4th International Conference on Operational Planning, Technological Innovations and Mathematical Applications (OPTIMA)

SHIELD

SECURING AGAINST INTRUDERS AND OTHER THREATS THROUGH A NFV-ENABLED ENVIRONMENT

[H2020 - Grant Agreement No. 700199]

SHIELD– Securing against intruders and other threats through a NFV-enabled environment

Dr. Antonis Litke
Dr. Nikos Papadakis
Dimitris Papadopoulos

Cybersecurity agenda

The EU is determined to safeguard an online environment providing the highest possible freedom and security, for the benefit of everyone.

The EU Objectives

- Increasing cybersecurity capabilities and cooperation
- Making the EU a strong player in cybersecurity

The EU strategy

- Increasing cyber resilience;
- Drastically reducing cybercrime;
- Developing EU cyber defence policy and capabilities related to the Common Security and Defence Policy (CSDP);
- Developing the industrial and technological resources for cybersecurity;
- Establishing a coherent international cyberspace policy for the EU and promote core EU values

The motivation for SHIELD

Lack of open-source tools for cybersecurity leveraging massive analytics capabilities

Huge momentum of open technologies for big data

Requirement for expensive, specialized hardware for information security (high CAPEX)

Emergence of the "Security as-a-Service" paradigm, based on cloud and NFV

Proposed solution

SHIELD was awarded with a **15/15** score.

Security as a Service based on NFV+SDN architecture

Virtualized Network Security Functions (**vNSF**)

Trustworthiness

Pervasive Trust Computing in VM and Containers

Big Data engine & **Machine Learning** capabilities

Real-time incident detection and mitigation

Goal

- Create an IDP platform that:

Retrieves information from vNSFs deployed at strategic locations of the network

Transmits such information to be properly **processed by Big Data engines**

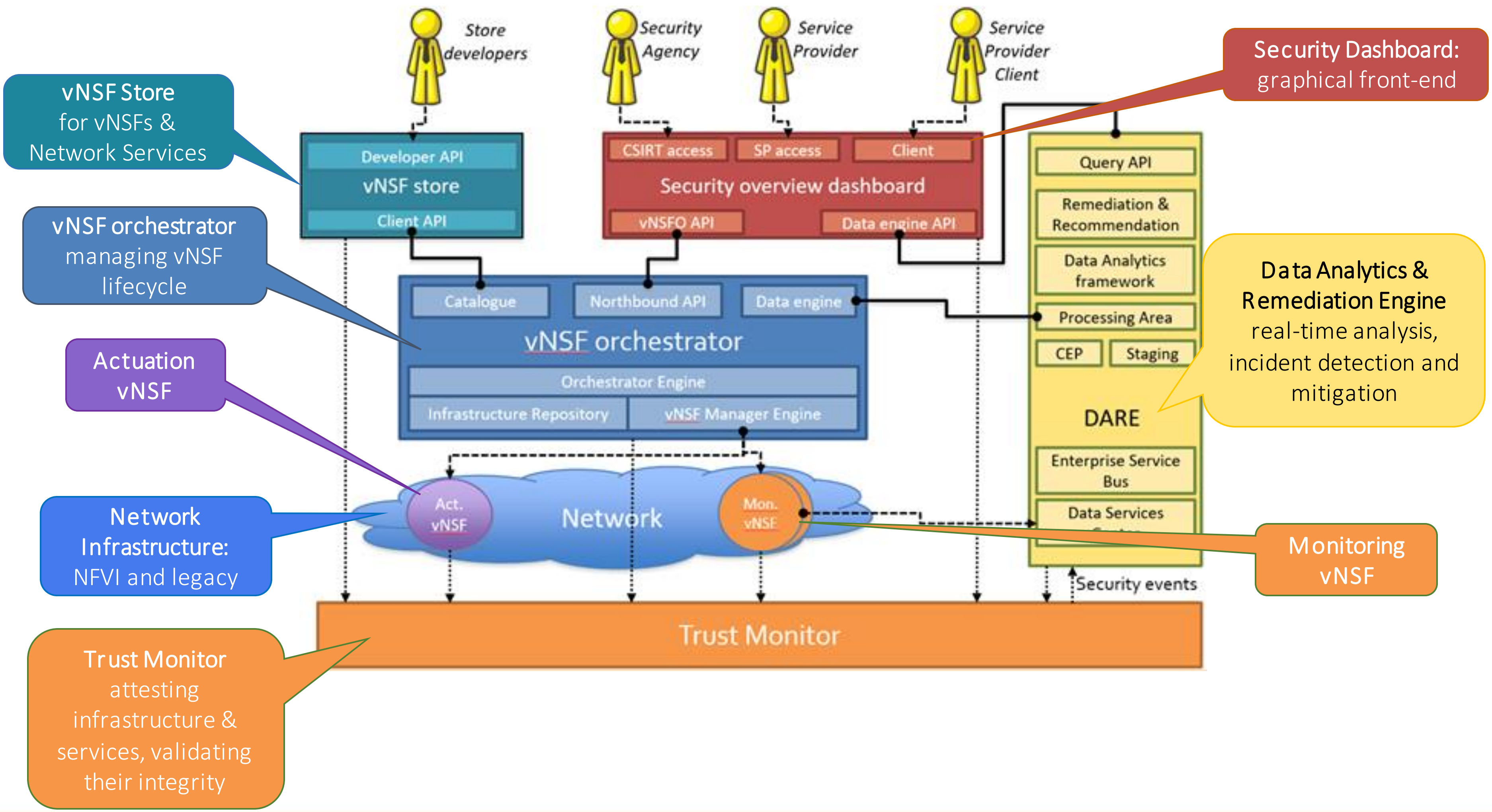
Visualises info and **recommends actions** by means of a dashboard and accessible API

