

Needles and Haystacks

Using Machine Learning and Threat Intelligence to detect, prevent & mitigate advanced cyber-physical threats to the communications critical infrastructure of Europe

How do you go about finding a needle in a haystack?

It's pretty simple:
You bring a magnet

Agenda

- 1 New Threats: Advanced and persistent
- 2 Machine Learning: Platform, Challenges, Sources, Data, Algorithms
- 3 Threat Intelligence: Crowd-Sourced Knowledge
- 4 Detection of cyber-threats: Beyond Firewalls and IPSs
- 5 Prevention and mitigation of cyber-threats: Closing the circle by automation
- 6 Critical Infrastructure Operators: TELCOs
- 7 Resilience as a function of threat detection
- 8 RESISTO as a holistic approach to Resilience
- 9 Orange Romania: Research, Education, Innovation
- 10 Key Takeaways, Q&A

APTs

New threats

Advanced and Persistent

Advanced: uses sophisticated vulnerability exploits, 0-day exploits. Stealthy.

Persistent: external C2C system that continuously monitors target(s) and exfiltrates data from target(s)

Highly dependent on social engineering, surveillance, supply-chain compromise

Usually infects multiple end-points in large networks.

Requires organization-wide efforts to detect, stop & clean-up.

New threats

Advanced and Persistent

STAGES

Initial Compromise: use of social engineering, spear phishing, media infestation with zero-day malware

Foothold in the stronghold: create C2C controllable 'victim' end-point in organization's network

Reconnaissance: Look around, find other vulnerable 'victims', gather info about network architecture;

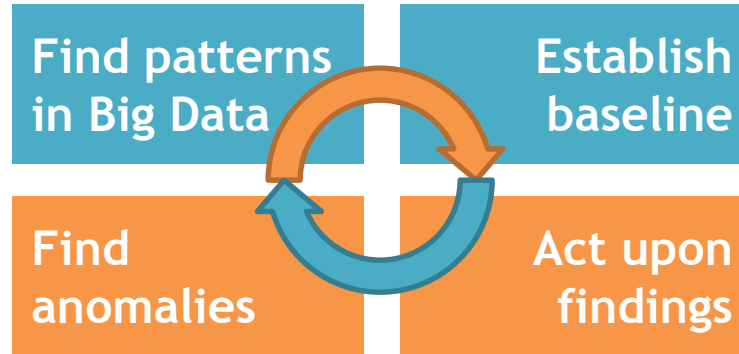
Lateral movement: expand control to other endpoints, servers, active equipment, perform data harvesting on them;

Keep a low profile: maintain presence, use stealthy / encrypted channels to communicate with C2C;

Complete mission: exfiltrate data, delete all tracks

Machine Learning

- **Machine learning:** “*Getting computers to act without being explicitly programmed*” (Stanford University)
- **Machine learning:** “*(...)algorithms can figure out how to perform important tasks by generalizing from examples*” (University of Washington)



Machine Learning

Where's this *big data* coming from?

- We generate a huge amount of data with various logical structures within multiple formats
- We generate “raw” unstructured data
- We have dozen of devices and technologies that analyze and correlate specific log and flow data. Most of them are proprietary. They speak they're own language. And Syslog 😊
- We discuss threats using natural language
- We disseminate using language constructs that computers usually don't understand (i.e. - posting about threats on social media platforms like twitter, using sarcasm etc.)

Metaphorically speaking, we're losing needles in haystacks when it comes to “blips on the radar” that might lead to something big

Machine Learning

What do we add to the mix?



