Towards an AI-powered Player in Cyber Defence Exercises

Roland Meier[1], Artūrs Lavrenovs[2], Kimmo Heinäaro[2],

Luca Gambazzi[3], Vincent Lenders[3]

26.05.2021

CYCON

(1) ETH zürich

(2) CCDCOE NATO COOPERATIVE CYBER DEFENCE CENTRE OF EXCELLENCE

(3) Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
armasuisse

2’500 Attacks

CYCON

2’500 Attacks

against

4’000 Systems

CYCON

Locked Shields is the largest live-fire global cyber defense exercise

Picture: NATO CCDCOE

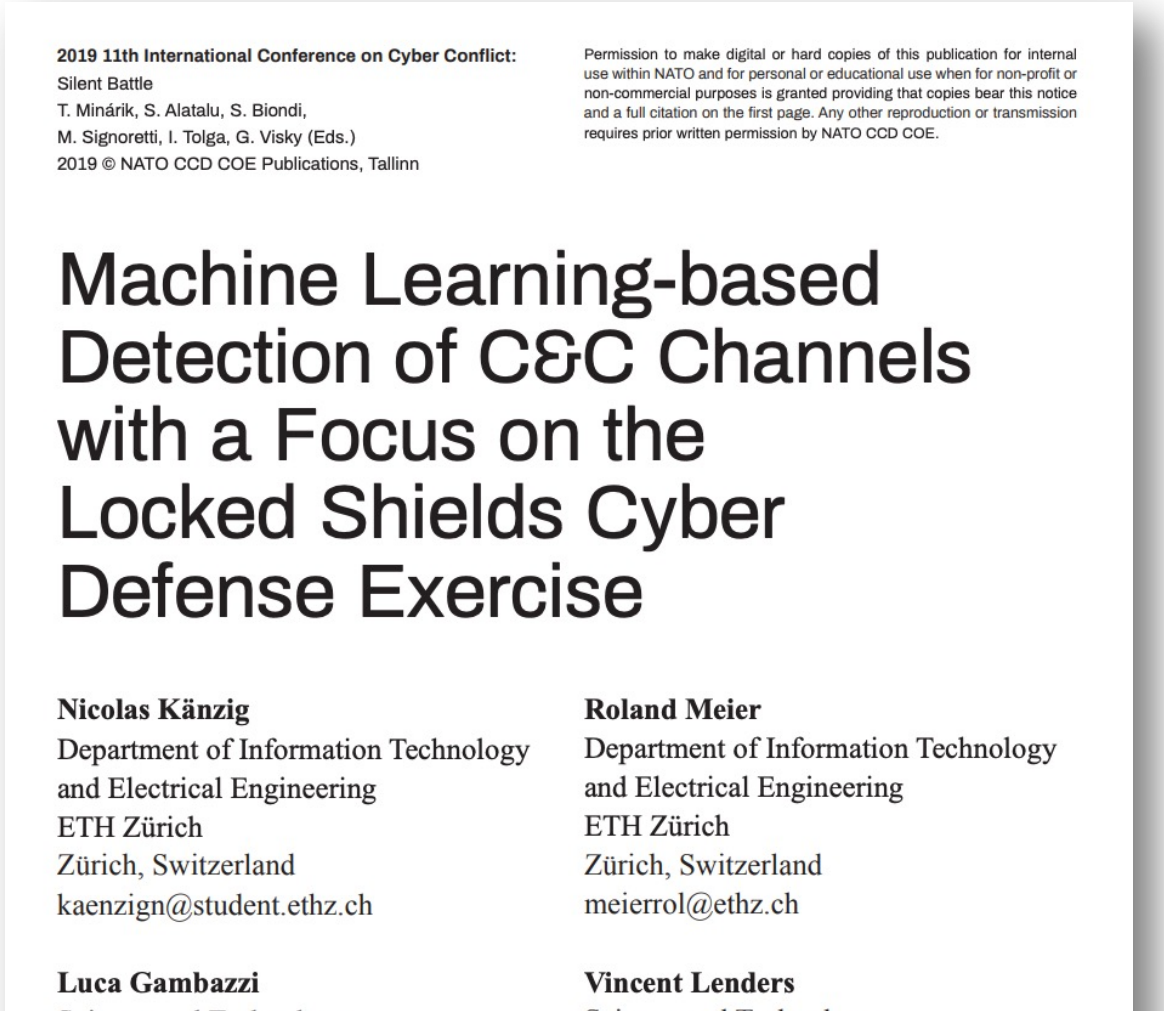
Locked Shields is the largest live-fire global cyber defense exercise

- Red Team vs. Blue Team exercise

Attackers	Defenders
1 Team	1 Team / country

Picture: NATO CCDCOE

2 years ago, we presented a system which uses AI to identify C&C channels



2019 11th International Conference on Cyber Conflict:
Silent Battle
T. Minárik, S. Alatalu, S. Biondi,
M. Signoretti, I. Tolga, G. Visky (Eds.)

Machine Learning-based Detection of C&C Channels with a Focus on the Locked Shields Cyber Defense Exercise

Nicolas Känzig
Department of Information Technology and Electrical Engineering
ETH Zürich
Zürich, Switzerland
kaenzign@student.ethz.ch

Roland Meier
Department of Information Technology and Electrical Engineering
ETH Zürich
Zürich, Switzerland
meierrol@ethz.ch

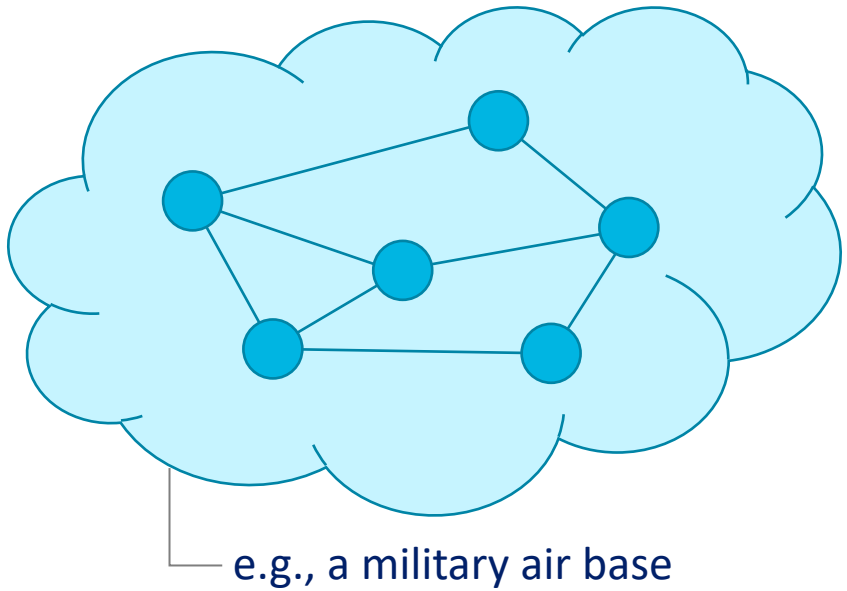
Luca Gambazzi

Vincent Lenders

CYCON

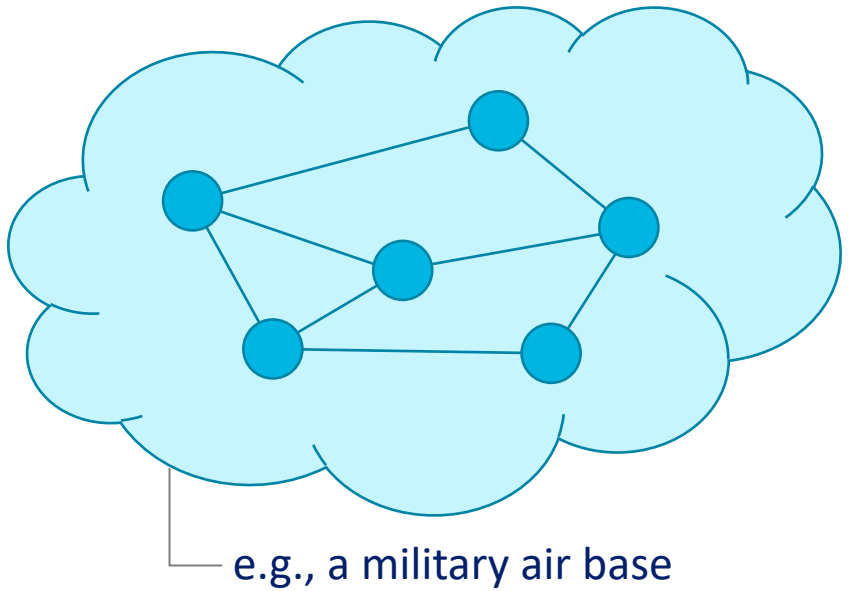
What would it take to have a fully automated Blue Team in a future iteration of Locked Shields?