Acts on the network by **taking effective counter-measures**

Provides **flexible support** for both new security capabilities and reconfiguration of existing security

Is built on top of **attested hardware**

SHIELD High Level Architecture



Creation of an information-driven, Data Analytics and Remediation Engine DARE

The engine will store and analyze heterogeneous network information, previously collected via monitoring vNSFs. It will combine techniques related to big data, data analysis and cognitive learning altogether.

The DARE will use machine learning and threat monitoring techniques to:

- Detect threats with pattern discovery techniques.
- Exploit feedback to update cybersecurity data topologies.
- Perform remediation activities by triggering vNSFs.

• Data Analytics engine

It will primarily include two modules that will work in parallel:

- a **cognitive Data Analysis module** implementing open-source technologies (Apache Hadoop, Spark, Mahout etc.) for big data analysis.
- a proprietary **security Data Analysis module** including algorithms for the detection of malicious network behavior.

The engine will be open for the future inclusion of additional data analysis modules.

Focusing on our main task: Cognitive analytics

Cognitive Data Analytics module

a Cognitive Data Analysis module

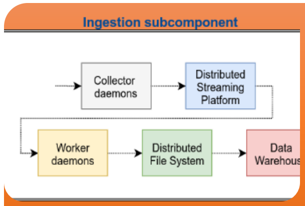
An entity of elements that is able to produce packet and flow analytics by using scalable machine-learning techniques.

- state-of-the-art **big data** solutions
- **distributed computing** technologies for batch and streaming processing
- **scalability** and **load balancing**
- utilization of **open data models** (ODM)
- **concurrent** running of multiple machine-learning applications on a single, shared, enriched data set
- **tailor-made** security analytics
- **attack prediction** by correlating network anomalies to specific threats

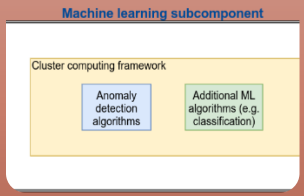
Based on the Apache Spot framework

Focusing on our main task: Cognitive analytics

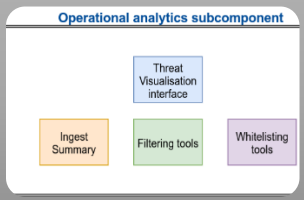
The Cognitive Data Analytics module contains 3 main subcomponents:



Ingestion



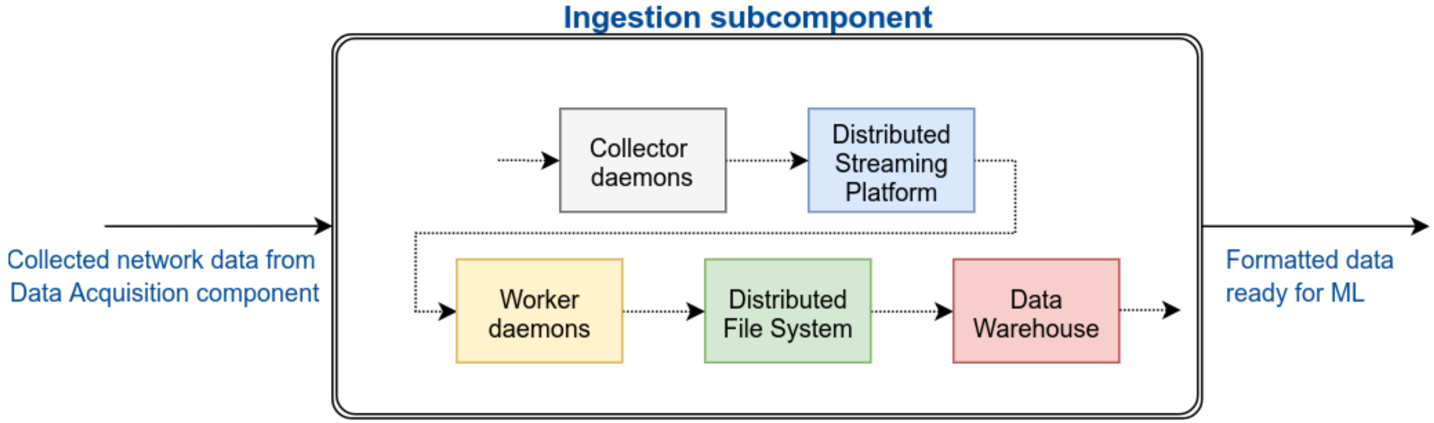
Machine learning



Operational Analytics

Focusing on our main task: Cognitive analytics

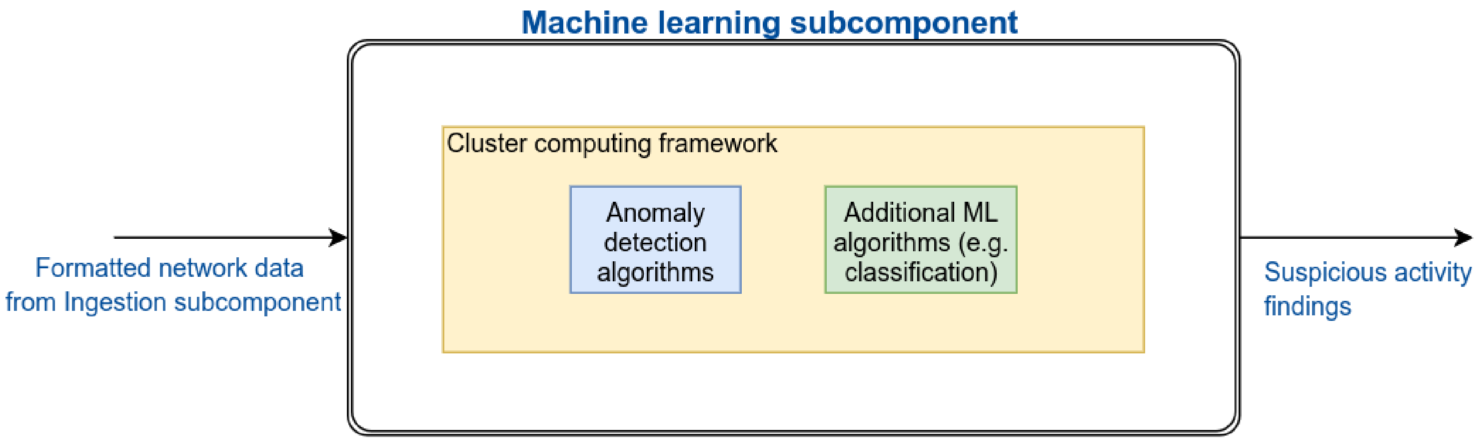
Ingestion



- **Collectors** detect new data (flow, DNS, proxy) generated by network tools.
- Data is sent to a Kafka **streaming platform** that splits it into specific topics and smaller partitions, while creating a data pipeline for each instance of the ingest process
- **Workers** subscribed to a specific topic and partition, are reading, parsing and storing the data in a specific distributed format to be consumed by the machine learning algorithms.
- Once the data has been transformed, it is **stored in HDFS** with the original format and it's made available to **Hive tables** so that it can be accessible by SQL queries.

Focusing on our main task: Cognitive analytics

Machine learning



The machine learning entity consumes a collection of network events in order to produce a list of the events that are considered to be the **least probable**, and these are considered the **most suspicious**.