Cellular Data
Wi-Fi Data
Security Data
Device metrics



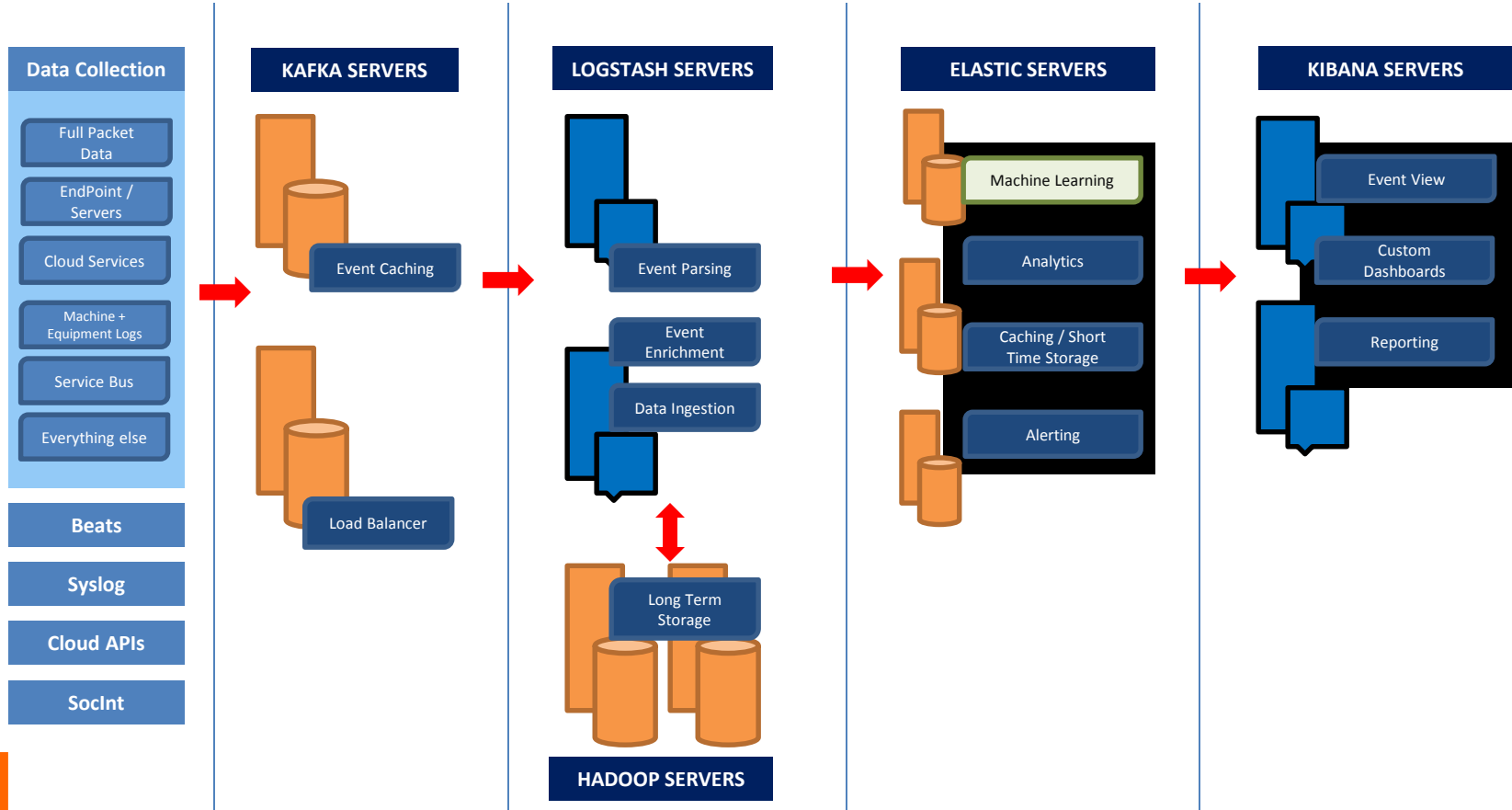
Threat Intelligence Feeds
Malware Analysis
Threat Hunting



Vulnerability Scans
CVEs & Zero Day Feeds
Pentest Results

Machine Learning

Architecture - The Big & Mean data-crunching machine ... is actually virtualized and lean ☺