Each Blue Team is responsible for its network (“Gamenet”)



CYCON

Each Blue Team is responsible for its network (“Gamenet”)



3 main tasks:

- Perform initial hardening
- Defend against attacks
- Communicate with other teams

CYCON

Besides defending its network, a Blue Team needs to communicate with other teams

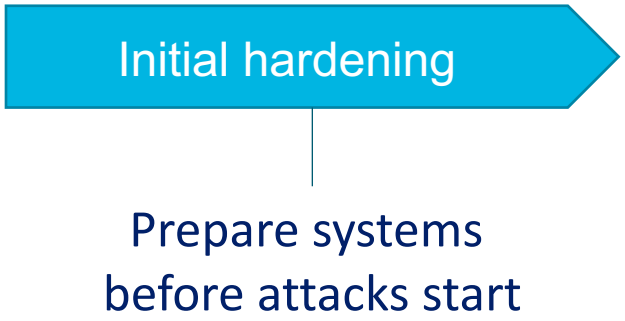
User Simulation Team	Read, address and respond to support tickets
Yellow Team	Provide periodic reports
White team	Voice or video calls through the Gamenet
Green Team	Gamenet status, reverting of devices

CYCON

Four stages in the exercise

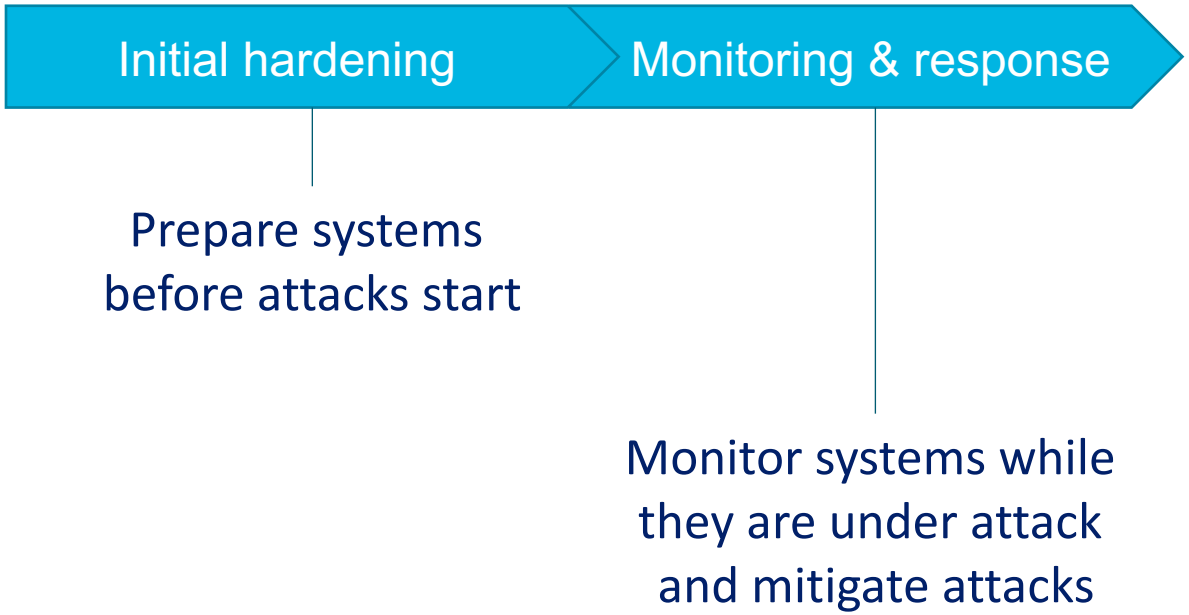
CYCON

Four stages in the exercise



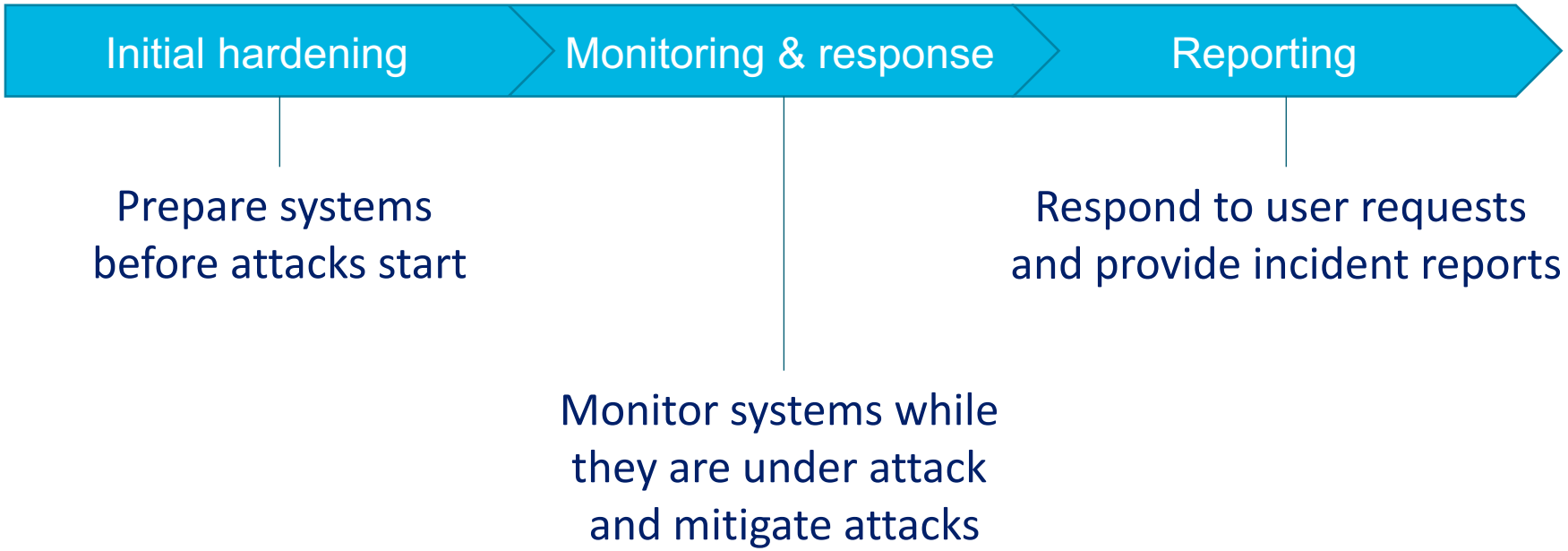
CYCON

Four stages in the exercise



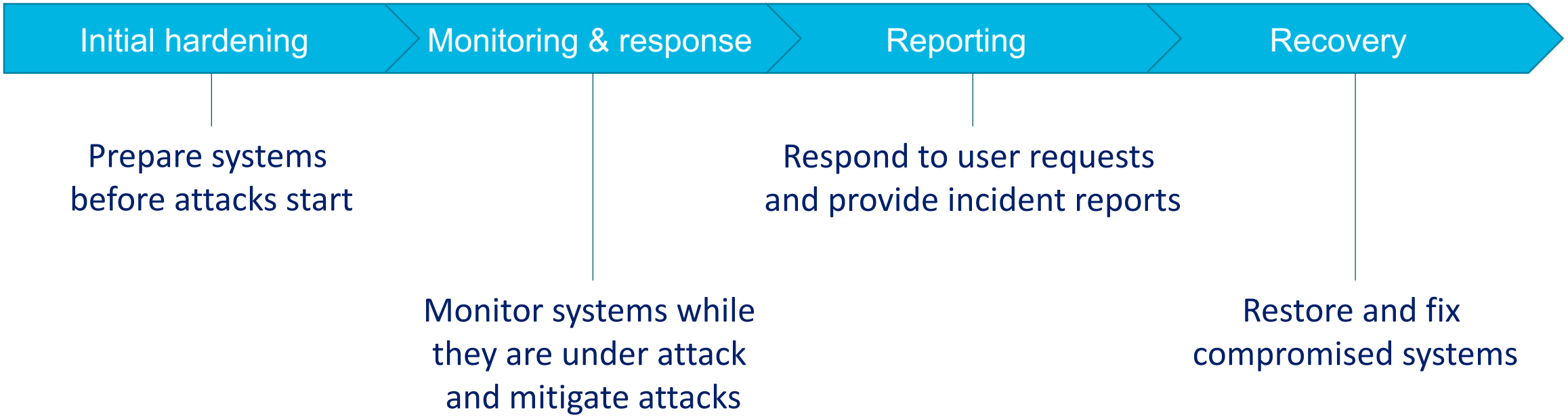
