Turning data into actionable intelligence

advanced features in MISP supporting your ana-

CIRCL / Team MISP Project

13th ENISA-EC3 Workshop

Turning data into actionable intelligence
advanced features in MISP supporting your ana-
CIRCL / Team MISP Project
13th ENISA-EC3 Workshop

The aim of this presentation

- Why is **contextualisation** important?
- What options do we have in MISP?
- How can we **leverage** this in the end?

The aim of this presentation
- Why is **contextualisation** important?
- What options do we have in MISP?
- How can we **leverage** this in the end?

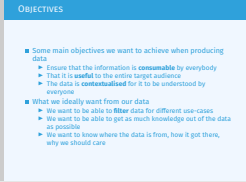
The growing need to contextualise data

- Contextualisation became more and more important as we as a community matured
  - **Growth and diversification** of our communities
  - Distinguish between information of interest and raw data
  - **False-positive** management
  - TTPs and aggregate information may be prevalent compared to raw data (risk assessment)
  - **Increased data volumes** leads to a need to be able to prioritise
- These help with filtering your TI based on your **requirements**...
- ...as highlighted by Pasquale Stirparo *Your Requirements Are Not My Requirements*

The growing need to contextualise data
- Contextualisation became more and more important as we as a community matured
  - **Growth and diversification** of our communities
  - Distinguish between information of interest and raw data
  - **False-positive** management
  - TTPs and aggregate information may be prevalent compared to raw data (risk assessment)
  - **Increased data volumes** leads to a need to be able to prioritise
- These help with filtering your TI based on your **requirements**...
- ...as highlighted by Pasquale Stirparo *Your Requirements Are Not My Requirements*

Objectives

- Some main objectives we want to achieve when producing data
  - Ensure that the information is **consumable** by everybody
  - That it is **useful** to the entire target audience
  - The data is **contextualised** for it to be understood by everyone
- What we ideally want from our data
  - We want to be able to **filter** data for different use-cases
  - We want to be able to get as much knowledge out of the data as possible
  - We want to know where the data is from, how it got there, why we should care



Different layers of context

- Context added by analysts / tools
- Data that tells a story
- Encoding analyst knowledge to automatically leverage the above

Different layers of context
- Context added by analysts / tools
- Data that tells a story
- Encoding analyst knowledge to automatically leverage the above

Context added by analysts / tools

Expressing why data-points matter

- An IP address by itself is barely ever interesting
- We need to tell the recipient / machine why this is relevant
- All data in MISP has a bare minimum required context
- We differentiate between indicators and supporting data

Expressing why data-points matter
- An IP address by itself is barely ever interesting
- We need to tell the recipient / machine why this is relevant
- All data in MISP has a bare minimum required context
- We differentiate between indicators and supporting data

Broadening the scope of what sort of context we are interested in

- **Who** can receive our data? **What** can they do with it?
- **Data accuracy, source reliability**
- **Why** is this data relevant to us?
- **Who** do we think is behind it, **what tools** were used?
- What sort of **motivations** are we dealing with? Who are the **targets**?
- How can we **block/detect/remediate** the attack?
- What sort of **impact** are we dealing with?

Broadening the scope of what sort of context we are interested in
Who can receive our data? What can they do with it?
Data accuracy, source reliability
Why is this data relevant to us?
Who do we think is behind it, what tools were used?
What sort of motivations are we dealing with? Who are the targets?
How can we block/detect/remediate the attack?
What sort of impact are we dealing with?

Tagging and taxonomies

- Simple labels
- Standardising on vocabularies
- Different organisational/community cultures require different nomenclatures
- Triple tag system - taxonomies
- JSON libraries that can easily be defined without our intervention

Tag	Events	Attributes	Tags
workflow:state="complete"	11	0	workflow:state="complete"
workflow:state="draft"	0	0	workflow:state="draft"
workflow:state="incomplete"	55	10	workflow:state="incomplete"
workflow:state="ongoing"	0	0	workflow:state="ongoing"

Galaxies

- Taxonomy tags often **non self-explanatory**
  - Example: universal understanding of tlp:green vs APT 28
- For the latter, a single string was ill-suited
- So we needed something new in addition to taxonomies - **Galaxies**
  - Community driven **knowledge-base libraries used as tags**
  - Including descriptions, links, synonyms, meta information, etc.
  - Goal was to keep it **simple and make it reusable**
  - Internally it works the exact same way as taxonomies (stick to **JSON**)

Ransomware galaxy

Galaxy ID	373
Name	Ransomware
Namespace	misp
Uuid	3f44af2e-1480-4b6b-9aa8-f9bb21341078
Description	Ransomware galaxy based on...
Version	4

