THREAT ONTOLOGIES FOR CYBER SECURITY ANALYTICS (TOCSA)

FIRST Conference 2017

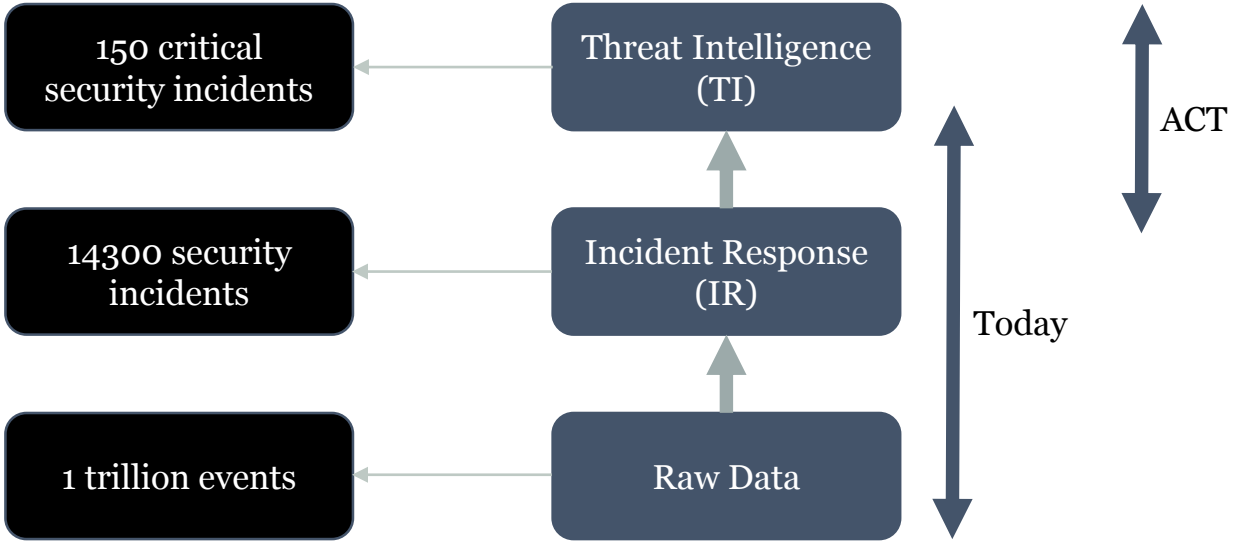
Martin Eian
meian@mnemonic.no

mnemonic

About Me

- Senior Security Analyst at mnemonic
- Project Manager «Semi-Automated Cyber Threat Intelligence (ACT)»
- Project Manager «Threat Ontologies for Cyber Security Analytics (TOCSA)»
- Member of the Europol EC3 Advisory Group on Internet Security

Motivation – mnemonic statistics from 2014



ACT, TOCSA and Oslo Analytics

- Semi-Automated Cyber Threat Intelligence (ACT)
  - Open Source Threat Intelligence Platform
  - https://www.mnemonic.no/research-and-development/semi-automated-cyber-threat-intelligence/
- Threat Ontologies for Cyber Security Analytics (TOCSA)
  - Ontologies
  - PhD Project
  - https://www.mnemonic.no/no/research-and-development/threat-ontologies-for-cybersecurity-analytics/
  - http://www.mn.uio.no/ifi/english/research/projects/tocsa/
- Operable Subjective Logic Analysis Technology for Intelligence in Cybersecurity (Oslo Analytics)
  - Analytics
  - Subjective Logic (quantifying uncertainty)
  - Trust Networks
  - Academic
  - http://www.mn.uio.no/ifi/english/research/projects/oslo-analytics/

Academic Paper: «Semantic Cyberthreat Modelling»

- Extended abstract presented at the Semantic Technology for Intelligence, Defense, and Security (STIDS) 2016 conference
  - http://stids.c4i.gmu.edu/
- Collaborative work:
  - Threat Ontologies in Cyber Security Analytics (TOCSA)
  - Operable Subjective Logic Analysis Technology for Intelligence in Cybersecurity (Oslo Analytics)
  - Semi-Automated Cyber Threat Intelligence (ACT)

THREAT INTELLIGENCE

mnemonic

What is Threat Intelligence?