CYCON

Four stages in the exercise



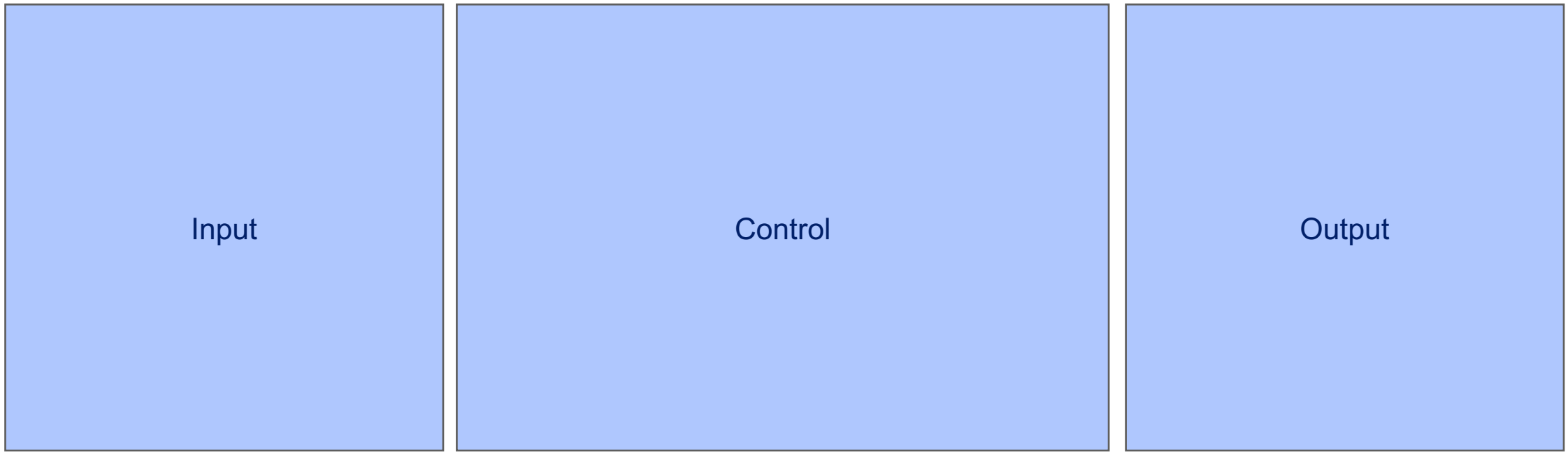
CYCON

Four stages in the exercise



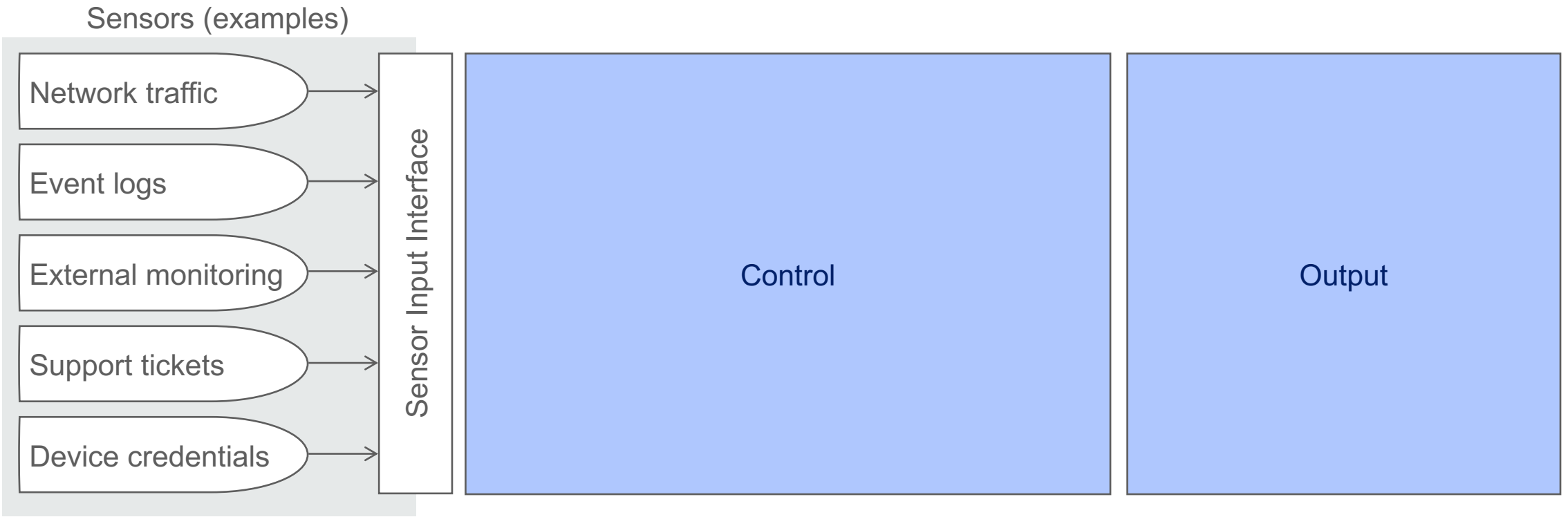
CYCON

Automated Blue Team framework overview



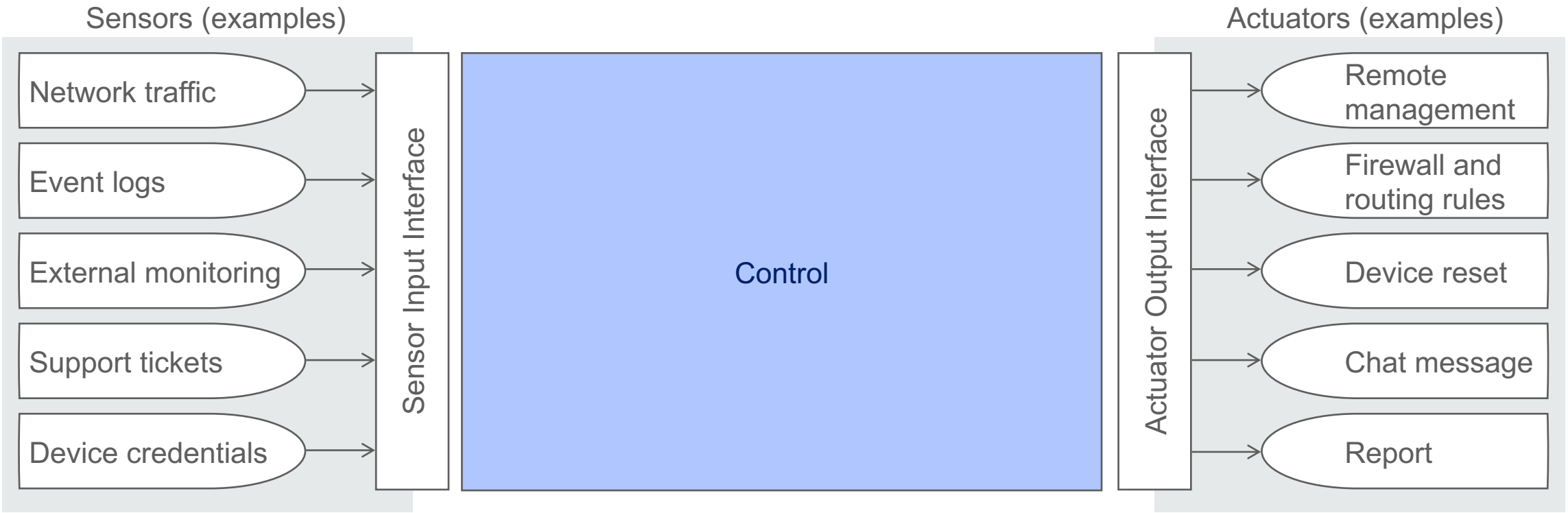
CYCON

Automated Blue Team framework overview



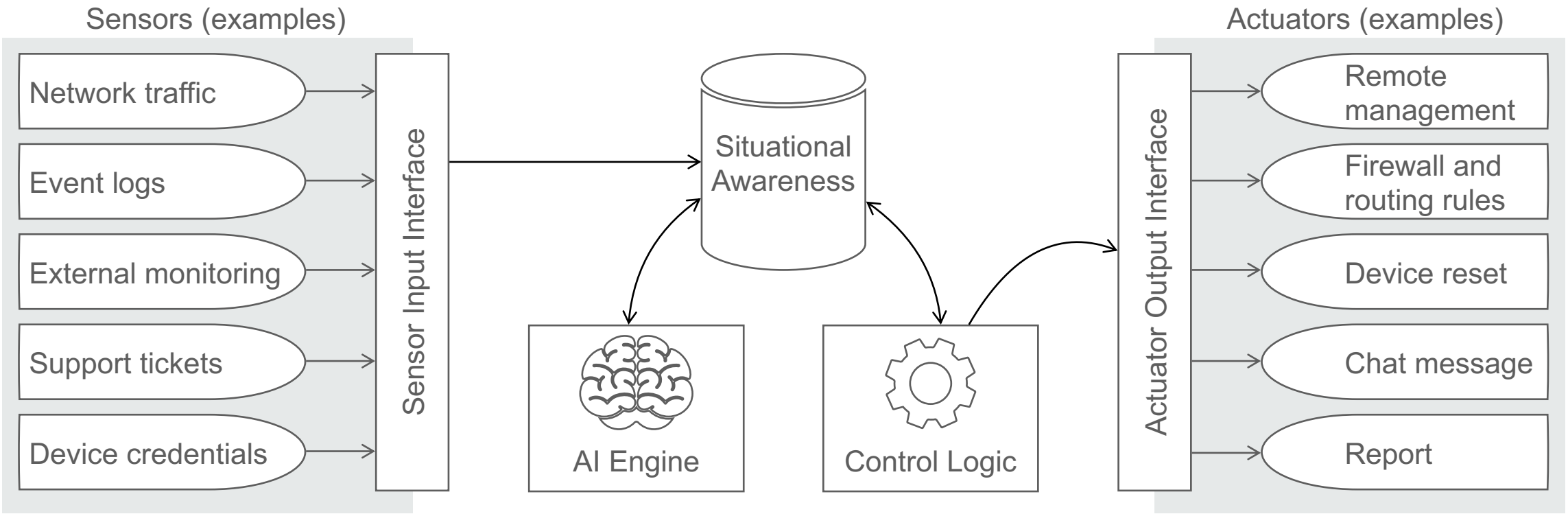
CYCON

Automated Blue Team framework overview



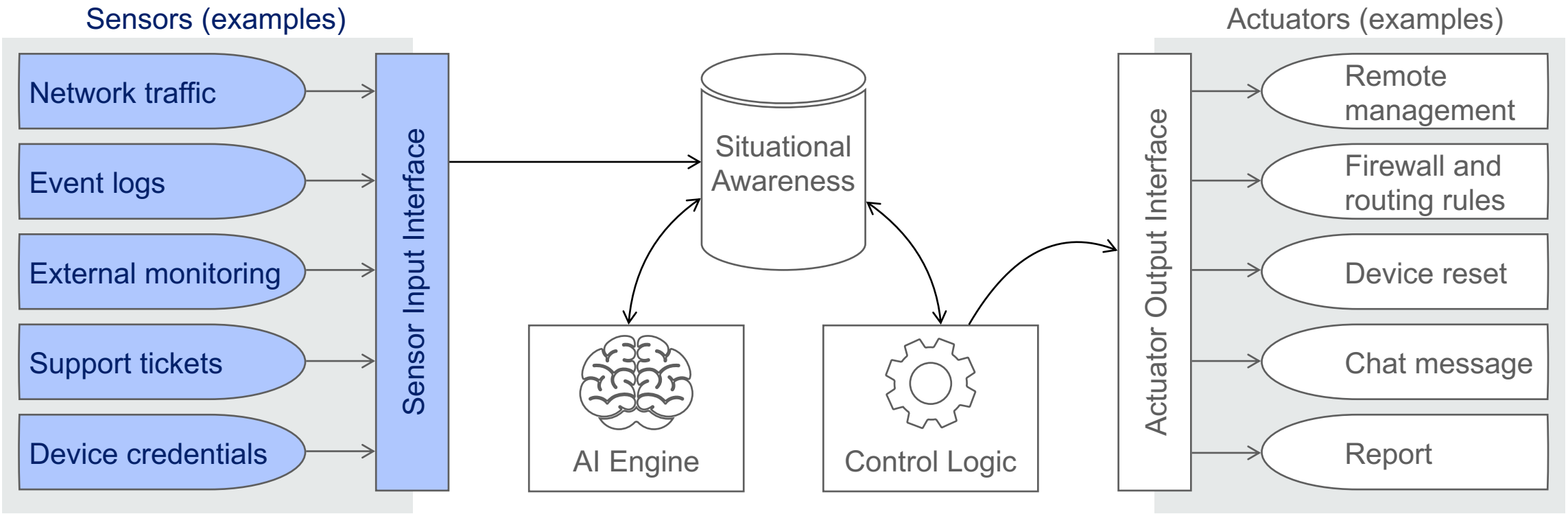
CYCON

Automated Blue Team framework overview



CYCON

Automated Blue Team framework overview



CYCON

There are five points where sensors can collect information

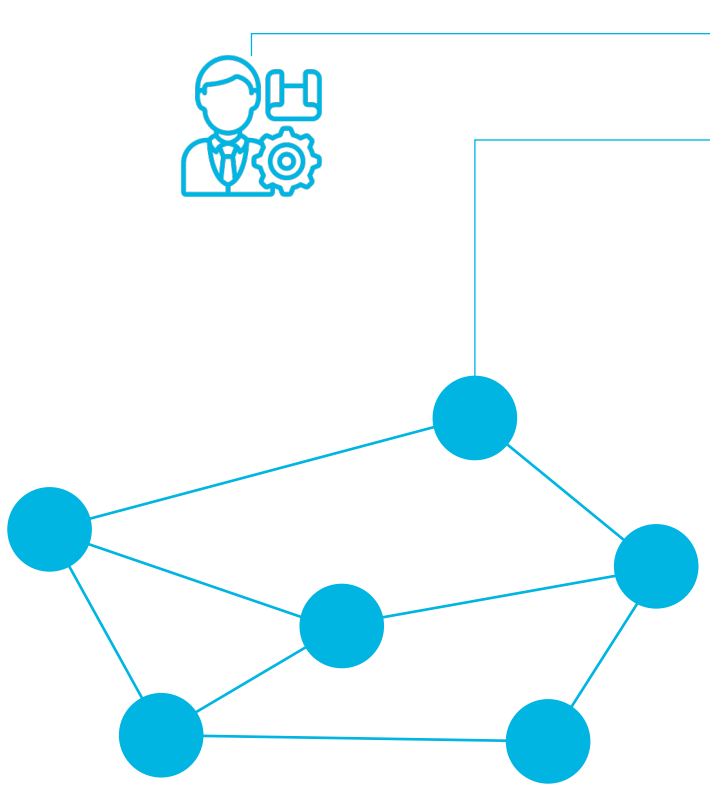
CYCON

There are five points where sensors can collect information

Organizational inputs

CYCON

There are five points where sensors can collect information