Machine Learning

The challenges

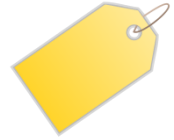
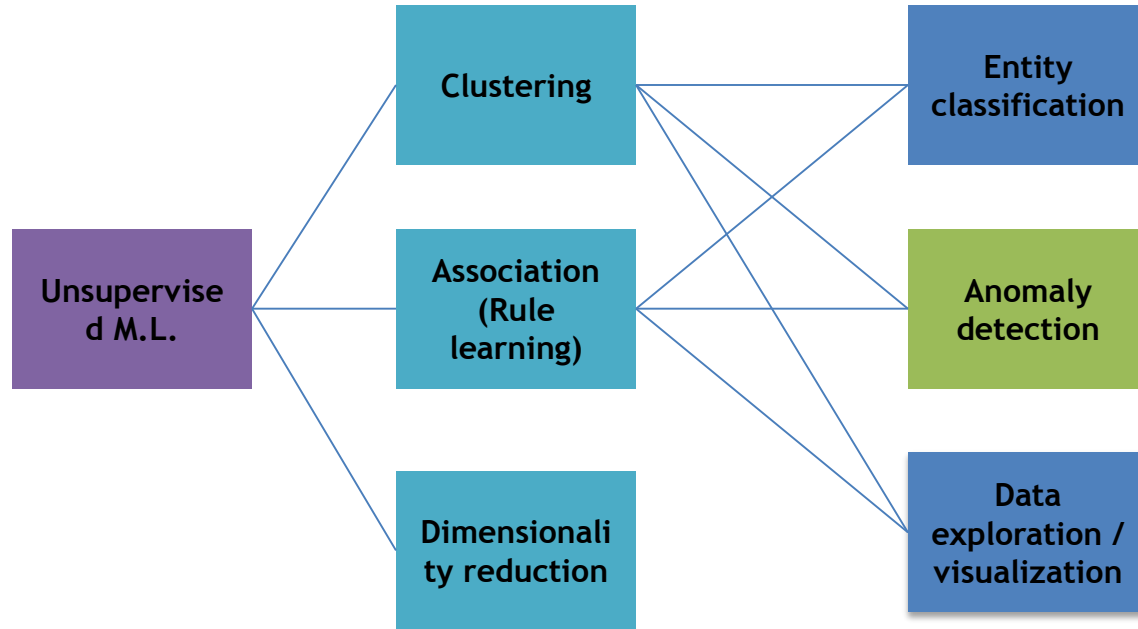
1. Deep Learning / Machine Learning are **JUST TOOLS**. They are NOT Silver Bullets.
2. For supervised M.L. you need reliable labeled sample data. And *reliable* is hard to find
3. The baseline is constantly moving. 'Moving averages' won't help. Expect **LOTS** of false positives, be prepared to invest in TLC for your Data Crunching setup
4. You **NEED** expert input and lots of context. The machines, by themselves won't turn dust to gold.



Machine Learning

Architecture - Unsupervised M.L.

CONTEXT

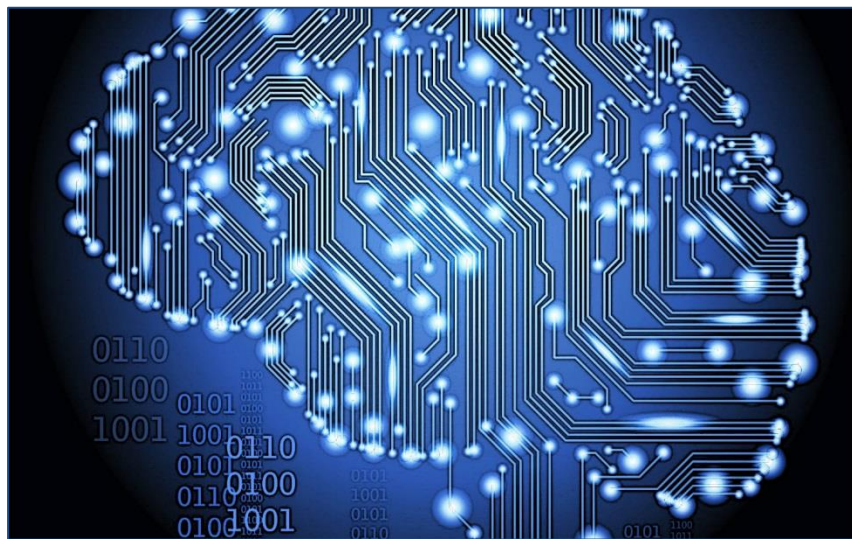


silvianstock - 189956502

Threat Intelligence

& Crowd-Sourced Knowledge

1. We're enriching data and adding context
2. We're integrating feeds from sources like **OTX** and **VirusTotal**
3. We're working on 'reading' tweets relevant to new threats. This is tricky as machines can't understand concepts like humor or irony. It's WIP 😊
4. We're developing a Threat Sharing platform with local key-players in the cybersec space, both business and public.