Value ↓	Synonyms
.CryptoHasYou.	
777	Sevleg
7ev3n	7ev3n-HONE$T

The emergence of ATT&CK and similar galaxies

- Standardising on high-level **TTPs** was a solution to a long list of issues
- Adoption was rapid, tools producing ATT&CK data, familiar interface for users
- A much better take on kill-chain phases in general
- Feeds into our **filtering** and **situational awareness** needs extremely well
- Gave rise to other, ATT&CK-like systems tackling other concerns
  - **attck4fraud** [1] by Francesco Bigarella from ING
  - **Election guidelines** [2] by NIS Cooperation Group

[1] https://www.misp-project.org/galaxy.html#_attck4fraud
[2] https://www.misp-project.org/galaxy.html#_election_guidelines

The emergence of ATT&CK and similar galaxies
Standardising on high-level TTPs was a solution to a long list of issues
Adoption was rapid, tools producing ATT&CK data, familiar interface for users
A much better take on kill-chain phases in general
Feeds into our filtering and situational awareness needs extremely well
Gave rise to other, ATT&CK-like systems tackling other concerns
attck4fraud [1] by Francesco Bigarella from ING
Election guidelines [2] by NIS Cooperation Group
[1] https://www.misp-project.org/galaxy.html#_attck4fraud
[2] https://www.misp-project.org/galaxy.html#_election_guidelines

Data that tells a story

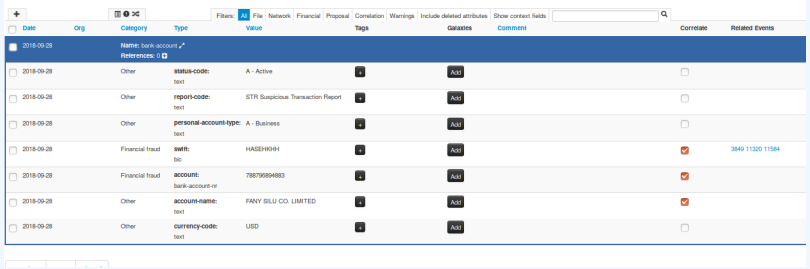
More complex data-structures for a modern age

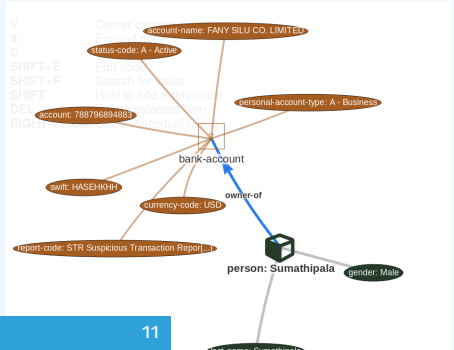
- Atomic attributes were a great starting point, but lacking in many aspects
- **MISP objects**[3] system
  - Simple **templating** approach
  - Use templating to build more complex structures
  - Decouple it from the core, allow users to **define their own** structures
  - MISP should understand the data without knowing the templates
  - Massive caveat: **Building blocks have to be MISP attribute types**
  - Allow **relationships** to be built between objects

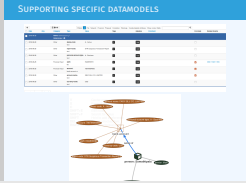
[3] https://github.com/MISP/misp-objects

More complex data-structures for a modern age
Atomic attributes were a great starting point, but lacking in many aspects
MISP objects[3] system
Simple templating approach
Use templating to build more complex structures
Decouple it from the core, allow users to define their own structures
MISP should understand the data without knowing the templates
Massive caveat: Building blocks have to be MISP attribute types
Allow relationships to be built between objects
[3] https://github.com/MISP/misp-objects

Supporting specific datamodels





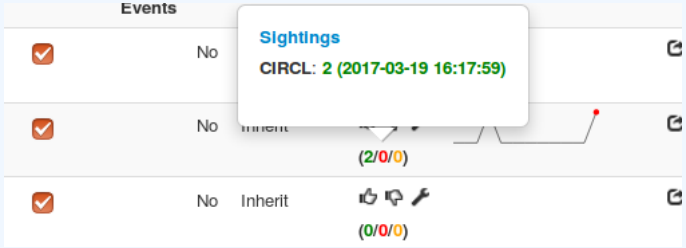


Continuous feedback loop

- Data shared was **frozen in time**
- All we had was a creation/modification timestamp
- Improved tooling and willingness allowed us to create a **feedback loop**
- Lead to the introduction of the **Sighting system**
- Signal the fact of an indicator sighting...
- ...as well as **when** and **where** it was sighted
- Vital component for IoC **lifecycle management**