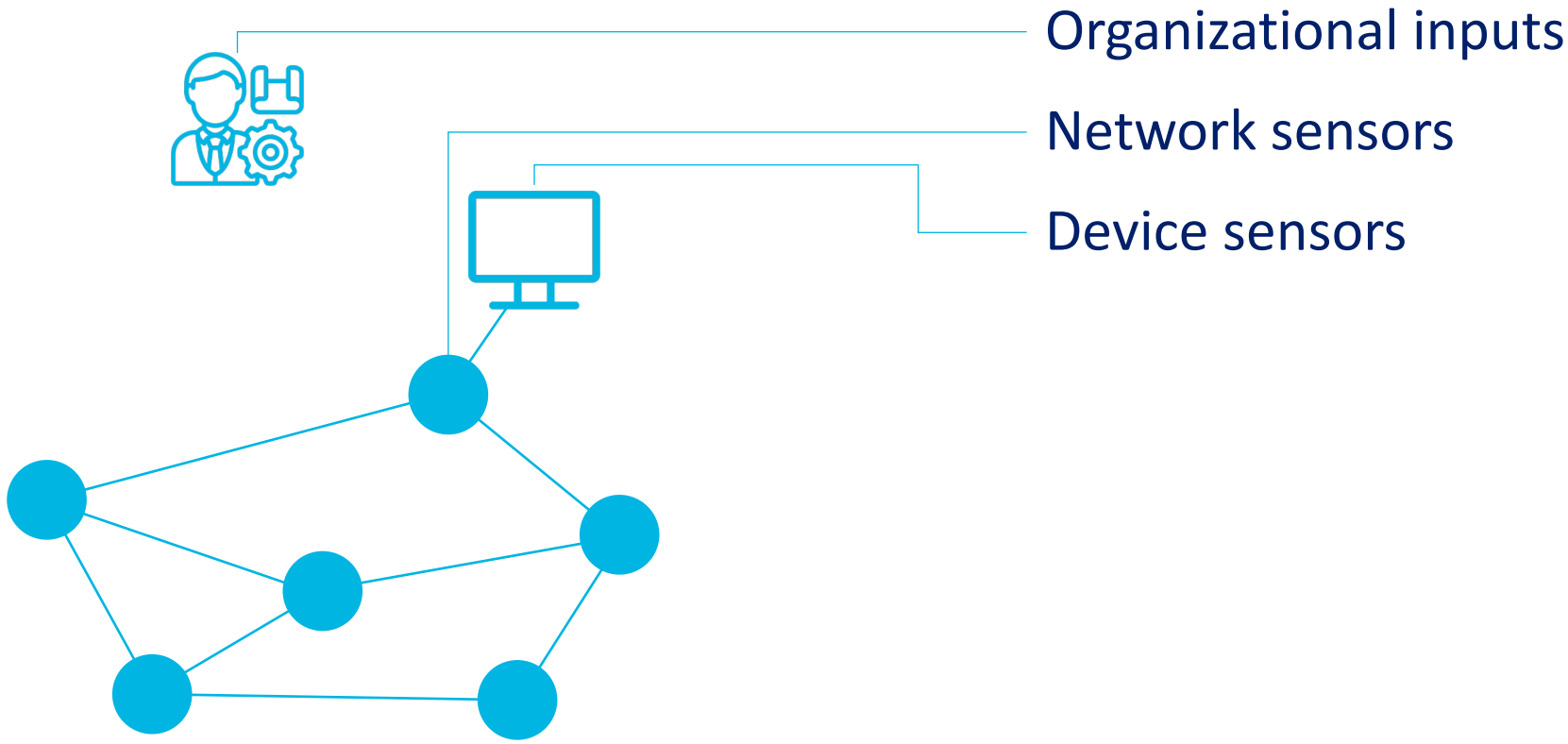


Organizational inputs

Network sensors

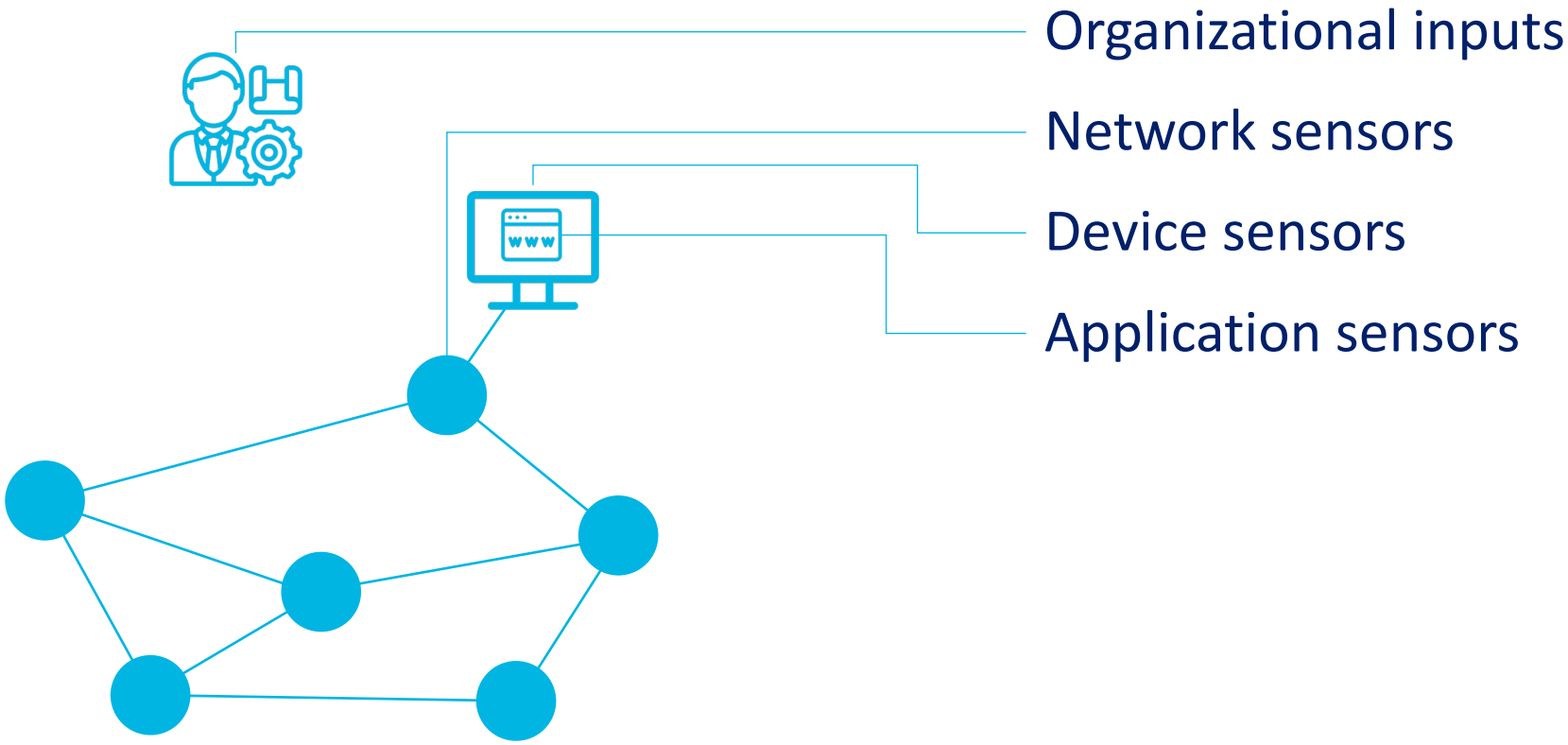
CYCON

There are five points where sensors can collect information



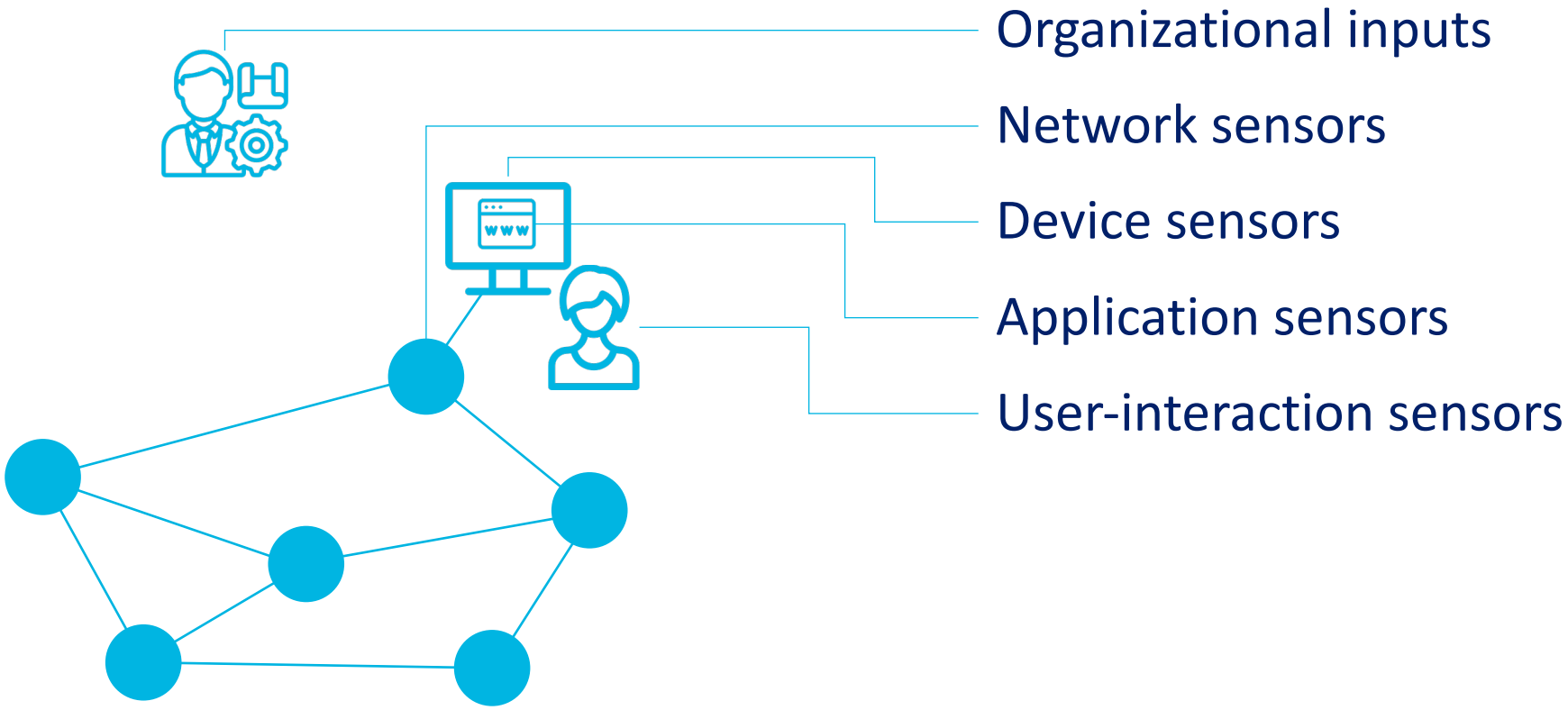
CYCON

There are five points where sensors can collect information



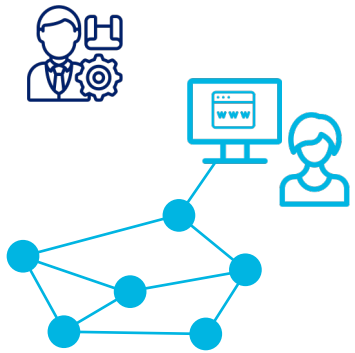
CYCON

There are five points where sensors can collect information



CYCON

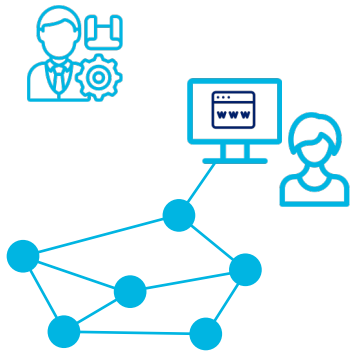
Organisational inputs define the environment and the objectives



- System documentation
  Devices and services running in the Gamenet
- Credentials
  To access and configure devices
- Scoring system
  Defines the objectives of the Blue Team
- External monitoring
  Information about the current performance

CYCON

