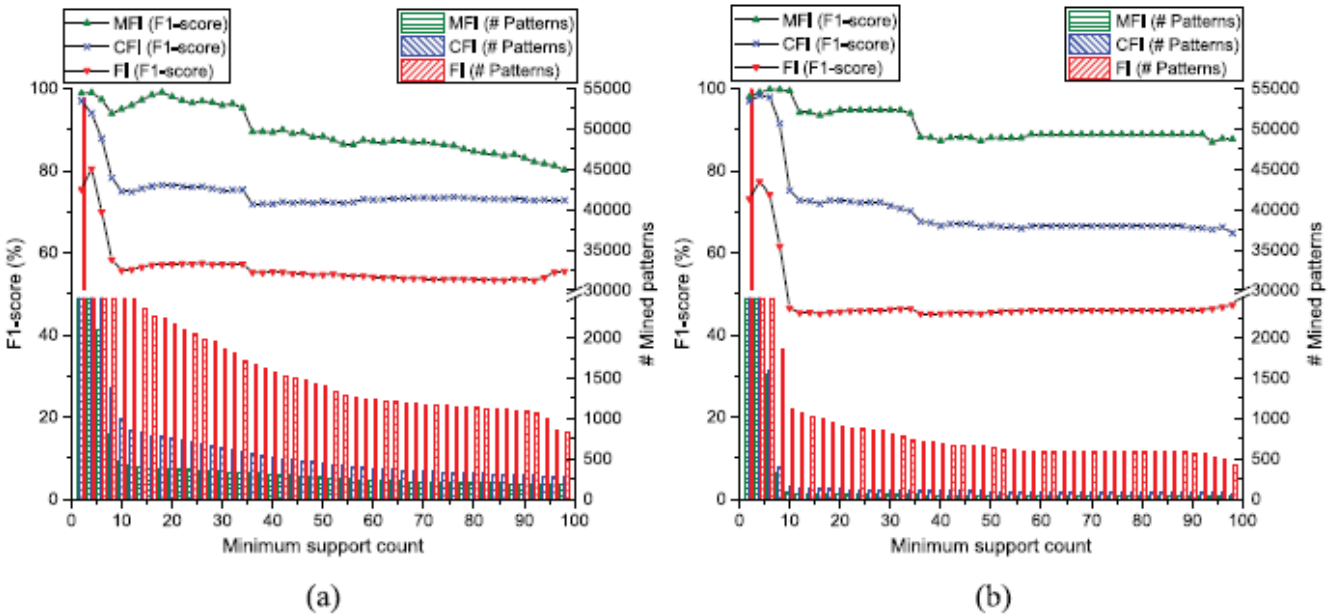
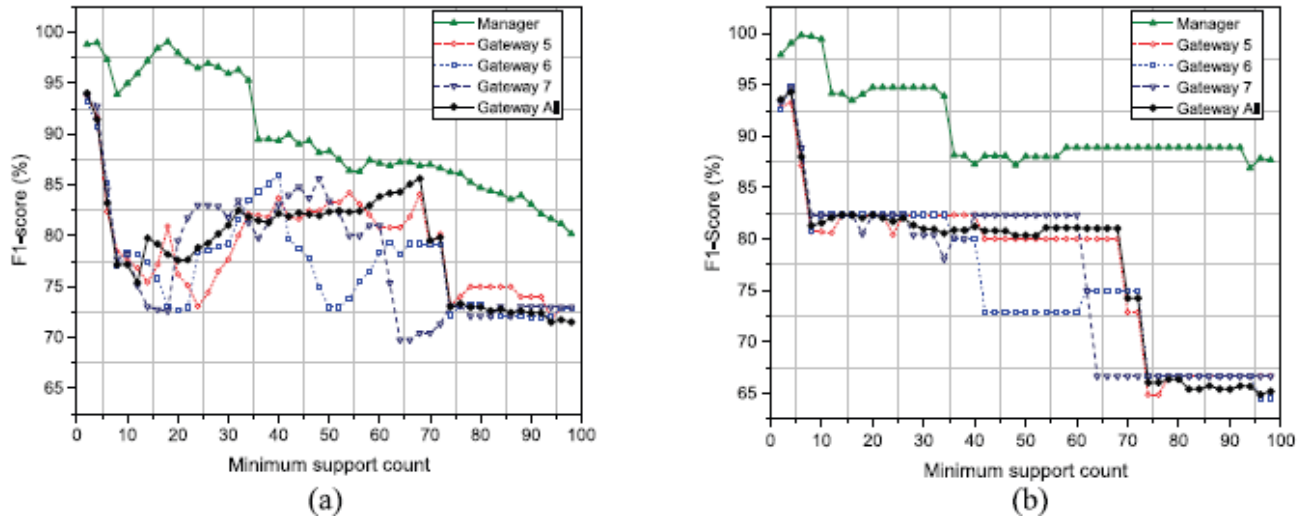


Fig. 6. Performance comparison between MFI, CFI, and FI. (a) Subnet-independent. (b) Subnet-dependent.




thank you!