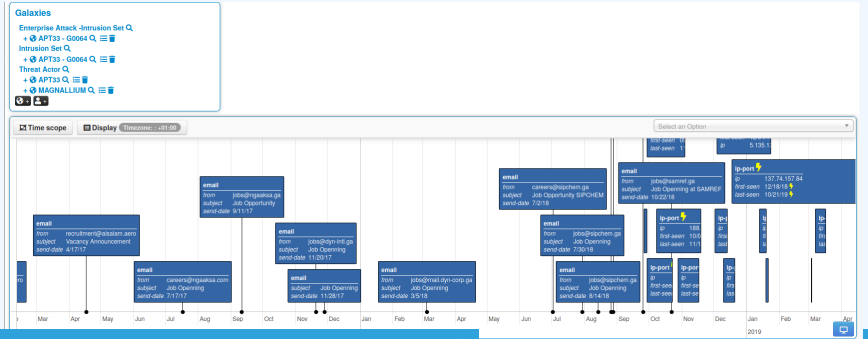
Continuous feedback loop
- Data shared was **frozen in time**
- All we had was a creation/modification timestamp
- Improved tooling and willingness allowed us to create a **feedback loop**
- Lead to the introduction of the **Sighting system**
- Signal the fact of an indicator sighting...
- ...as well as **when** and **where** it was sighted
- Vital component for IoC **lifecycle management**

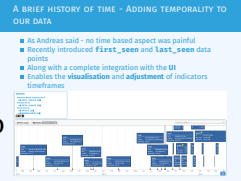
Continuous feedback loop (2)



A brief history of time - Adding temporality to our data

- As Andreas said - no time based aspect was painful
- Recently introduced **`first_seen`** and **`last_seen`** data points
- Along with a complete integration with the **UI**
- Enables the **visualisation** and **adjustment** of indicators timeframes





The various ways of encoding analyst knowledge to automatically leverage our TI

The various ways of encoding analyst knowledge to automatically leverage our TI

False positive handling

- Low quality / false positive prone information being shared
- Lead to **alert-fatigue**
- Exclude organisation xy out of the community?
- FPs are often obvious - **can be encoded**
- **Warninglist system**[4] aims to do that
- Lists of well-known indicators which are often false-positives like RFC1918 networks, ...

LIST OF KNOWN IPV4 PUBLIC DNS RESOLVERS

Id	89
Name	List of known IPv4 public DNS resolvers
Description	Event contains one or more public IPv4 DNS resolvers as attribute with an IDS flag set
Version	20181114
Type	string
Accepted attribute types	ip-src, ip-dst, domain|ip
Enabled	Yes (disable)

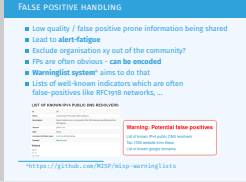
Values

1.0.0.1
1.1.1.1
1.11.71.4

Warning: Potential false positives

List of known IPv4 public DNS resolvers
Top 1000 website from Alexa
List of known google domains

[4] https://github.com/MISP/misp-warninglists



Making use of all this context

- Providing advanced ways of querying data
  - Unified export APIs
  - Incorporating all contextualisation options into **API filters**
  - Allowing for an **on-demand** way of **excluding potential false positives**
  - Allowing users to easily **build their own** export modules feed their various tools

Making use of all this context
Providing advanced ways of querying data
Unified export APIs
Incorporating all contextualisation options into API filters
Allowing for an on-demand way of excluding potential false positives
Allowing users to easily build their own export modules feed their various tools

Example query

```
/attributes/restSearch
{
    "returnFormat": "netfilter",
    "enforceWarninglist": 1,
    "tags": {
      "NOT": [
        "tlp:white",
        "type:OSINT"
      ],
      "OR": [
        "misp-galaxy:threat-actor=\"Sofacy\"",
        "misp-galaxy:sector=\"Chemical\""
      ],
    }
}
```


Example query
/attributes/restSearch
{
"returnFormat": "netfilter",
"enforceWarninglist": 1,
"tags": {
"NOT": [
"tlp:white",
"type:OSINT"
],
"OR": [
"misp-galaxy:threat-actor=\"Sofacy\"",
"misp-galaxy:sector=\"Chemical\""
],
}
}

Example query to generate ATT&CK heatmaps

```
/events/restSearch
{
    "returnFormat": "attack",
    "tags": [
        "misp-galaxy:sector=\"Chemical\""
    ],
    "timestamp": "365d"
}
```