Threat intelligence is *evidence-based knowledge*, including context, mechanisms, indicators, implications and actionable advice, about an existing or emerging *menace or hazard* to assets that can be used to *inform decisions* regarding the subject's response to that menace or hazard.

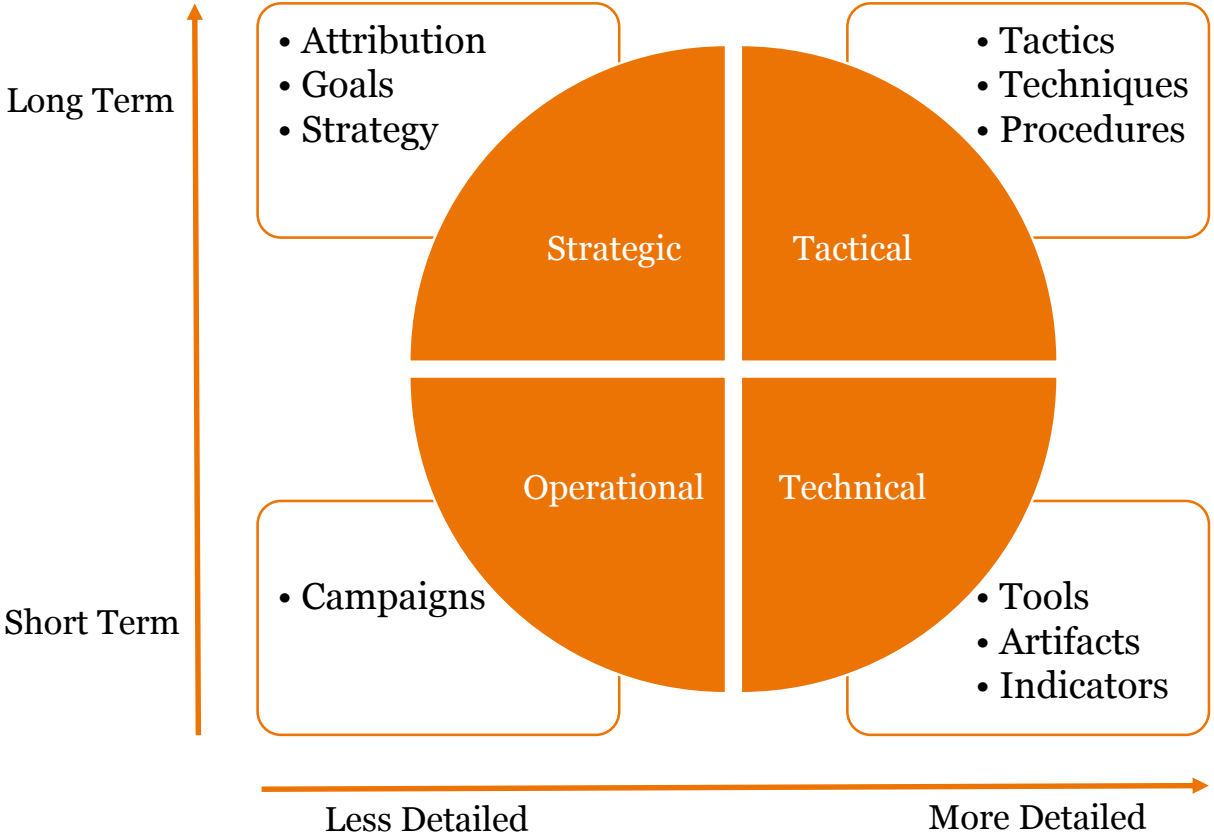
- Gartner (2013)