Threat Intelligence

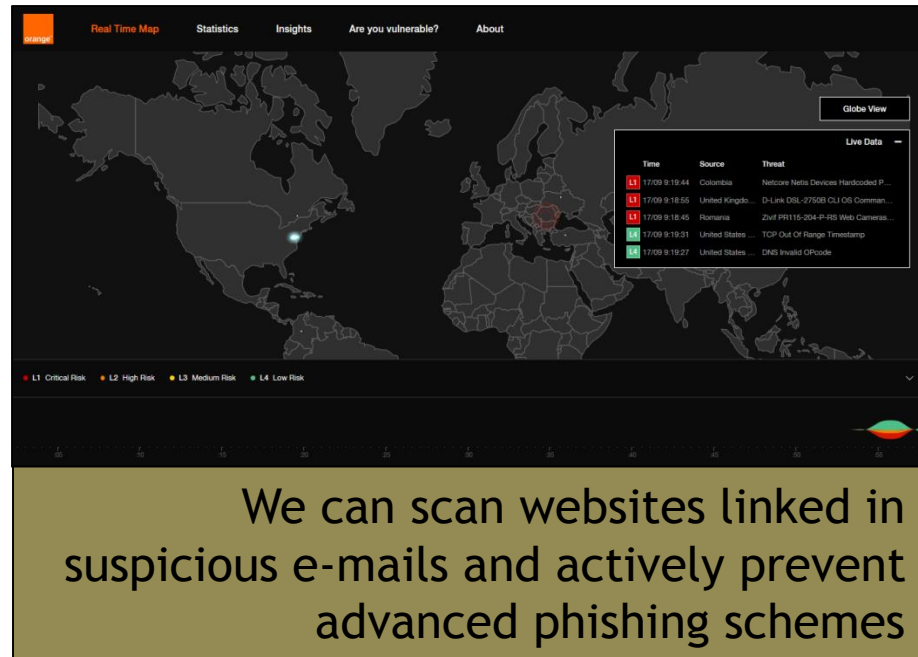
Web Sites & IoCs

Business Internet Security ThreatMap

We're using our own platform to monitor websites for IoCs

We're using 4 scanning engines:

- OWASP web vulnerability scanner
- CMS-specific vulnerability scanner
- APT - Watering Hole Malware scanner
- Previously Hacked? We'll interrogate the largest .ro database of hacked websites



Threat Detection

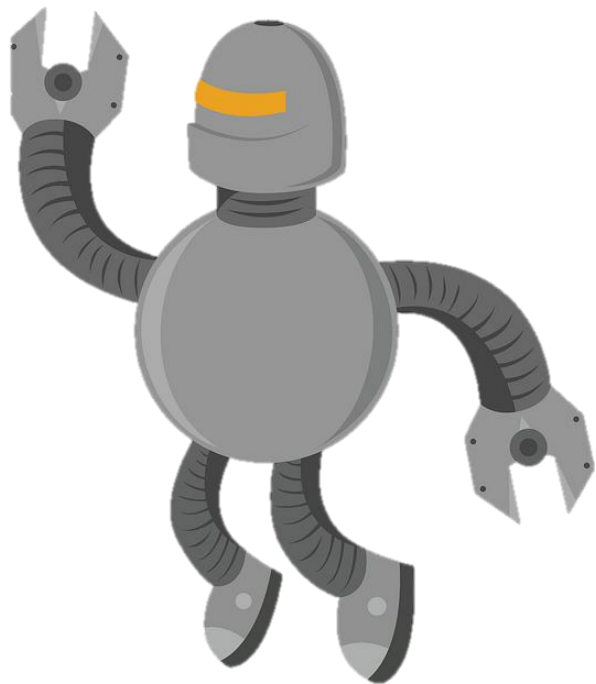
Beyond NGFWs and IPSs



1. Signature-based detection works for most known threats
2. It won't work for new threats and APTs
3. There's a window of opportunity for attackers from zero-day to patch. We're trying to close it.
4. If you don't have a signature for your new type of threat you must rely on behavior analysis.
5. We use sandboxes with near-instant spin-up and complete deletion after use. We monitor their output.