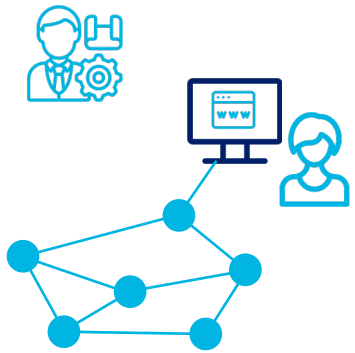
Application-level sensors provide information about running applications



- Application configuration
  E.g., users and permissions
- Application state
  E.g., update status, access logs

CYCON

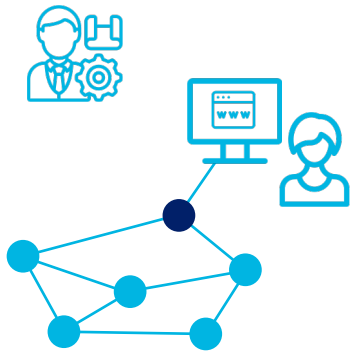
Device-level sensors report the status of all clients and servers



- System information
  E.g., running services, update status
- System logs
  E.g., login attempts, executed commands
- Active monitoring
  E.g., for additional logs, screen capture

CYCON

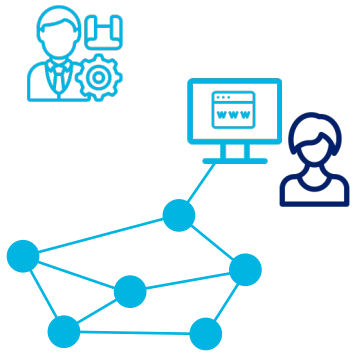
Network-level sensors provide access to network traffic



- Network configuration
  E.g., topology, open firewall ports
- Network state
  E.g., current load, reachability
- Network traffic
  Capture for inspection