Evidence-Based Knowledge

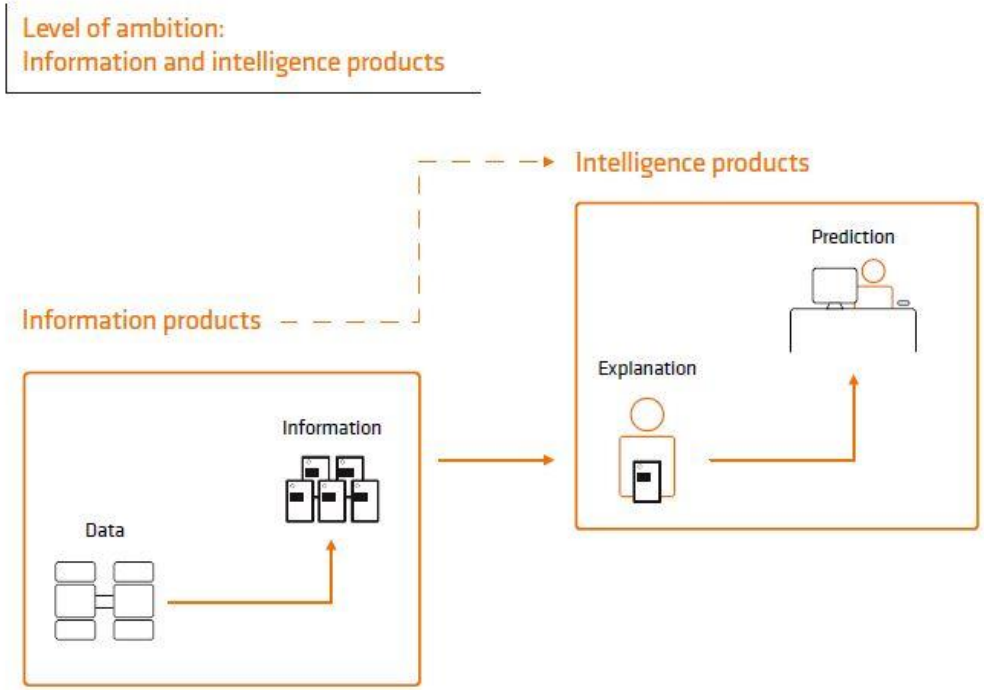
Evidence-Based *Knowledge*

www[.]iuqerfsodp9ifjaposdfjhgosurijfaewrwergwea[.]com

Threat Intelligence Categories



Threat Information vs Threat Intelligence



THREAT INTELLIGENCE PLATFORMS

mnemonic

Evalutation of existing platforms

Threat Intelligence Sharing Platforms: An Exploratory Study of Software Vendors and Research Perspectives

Clemens Sauerwein[1], Christian Sillaber[1], Andrea Mussmann[1], and Ruth Breu[1]

[1] University of Innsbruck, Department of Computer Science, Innsbruck, Austria
{Clemens.Sauerwein,Christian.Sillaber,Andrea.Mussmann,Ruth.Breu}
@uibk.ac.at

Abstract. In the last couple of years, organizations have demonstrated an increased willingness to exchange information and knowledge regarding vulnerabilities, threats, incidents and mitigation strategies in order to collectively protect against today's sophisticated cyberattacks. As a reaction to this trend, software vendors started to create offerings that facilitate this exchange and appear under the umbrella term "Threat Intelligence Sharing Platforms". To which extent these platforms provide the needed means for exchange and information sharing remains unclear as they lack a common definition, innovation in this area is mostly driven by vendors and empirical research is rare. To close this gap, we examine the state-of-the-art software vendor landscape of these platforms, identify gaps and present arising research perspectives. Therefore, we conducted a systematic study of 22 threat intelligence sharing platforms and compared them. We derived eight key findings and discuss how existing gaps should be addressed by future research.