Prevention & Mitigation

Closing the circle by automation



We're faced with challenges:

1. Is there a barrier of trust we must cross, hand in hand with our Analytics machines?
2. Will we let the machine decide what, when and why block a connection, delete a file, quarantine an e-mail?
3. Trust **IS** a issue.
4. Testing before production is a must but in this case we need to cover as much ground as we can,

Critical Infrastructure Operators

What is Critical Infrastructure?

Critical Infrastructure is a term used by governing bodies to describe assets that are essential for the functioning of a society and economy

TELCOs as C.I.s

Telecommunication Operators can be assets regarded as C.I. operators because they provide services necessary for coordination and basic inter-human communication

SPECIFIC VULNs & THREATS

Large, complex technical infrastructure;

Provides services for current-gen and at least N-2 Gen devices (such as 2.5G / EDGE);

Heavily reliant on human decisions

Cyber-Physical Threats

Specific vulnerabilities and threats (SS7)

APTs relying on Social Engineering (SIM Hijacking)

Resilience as a function of Threat Detection

RESILIENCE

The ability (capacity) to provide and maintain an acceptable level of service in the face of **threats**, faults and challenges to normal operations.

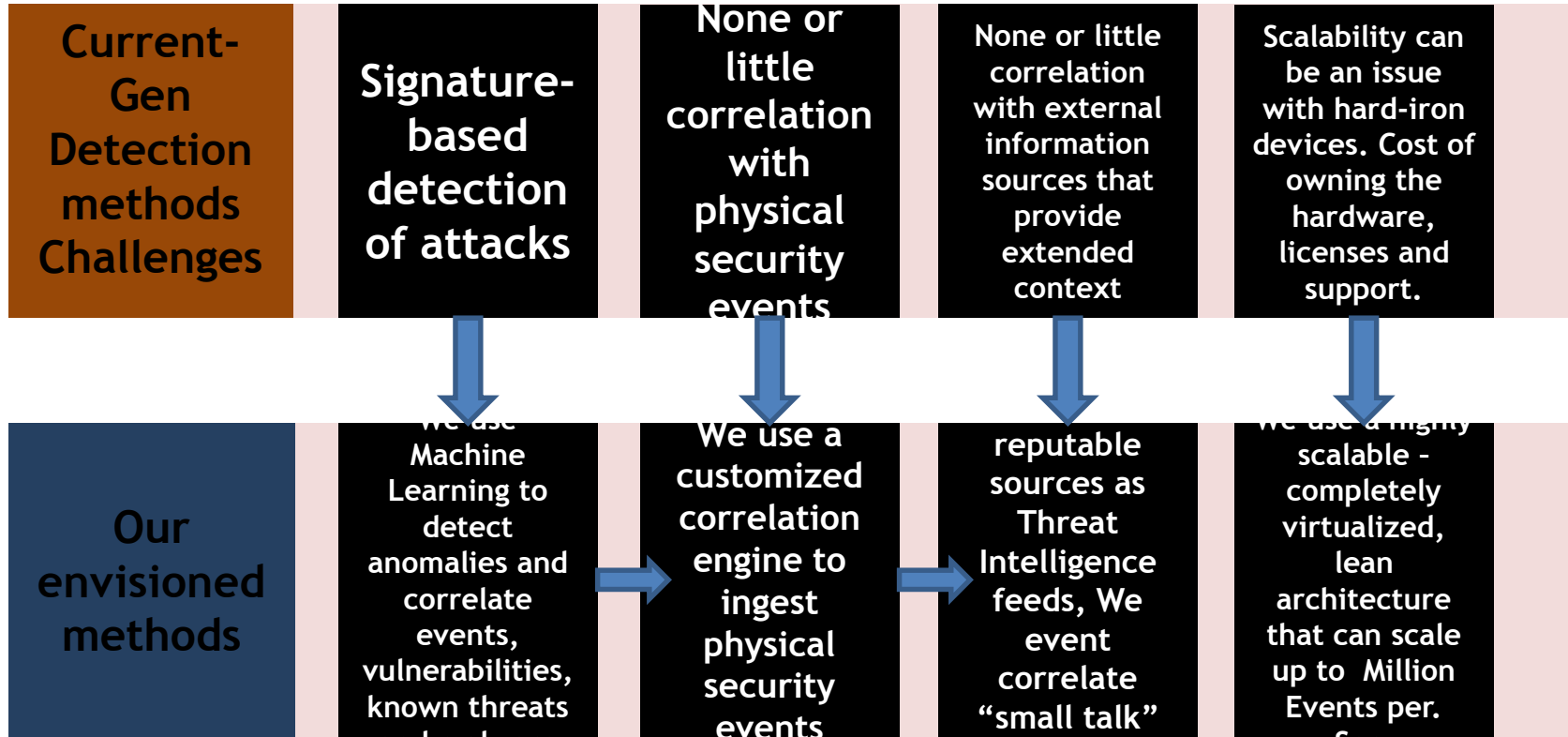
In order to increase resilience of a given system, the probable threats, challenges and risks have to be identified