Example query to generate ATT&CK heatmaps

/events/restSearch
{
"returnFormat": "attack",
"tags": [
"misp-galaxy:sector=\"Chemical\""
],
"timestamp": "365d"
}

A SAMPLE RESULT FOR THE ABOVE QUERY

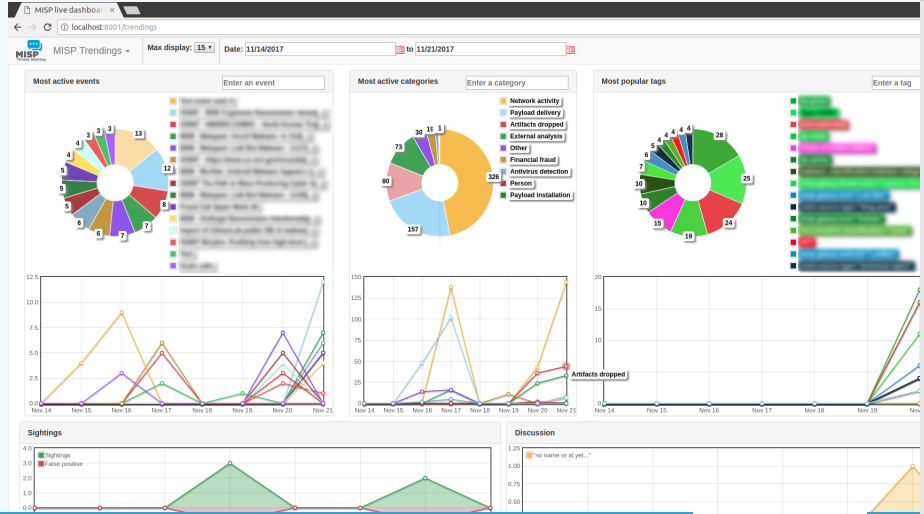
Pre Attack - Attack Pattern | **Enterprise Attack - Attack Pattern** | Mobile Attack - Attack Pattern

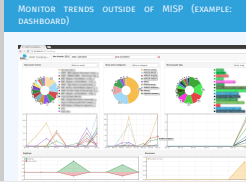
0 — 11 — Show all

Initial access	Execution	Persistence	Privilege escalation	Defense evasion	Credential access	Discovery	Lateral movement	Collection	Exfiltration	Command and control
Spearphishing Attachment	Scripting	Screensaver	File System Permissions Weakness	Process Hollowing	Securityd Memory	Password Policy Discovery	AppleScript	Data from Information Repositories	Exfiltration Over Alternative Protocol	Standard Application Layer Protocol
Spearphishing via Service	Command-Line Interface	Login Item	AppCert DLLs	Code Signing	Input Capture	System Network Configuration Discovery	Distributed Component Object Model	Data from Removable Media	Exfiltration Over Command and Control Channel	Communication Through Removable Media
Trusted Relationship	User Execution	Trap	Application Shimming	Rootkit	Bash History	Process Discovery	Pass the Hash	Man in the Browser	Data Compressed	Custom Command and Control Protocol
Replication Through Removable Media	Regsvcs/Regasm	System Firmware	Scheduled Task	NTFS File Attributes	Exploitation for Credential Access	Network Share Discovery	Exploitation of Remote Services	Data Staged	Automated Exfiltration	Multi-Stage Channels
Exploit Public-Facing Application	Trusted Developer Utilities	Registry Run Keys / Start Folder	Startup Items	Exploitation for Defense Evasion	Private Keys	Peripheral Device Discovery	Remote Desktop Protocol	Screen Capture	Scheduled Transfer	Remote Access Tools
Spearphishing Link	Windows Management Instrumentation	LC_LOAD_DYLIB Addition	New Service	Network Share Connection Removal	Brute Force	Account Discovery	Pass the Ticket	Email Collection	Data Encrypted	Uncommonly Used Port
Valid Accounts	Service Execution	LSASS Driver	Sudo Caching	Process Doppelgänging	Password Filter DLL	System Information Discovery	Windows Remote Management	Clipboard Data	Exfiltration Over Other Network Medium	Multilayer Encryption
Supply Chain Compromise	CMSTP	Rc.common	Process Injection	Disabling Security Tools	Two-Factor Authentication Interception	System Network Connections Discovery	Windows Admin Shares	Video Capture	Exfiltration Over Physical Medium	Domain Fronting
Drive-by Compromise	Control Panel Items	Authentication Package	Bypass User Account Control	Timestomp	LLMNR/NBT-NS Poisoning	Network Service Scanning	Remote Services	Audio Capture	Data Transfer Size Limits	Data Obfuscation
Hardware Additions	Dynamic Data Exchange	Component Firmware	Extra Window Memory Injection	Modify Registry	Credentials in Files	File and Directory Discovery	Taint Shared Content	Data from Network Shared Drive		Connection Proxy
	Source	Windows Management Instrumentation Event Subscription	Setuid and Setgid	Indicator Removal from Tools	Forced Authentication	Security Software Discovery	Application Deployment Software	Data from Local System		Commonly Used Port
	Space after Filename	Change Default File	Launch Daemon	Hidden Window	Keychain	System Service Discovery	Third-party Software	Automated Collection		Data Encoding