mnemonic

http://aisel.aisnet.org/wi2017/track08/paper/3/

Key findings

1. There is no common definition of threat intelligence sharing platforms
2. STIX is the de-facto standard for describing threat intelligence
3. **Platforms primarily focus on sharing of indicators of compromise**
4. **The majority of platforms is closed source**
5. **Most platforms focus on data collection instead of analysis**
6. **Trust issues between users and platform providers are mostly neglected**
7. Academic and commercial interest in threat intelligence sharing increases
8. **Many manual tasks make the user the bottleneck**

EXAMPLE: APT REPORT

mnemonic

Report Contents

ShimRat core		ShimRat core loader DLL	
Filename(s)	elogger.dat	Filename(s)	elogger.dll
Related campaign	"Myanmar", see section 5		
C2 URL	https://secure2.sophosrv.com/en-us/support/ms-cache_check.php		
MD5	3eb9d4c448cd5ec8cb-49fa1e3b42b7d5	MD5	f34c6239b7d70f-23ce02a8d207176637
SHA256	8ee3fc5ccef751e098c4e-64b36e8b5c95d-c48473ac83380b59d10e-a32f9946f9	SHA256	35589ce27c27d-d4407a79540f32031d752b774b4bd-6b8a3687e19a177ae6b18b

ShimRat core	
Filename(s)	vmware-vmx.exe
Related campaign	"Global campaign", see section 6.4
C2 URL	https://ie.update-windows-microsoft.com/my/js/index.php
MD5	2cc5bc69e24a13bfc8ea3dc679ab0efc
SHA256	36422e6ccaa50a9ecceb7fb709a9e383552732525cb579f8438237d87aaf8377

ShimRat core		ShimRat core loader DLL	
Filename(s)	elogger.dat	Filename(s)	elogger.dll
Related campaign	-		
C2 URL	http://www.tinroofpopcorn.com/admin/fckeditor/_samples/_plugins/samples.php		
MD5	a3f7895fae05fa121a4e23d-d3595c366	MD5	5965731f2f237a-12f7a4873e3e37658a
SHA256	3c5c4d68d0fa6520637fb4a-fe6a7097ec7d0f-1d6a738bb0064bb009e-a6344e8d	SHA256	a03bd56eeee9f376eb-59c6f4d19bf8a651eeb57b-b4ebb7f884192b22a6616e68

ShimRat core	
Filename(s)	svchost.exe
Related campaign	-
C2 URL	http://update.nfkllyuisyahooapis.com/js/js/js.php
MD5	f9c14a8e9ceb143d959743ad8c09fdc4
SHA256	b53b27bb3e9d02e3ec5404cf3e67debb90d9337dbb570ca8b8cfce1054428466

ShimRat core	
Filename(s)	svchost.exe
Related campaign	-
C2 URL	http://www.go-gga.com/ez/doc/company/log/logon.php
MD5	663e54e686842eb8f8bae2472cf01ba1
SHA256	ba0057a1b132ec16559efc832941455cc07f34c434da2a7434f73f1d2141bebf

ShimRat core	
Filename(s)	svchost.exe
Related campaign	"Myanmar", see section 5
C2 URL	http://www.commerce.gov.mm/templates/css1/logon.php
MD5	a4da3b820883e9808bd3ca2e02437a25
SHA256	2b111e287d356ac4561ba4f56135b7c1361b7da32e5825028a5e300e44b05579

ShimRat core	
Filename(s)	vmware-vmx.exe
Related campaign	-
C2 URL	http://www.ipacking.co.kr/ez/admin/data/403.php
MD5	ca41c19366bee737fe5bc5008250976a
SHA256	029e735581c38d66f03aa0e9d1c22959b0bc8dfe298b9e91b127c42c7f904b5e