THREAT DETECTION

- Key Factor for resilience
- Must adapt to specific threats
- Reliable in detecting complex and persistent threats

A word cloud on a black background featuring various terms related to resilience and threat detection. The words are in different colors (yellow, green, blue, purple) and sizes. The most prominent words are 'Analytic', 'Monitoring', 'Response', 'Adaptive', 'Positioning', and 'Integrity'. Other visible words include 'Restriction', 'Unpredictability', 'Segmentation', 'Diversity', 'Privilege', 'Dynamic', 'Substantiated', 'Non-persistence', and 'Deception'.

Resilience as a function of Threat Detection

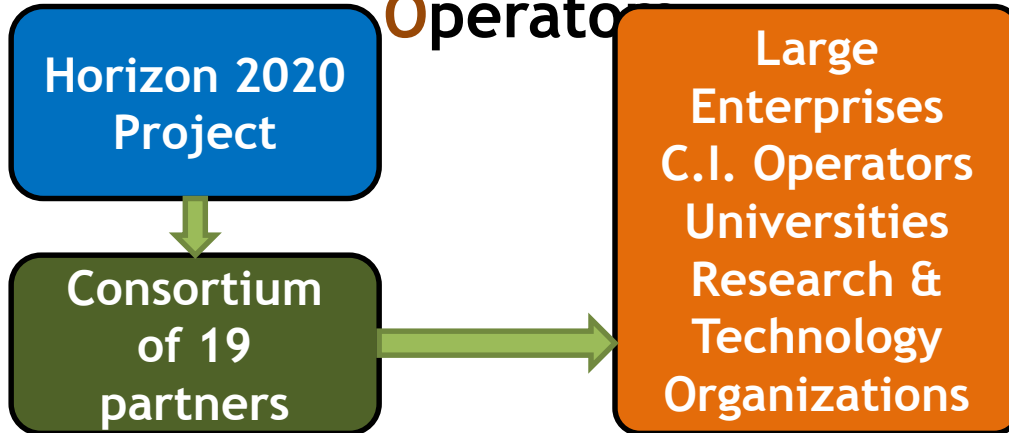




A holistic approach to resilience enhancement

RESilience enhancement and risk control platform for communication infraSTructure

Operator



- Holistic Approach to Situation Awareness
- Innovative Risk & Resilience & Improvement Process Management
- Decision Support System
- Protection against cyber-physical threats
- Modeled on state-of-the art technologies (Machine Learning, IoT, Block chain, Airborne Threat Detection, Holistic A-V analytics)