CYCON

User-interaction sensors receive chat and email messages

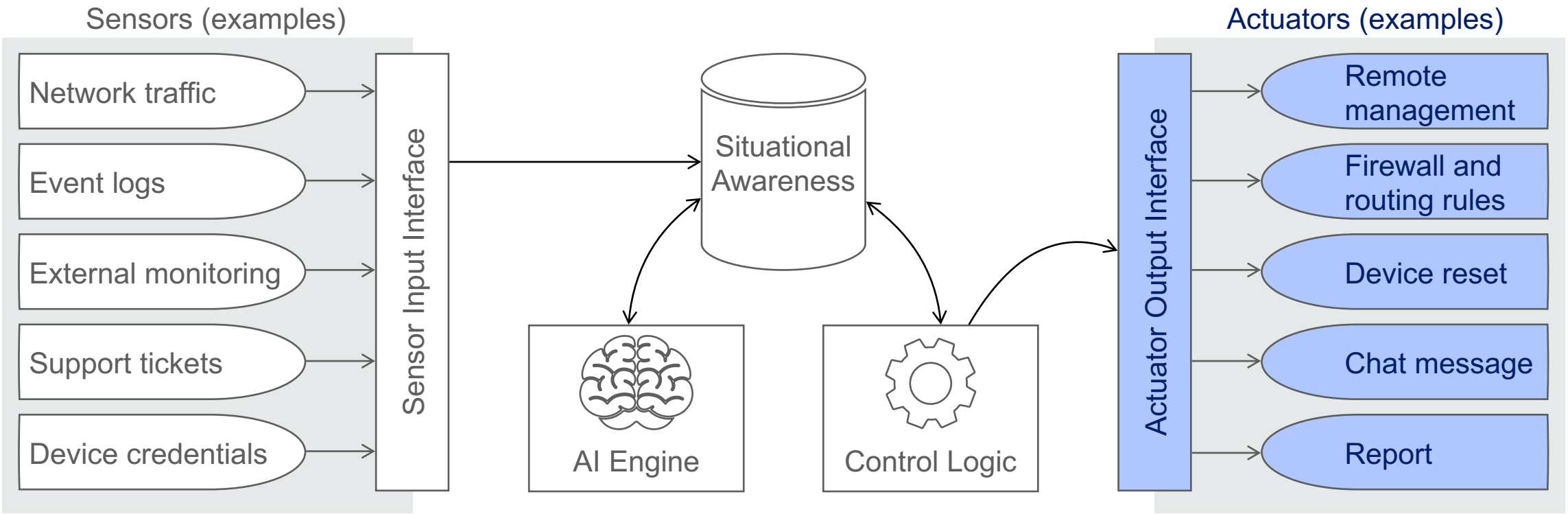


- Support tickets

  Notifications about observed problems

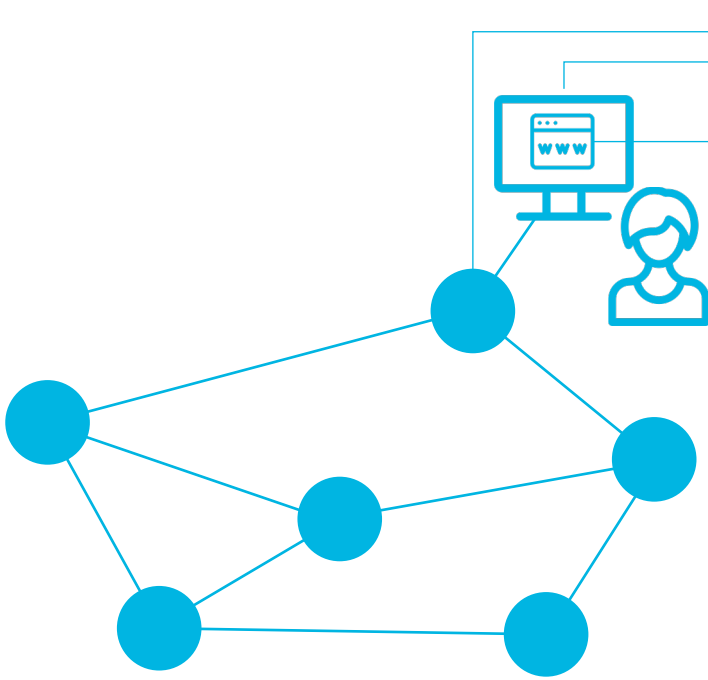
CYCON

Overview



CYCON

There are four points where actuators can act



Network actuators

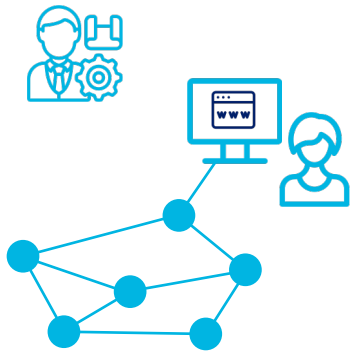
Device actuators

Application actuators

User-interaction actuators

CYCON

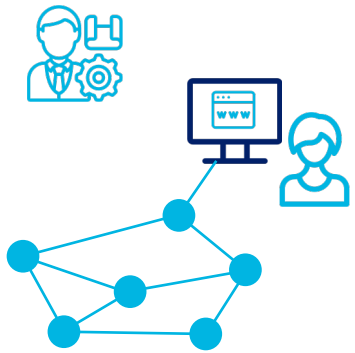
Application-level actuators



- Application firewall
  E.g., block malicious uploads in web application firewall
- Application configuration
  E.g., restrict access
- Application patching
  E.g., install updates

CYCON

Device-level actuators



- Remote management

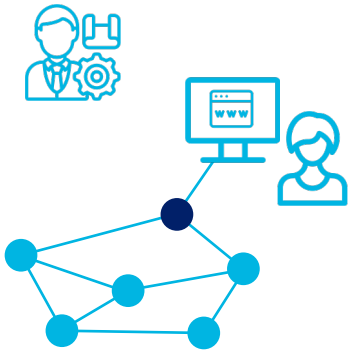
  Log in to every device and run commands

- Restore initial state

  Revert a device to a previous state
  (e.g., from a backup)

CYCON

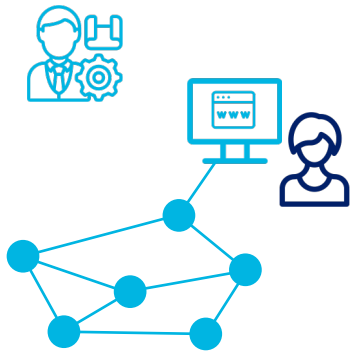
Network-level actuators



- Network configuration
  E.g., firewall rules

CYCON

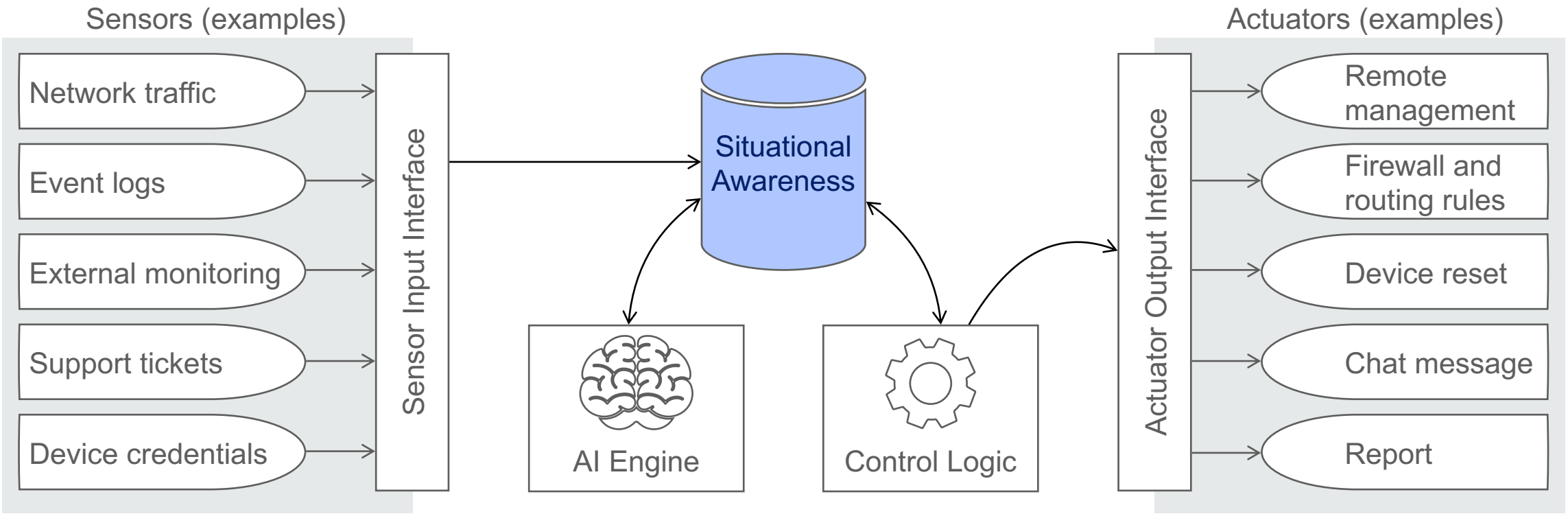
User-interaction actuators



- Chat and e-mail
  Respond to messages in a human-readable format
- Incident reports
  Report successful and failed attacks to the Green Team

CYCON

Overview



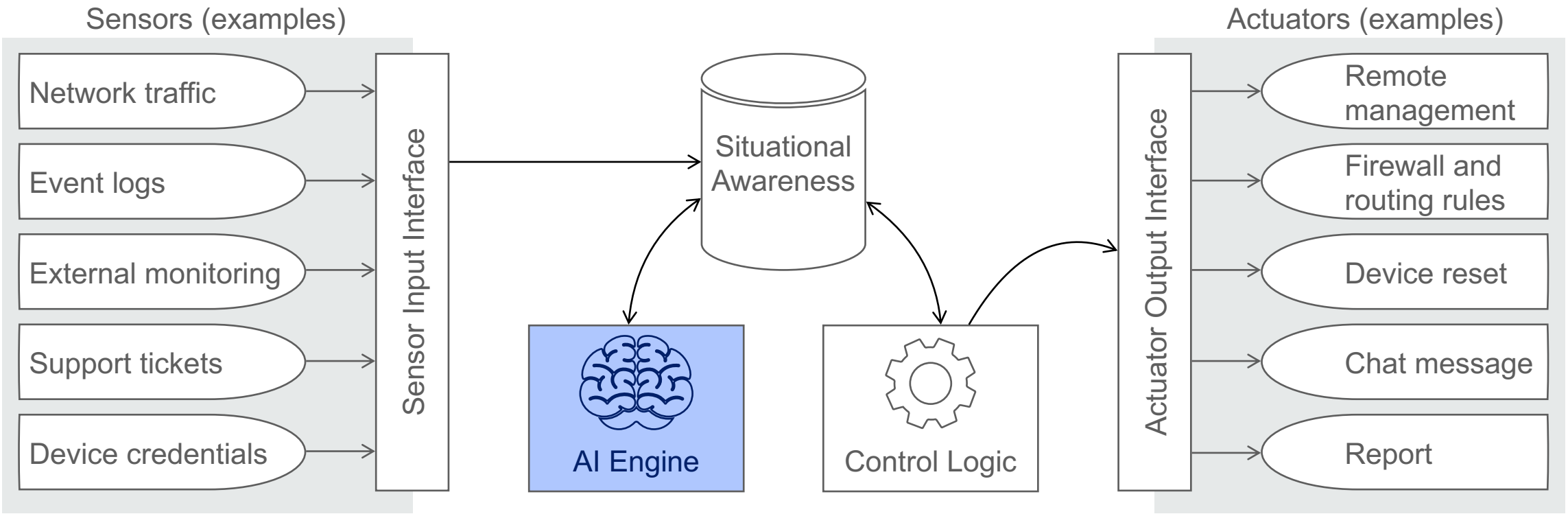
CYCON

The situational awareness database contains all the relevant information

- Sensor data
  E.g., access logs
- AI models
  E.g., anomaly detection
- Actuator outputs
  E.g., incident reports
- External information
  E.g., known software vulnerabilities