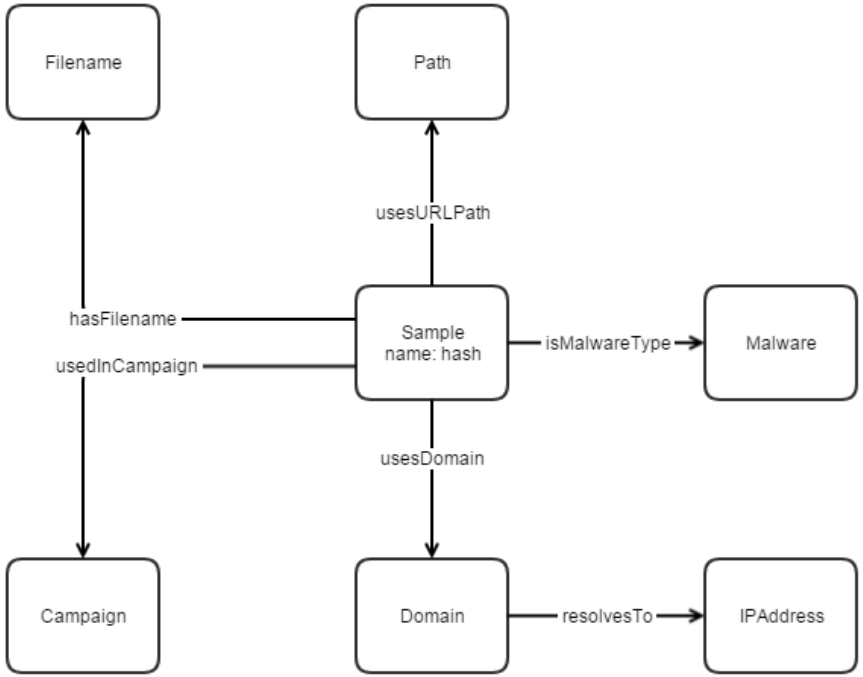
ShimRat core	
Filename(s)	-
Related campaign	"MSME DEFEXPO", see section 6.1
C2 URL	http://images.defexpoindia14.com/se/index.php
MD5	25e87e846bb969802e8db9b36d6cf67c
SHA256	33b288455c12bf7678fb5fd028ff3d42fcaf33cf833a147cb7f0f89f7dad0d8f

ShimRat core	
Filename(s)	helpservice.exe
Related campaign	"Global Campaign", see section 6.4
C2 URL	http://update.micrdsoft.com/image/image.php
MD5	cf883d04762b868b450275017ab3ccfa
SHA256	eb2d3c9e15b189dd02f753f805e90493254e17d40db6f1228a4e4095c5f260c1