- **Anomaly Detection algorithms:** Discovering abstract topics of events by examining network traffic and ultimately discovering normal and abnormal behaviour using a topic modelling algorithm (Latent Dirichlet Allocation - LDA).
- **Additional algorithms:** Correlating the detected anomalies with specific threats using classification algorithm.

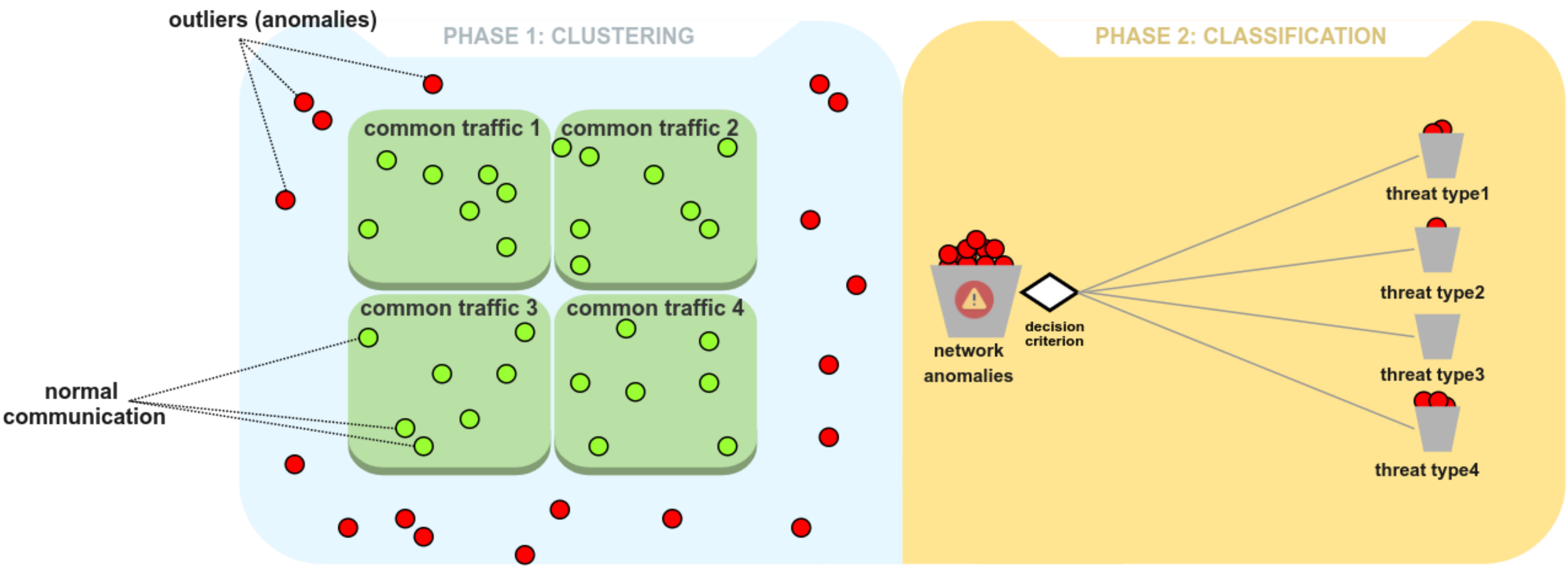
Focusing on our main task: Cognitive analytics

Machine learning

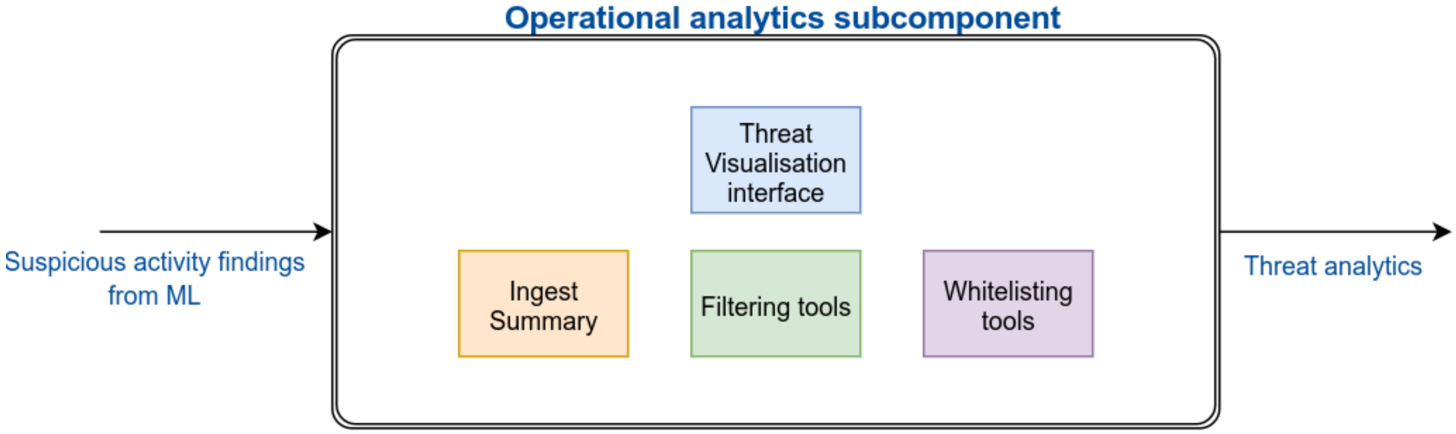
A two-phase approach:

Phase 1: Clustering of network traffic to detect anomalies

Phase 2: Classification of these anomalies to specific threats



Operational Analytics



Context enrichment, noise filtering, whitelisting and heuristics processes to produce a list of the most likely patterns.

- **Threat visualisation interface:** An interactive dashboard for a comprehensive view of the network anomalies
- **Ingest summary:** Presents the amount of network data that has been ingested on the cluster.
- **Filtering tools:** Visualisation tools for customized results based on time, source/destination, severity, type etc.
- **Whitelisting tools:** A set of tools to exclude some of the results, dealing with potential false-positives.

- Remediation engine
  - Uses the feedback from the data analysis modules to **activate remediation activities** (e.g. block traffic, isolate user, sandboxing etc.)
  - **Provides recommendations** to the users and collaborates with the vNSF orchestrator for the selection of the appropriate vNSF.
  - Performs in near **real-time**, using open-source solutions.

Key project milestones

Open-source release of the SHIELD system (alpha version):

September 2017

Open-source release of the SHIELD system (beta version):

September 2018

System tested & validated – Final release:

February 2019

Numbers of SHIELD

European R&D project

Co-funded by the EU under H2020 "Secure Societies" program

11 partners

4.56 M€ total budget

Sep 2016 – Feb 2019

Our team

Follow us!

https://www.shield-h2020.eu/

@shield_h2020

SHIELD EU Project

info@shield_h2020.eu

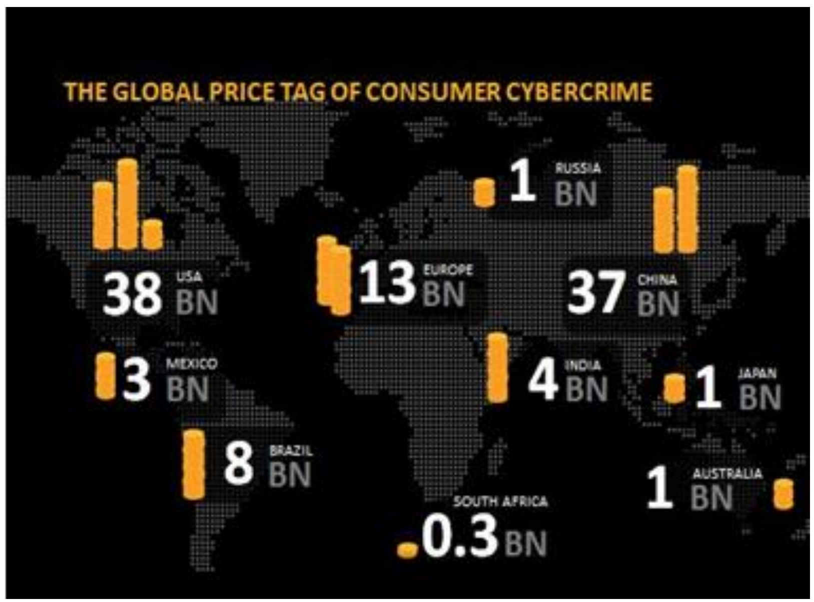
SHIELD has received financial support from the European Commission
under Grant Agreement No. 700199

Cybersecurity, a funded area

As part of the EU cybersecurity strategy, the European Commission and the European Cyber Security Organisation (ECSO) signed a cPPP on 5 July 2016.

- The EU will invest up to €450 million in this partnership, under its research and innovation programme Horizon 2020. Cybersecurity market players are expected to invest three times more.