A SAMPLE RESULT FOR THE ABOVE QUERY

Monitor trends outside of MISP (example: dashboard)



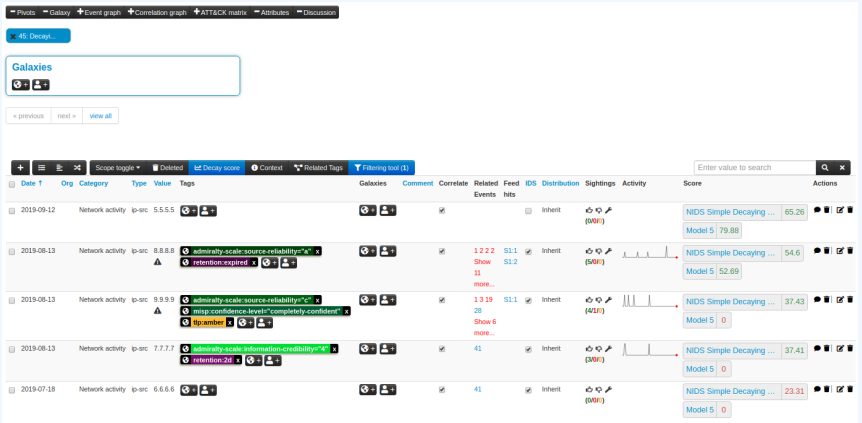


Decaying of indicators

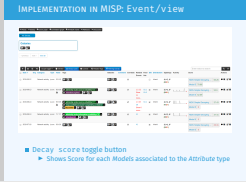
- We were still missing a way to use all of these systems in combination to decay indicators
- Move the decision making **from complex filter options to** complex **decay models**
- Decay models would take into account various available **context**
  - Taxonomies
  - Sightings
  - type of each indicator
  - Creation date
  - ...

Decaying of indicators
- We were still missing a way to use all of these systems in combination to decay indicators
- Move the decision making **from complex filter options to** complex **decay models**
- Decay models would take into account various available **context**
  - Taxonomies
  - Sightings
  - type of each indicator
  - Creation date
  - ...

Implementation in MISP: Event/view



- Decay score toggle button
  - Shows Score for each *Models* associated to the *Attribute* type



Implementation in MISP: API result

```
/attributes/restSearch
"Attribute": [
  {
    "category": "Network activity",
    "type": "ip-src",
    "to_ids": true,
    "timestamp": "1565703507",
    [...]
    "value": "8.8.8.8",
    "decay_score": [
      {
        "score": 54.475223849544456,
        "decayed": false,
        "DecayingModel": {
          "id": "85",
          "name": "NIDS Simple Decaying Model"
```


Implementation in MISP: API result
/attributes/restSearch
"Attribute": [
{
"category": "Network activity",
"type": "ip-src",
"to_ids": true,
"timestamp": "1565703507",
[...]
"value": "8.8.8.8",
"decay_score": [
{
"score": 54.475223849544456,
"decayed": false,
"DecayingModel": {
"id": "85",
"name": "NIDS Simple Decaying Model"
}

To sum it all up...

- Massive rise in **user capabilities**
- Growing need for truly **actionable threat intel**
- Lessons learned:
  - **Context is king** - Enables better decision making
  - **Intelligence and situational awareness** are natural by-products of context
  - Don't lock users into your **workflows**, build tools that enable theirs

To sum it all up...
Massive rise in user capabilities
Growing need for truly actionable threat intel
Lessons learned:
Context is king - Enables better decision making
Intelligence and situational awareness are natural by-products of context
Don't lock users into your workflows, build tools that enable theirs

Get in touch if you have any questions

- Contact us
  - https://twitter.com/mokaddem_sami
  - https://twitter.com/iglocska
- Contact CIRCL
  - info@circl.lu
  - https://twitter.com/circl_lu
  - https://www.circl.lu/
- Contact MISPProject
  - https://github.com/MISP
  - https://gitter.im/MISP/MISP
  - https://twitter.com/MISPProject