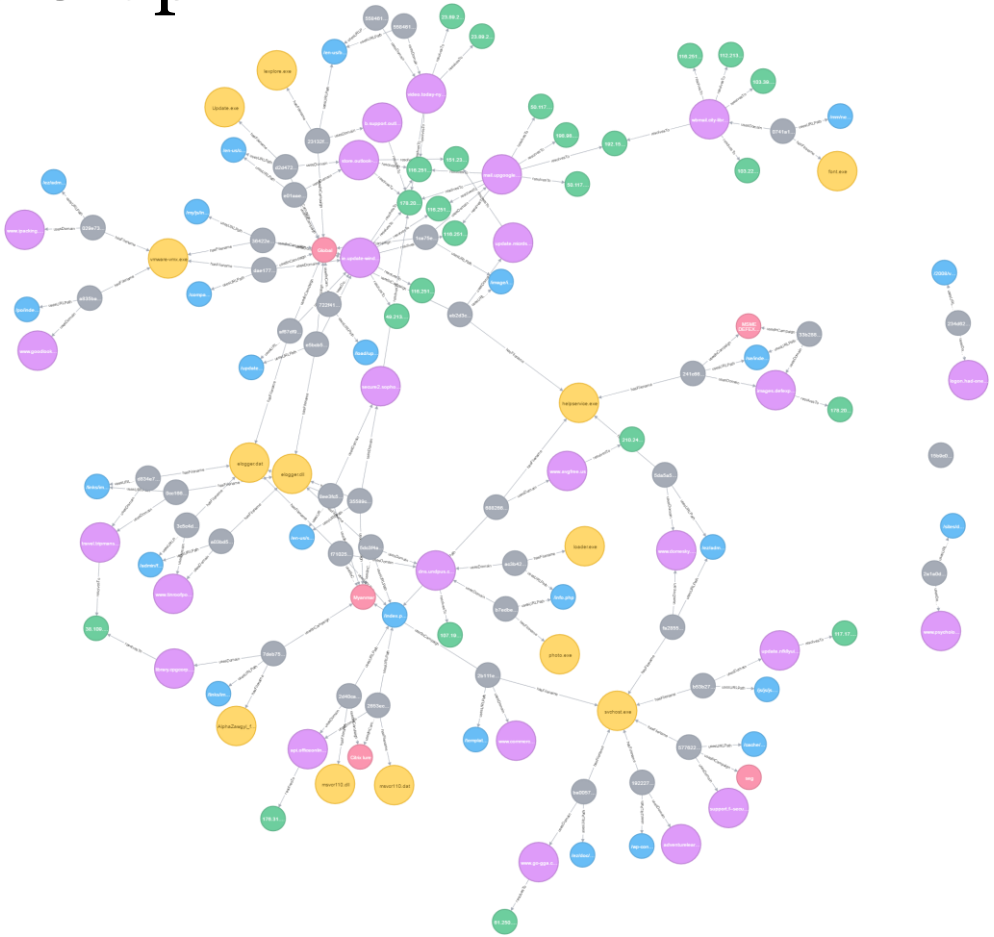
Approach

- Manually create csv files
- Design simple graph structure
- Transform csv files to graph DB using Python