European Commission | Horizon 2020 European Union funding for Research & Innovation



Numbers of SHIELD

1 Innovation Action

5 Components (TRL 6 and above)

30 Months

6 WP

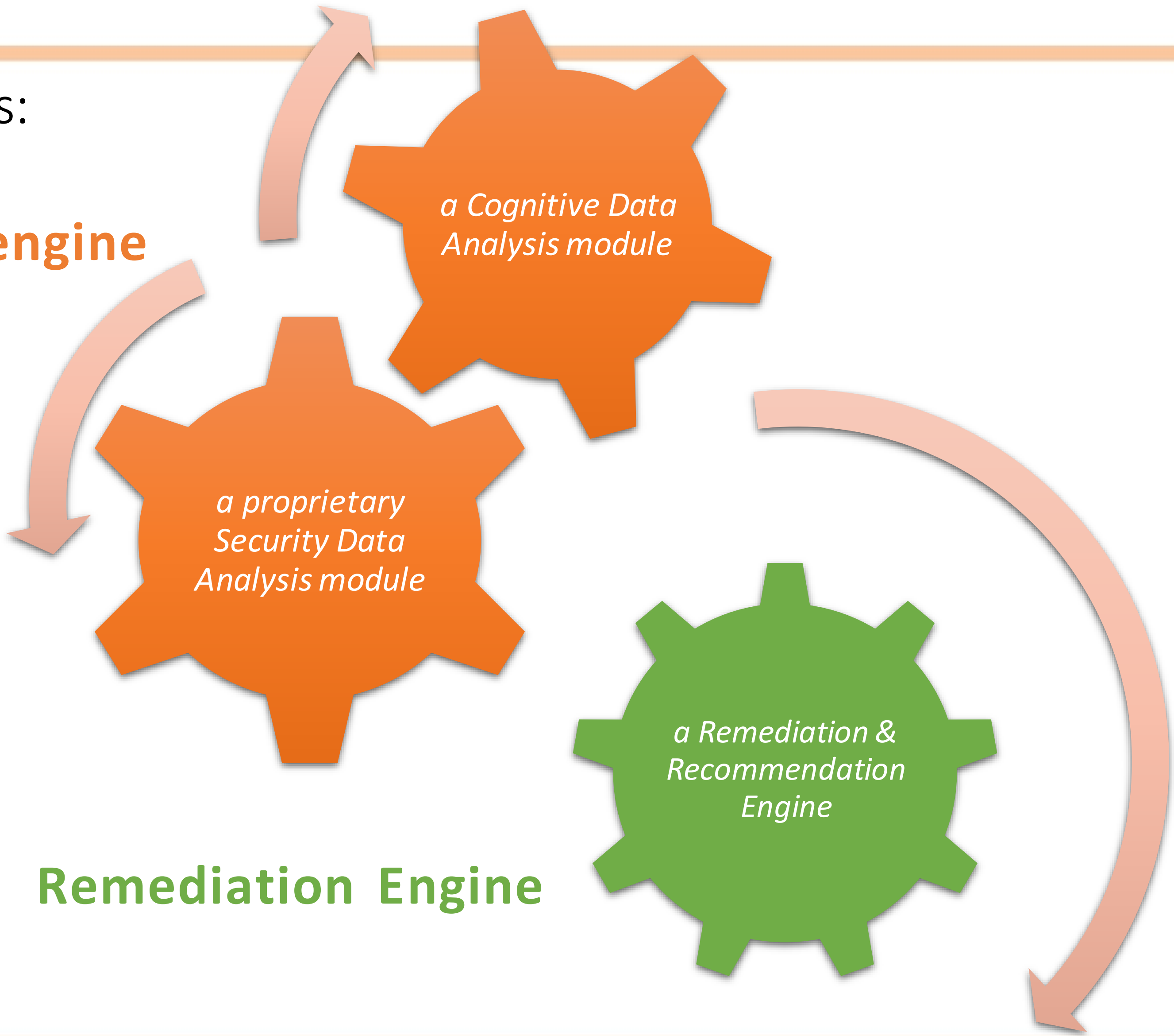
1 Project Management
1 Requirements
2 Developing
1 Integration and experimentation
1 Exploitation

2 Research centres
1 University
1 Public Cybersecurity Agency
3 Companies
4 SMEs

The DARE overview

Main components:

Data Analytics engine



Remediation Engine