CYCON

Overview



CYCON

Four levels of AI

CYCON

Four levels of AI

Level 0: Reactive narrow AI

Signature-based decisions

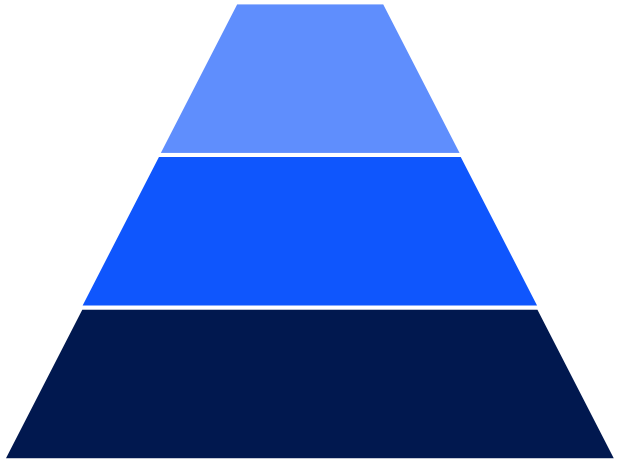
CYCON

Four levels of AI

Level 1: Limited-memory narrow AI	“Machine learning” today
Level 0: Reactive narrow AI	Signature-based decisions

CYCON

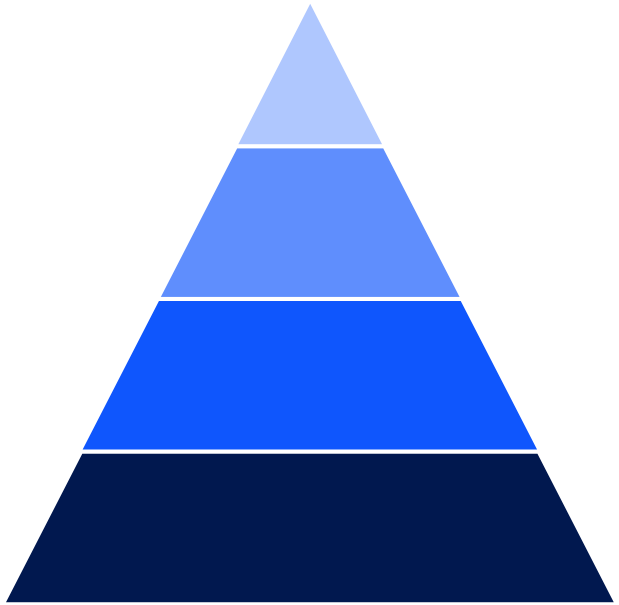
Four levels of AI



Level 2: General AI	Mimics human intelligence
Level 1: Limited-memory narrow AI	“Machine learning” today
Level 0: Reactive narrow AI	Signature-based decisions

CYCON

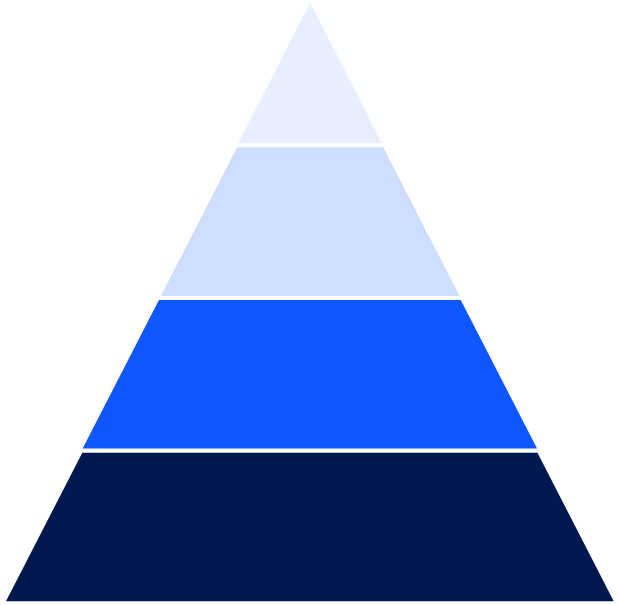
Four levels of AI



Level 3: Super AI	Surpasses human intelligence
Level 2: General AI	Mimics human intelligence
Level 1: Limited-memory narrow AI	“Machine learning” today
Level 0: Reactive narrow AI	Signature-based decisions

CYCON

Only level 0 and level 1 exist today



Level 3: Super AI	Surpasses human intelligence
Level 2: General AI	Mimics human intelligence
Level 1: Limited-memory narrow AI	“Machine learning” today
Level 0: Reactive narrow AI	Signature-based decisions

CYCON

Five tasks for AI

Identification / classification	What is it?
Categorisation	What belongs together?
Assessment	What is important?
Recommendation	What should be done?
Prediction	What will happen?

CYCON

AI for initial hardening

Identification / classification

Categorisation

Find groups of similar devices

Assessment

Recommendation

Prediction

CYCON

AI for monitoring and response

Identification / classification	Detect malicious network traffic
Categorisation	Detect malicious patterns in log files
Assessment	
Recommendation	
Prediction	

CYCON

AI for reporting

Identification / classification	
Categorisation	Link support tickets and monitoring alerts
Assessment	Prioritise support tickets
Recommendation	Formulate response to support tickets
Prediction	Predict impact on the scoring

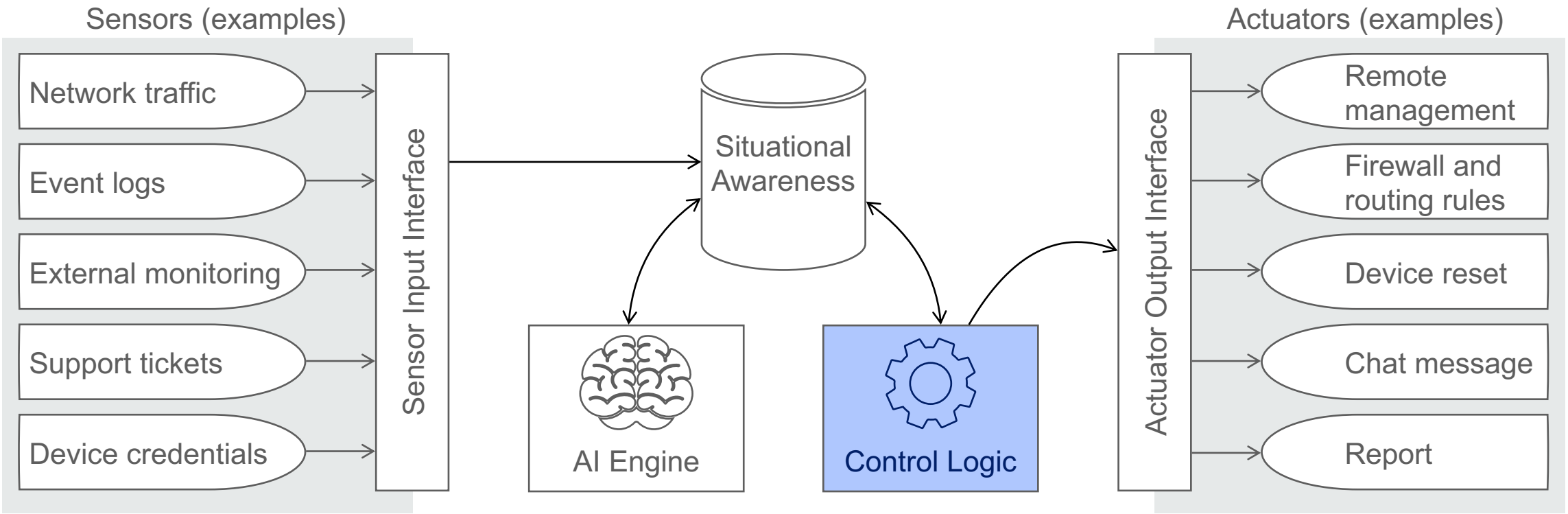
CYCON

AI for recovery

Identification / classification	Find devices that need to be recovered
Categorisation	Find similar devices as a template for recovery
Assessment	
Recommendation	
Prediction	

CYCON

Overview



CYCON

Control logic

CYCON

Actions for initial hardening

- Software deployment

  E.g., web application firewall, antivirus, monitoring agents

- Patch systems

  E.g., install updates, find backdoors

- Secure initial configuration

  E.g., change passwords, add firewall rules

CYCON

Actions for monitoring and response

- Disable exploited services
  E.g., compromised websites
- Monitor and block malicious system calls
  E.g., process creations or opening of files
- Monitor and block malicious network connections
  E.g., command and control flows