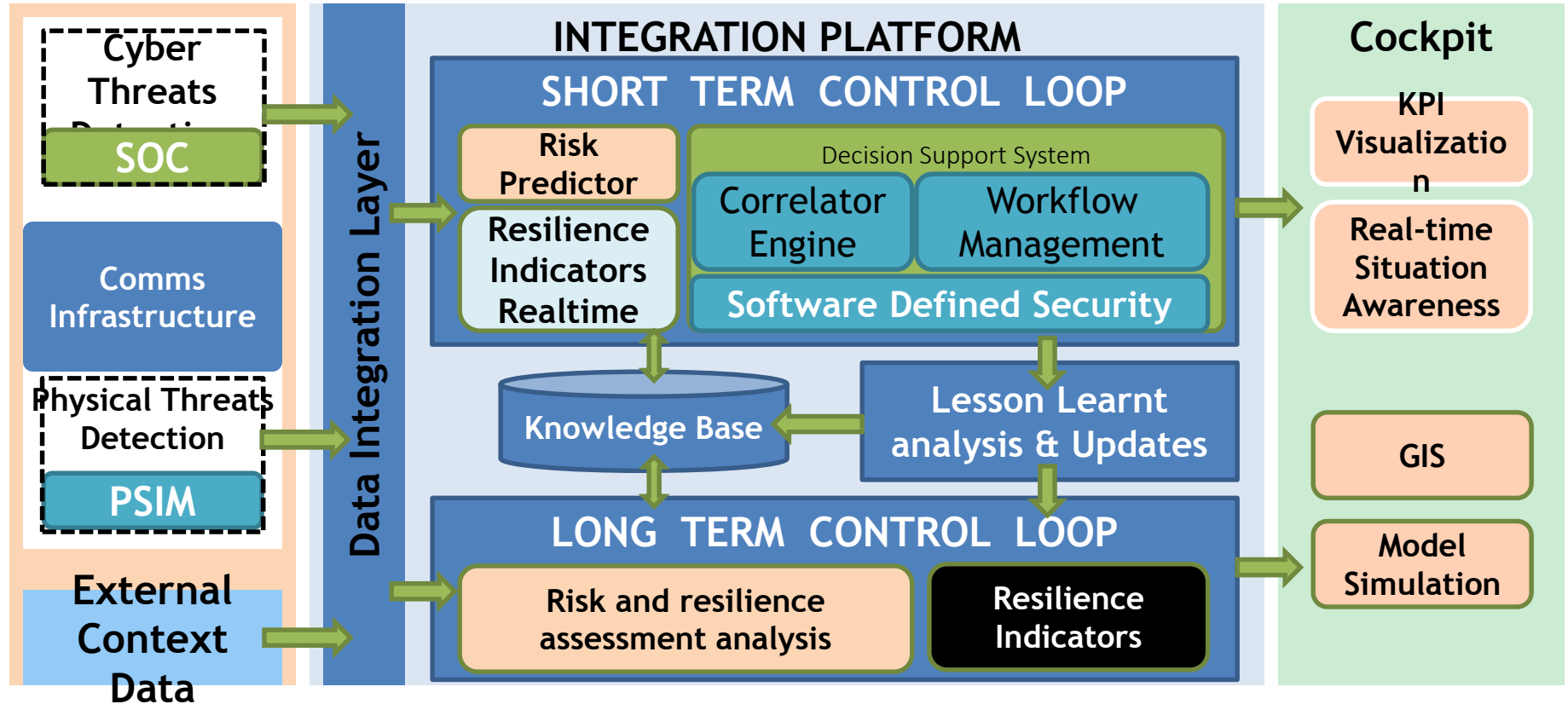
RESIST

Key Innovation Areas

1. Enlarged Threat Landscape considered (**Cyber / Physical / Cyber + Physical**)
2. Holistic Approach to System Modelling
3. Integrated Risk and Resilience management
4. Convergence of **PSIM** and **Cyber Protection Technology**
5. Perspective - New challenges posed by 5G evolution (**IoT/IoE, LPWAN**)
6. New Technology for detection, protection and response (**blockchain, drones, machine learning algorithms, software defined security**)
7. Cyber Intelligence

RESISTO

Architecture overview



Orange

Education, Research, Innovation

Orange Fab

Startups in cyber security, future of life and smart territories.
From secure remote access to tissue-printing 3D printers

Orange Education Program

We are mentoring and training 40 scholars, students of ETTI, providing scholarships, internships, bachelors and masters degree project coordination

Horizon 2020 Projects

We are participating in Consortiums researching new technologies, tools and methods in the IoT, 5G, Cyber-Security & Resilience, Computer Forensics,

Thanks 😊