CYCON

Actions for reporting

- Generate human-readable reports

  E.g., about mitigated and successful attacks

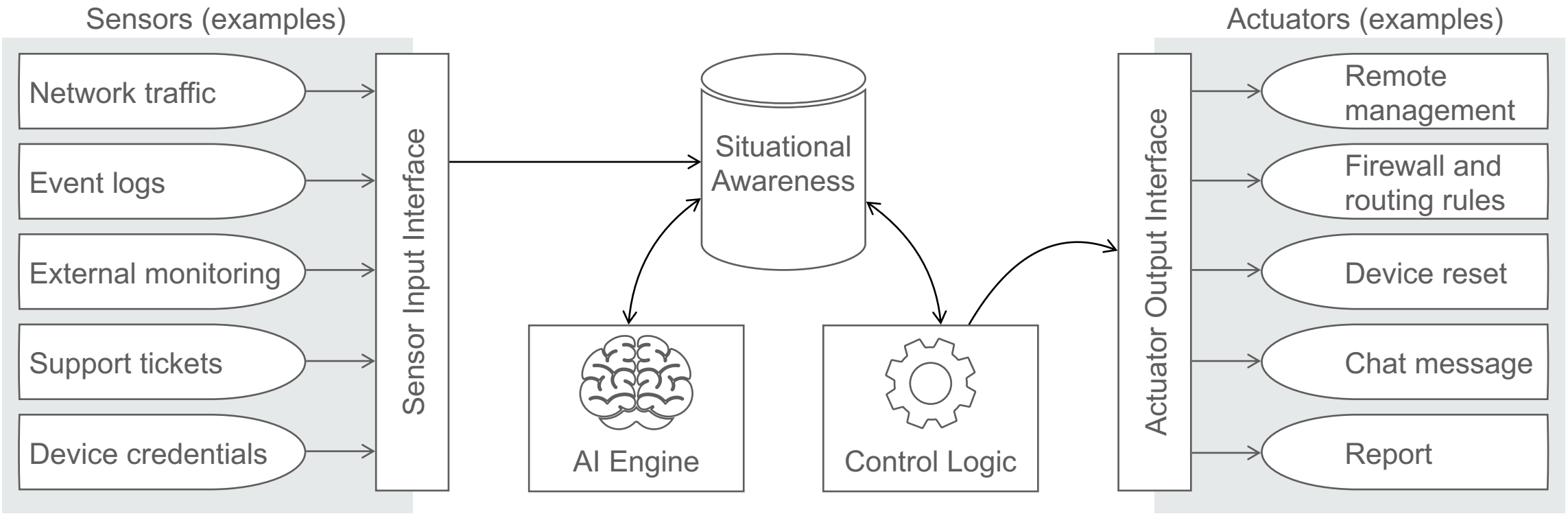
CYCON

Actions for recovery

- Self-recovery from previous backups
  E.g., database snapshots
- Revert devices to their initial state
  E.g., special systems without backup possibilities

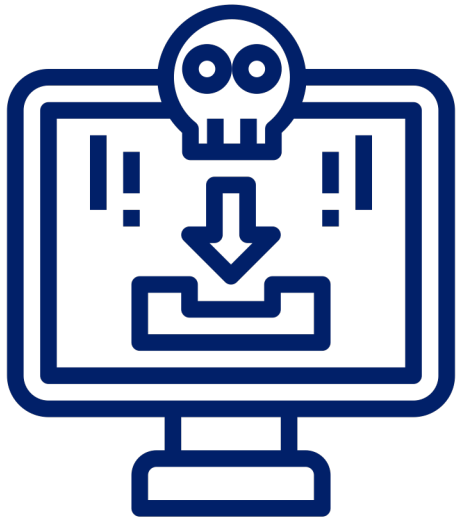
CYCON

Overview



CYCON

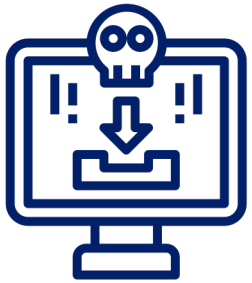
Case study: A user mistakenly downloads and executes a malicious file



CYCON

Case study

Initial hardening | Monitoring & response | Reporting | Recovery

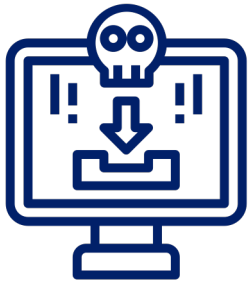


- Configure all clients to send HTTP(s) traffic via a proxy
- Enable detailed logging and send logs to a central server
- Set up recording of all network traffic and feature extraction

CYCON

Case study

Initial hardening | **Monitoring & response** | Reporting | Recovery

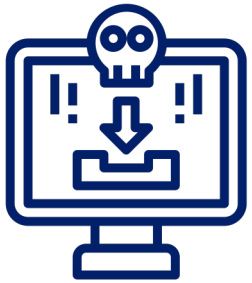


- Proxy detects the malicious payload
- Logging reports the execution of an unknown file
- Sniffer detects connection to C&C server

CYCON

Case study

Initial hardening | Monitoring & response | Reporting | Recovery

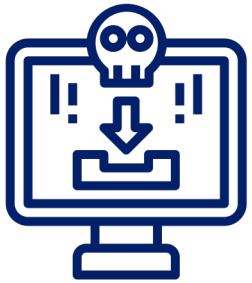


- Proxy detects the malicious payload → Remove payload
- Logging reports the execution of an unknown file → Block execution
- Sniffer detects connection to C&C server → Drop packets

CYCON

Case study

Initial hardening | Monitoring & response | **Reporting** | Recovery

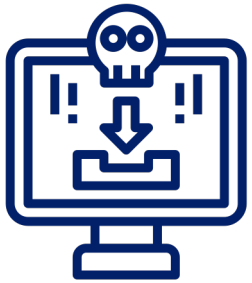


- Human-readable report with information about the incident
  - Malware source (compromised webserver)
  - Malware type
  - …

CYCON

Case study

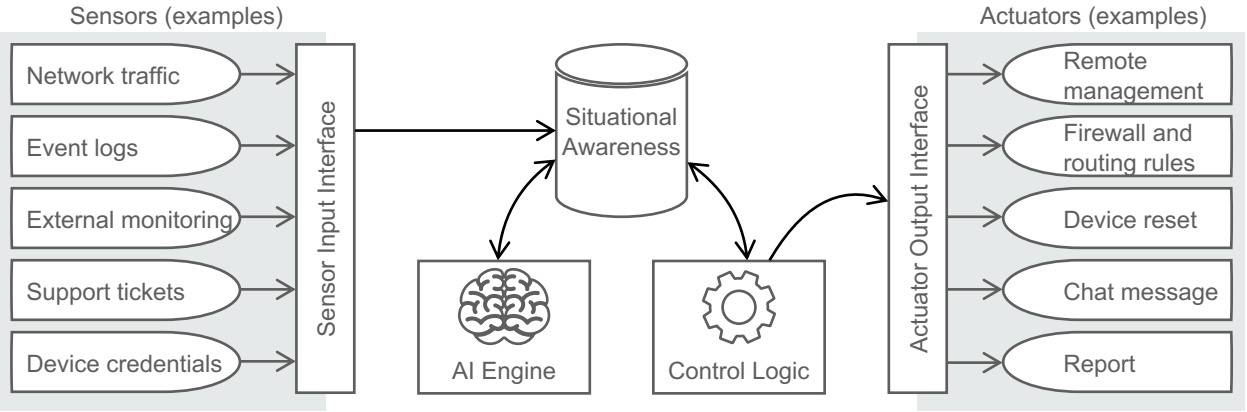
Initial hardening | Monitoring & response | Reporting | **Recovery**



- Restore a device if the malware was executed
  E.g., from a backup

CYCON

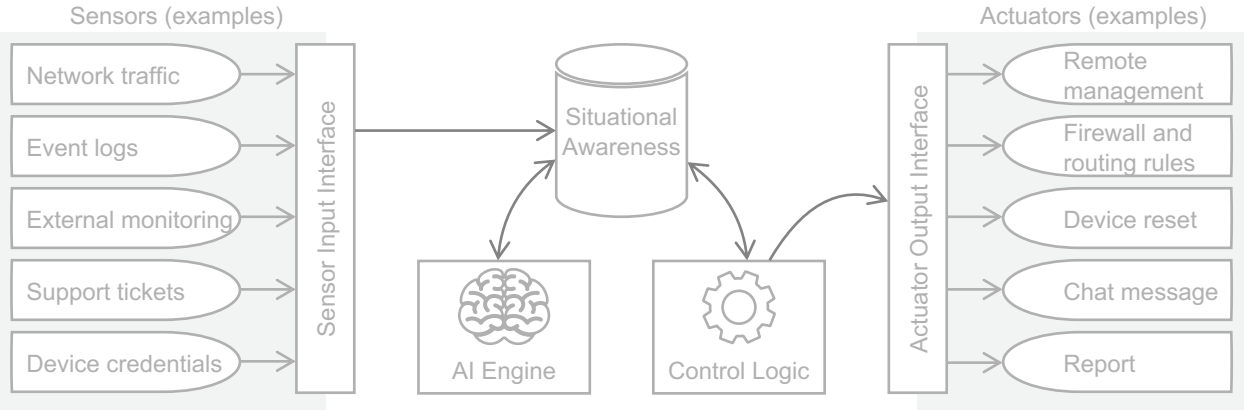
What comes next?



- Implementation
- Evaluation in Locked Shields 20xx
- Adaptation for other exercises

CYCON

What comes next?



- Implementation
- Evaluation in Locked Shields 20xx
- Adaptation for other exercises

Questions?

CYCON

Roland Meier
meierrol@ethz.ch