Graph Structure

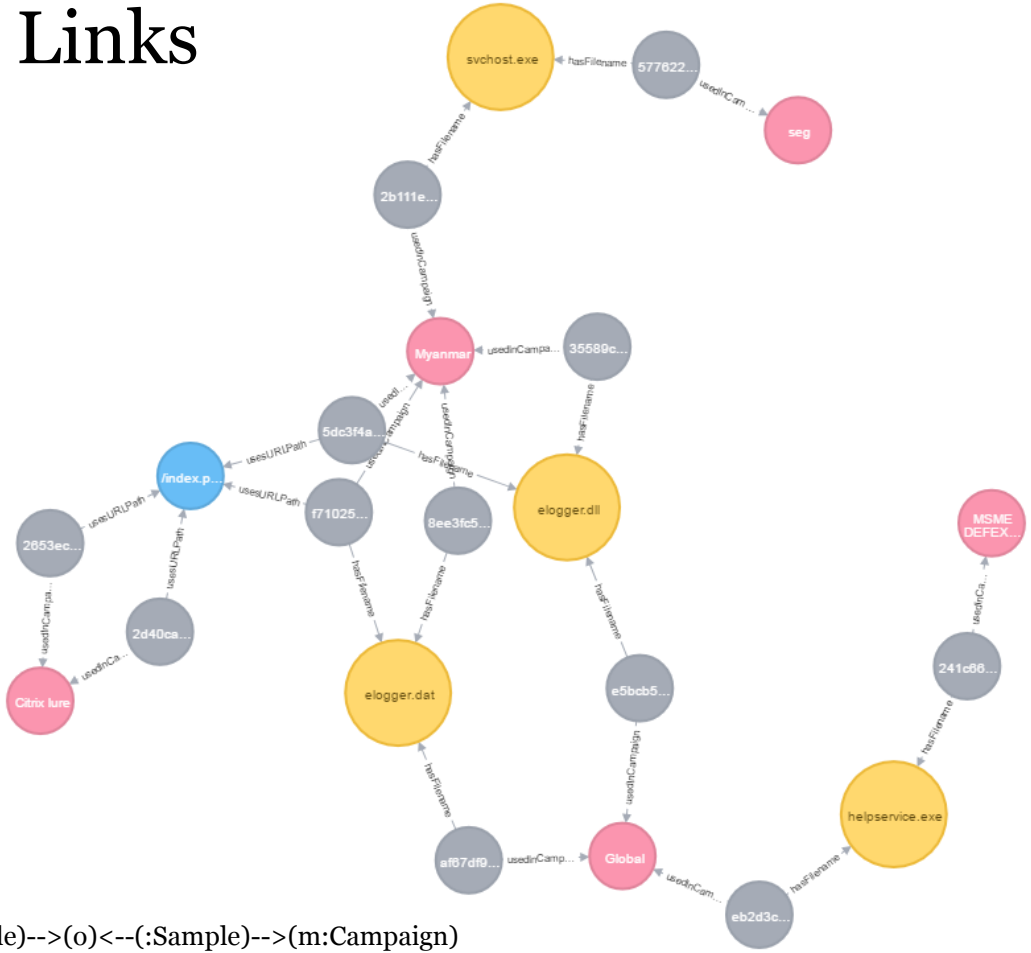


Knowledge Graph



mnemonic

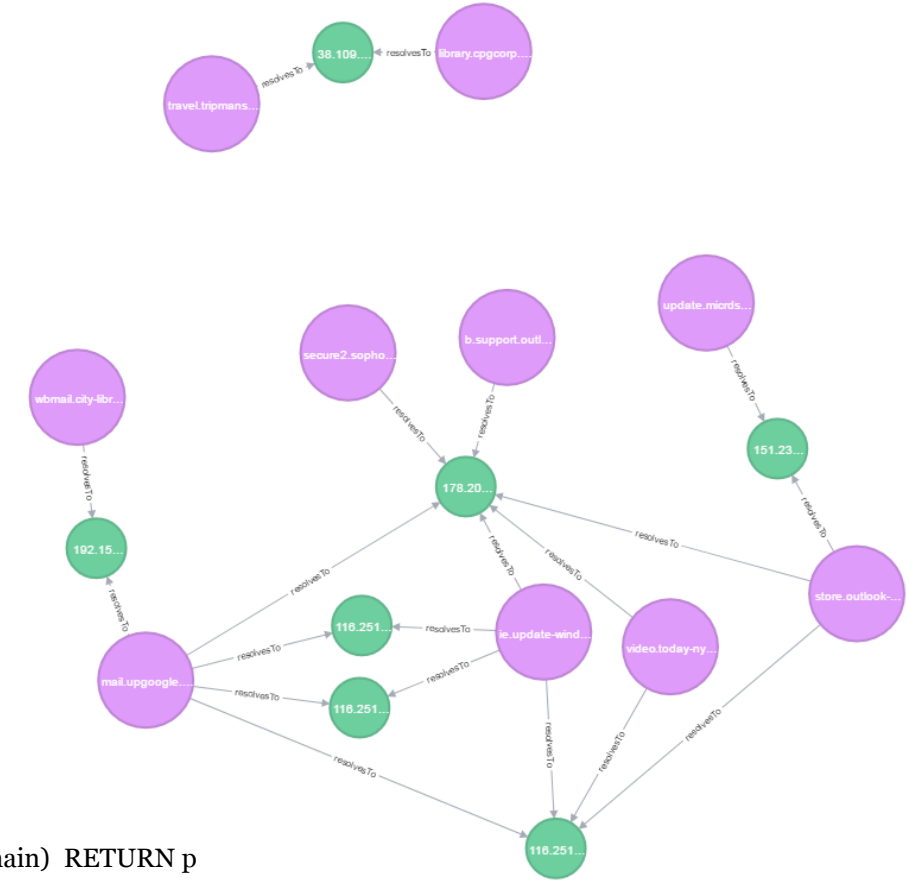
Campaign Links




```
MATCH p=(n:Campaign)<--(:Sample)-->(o)<--(:Sample)-->(m:Campaign)
WHERE NOT o:Malware AND m <> n  RETURN p
```

mnemonic

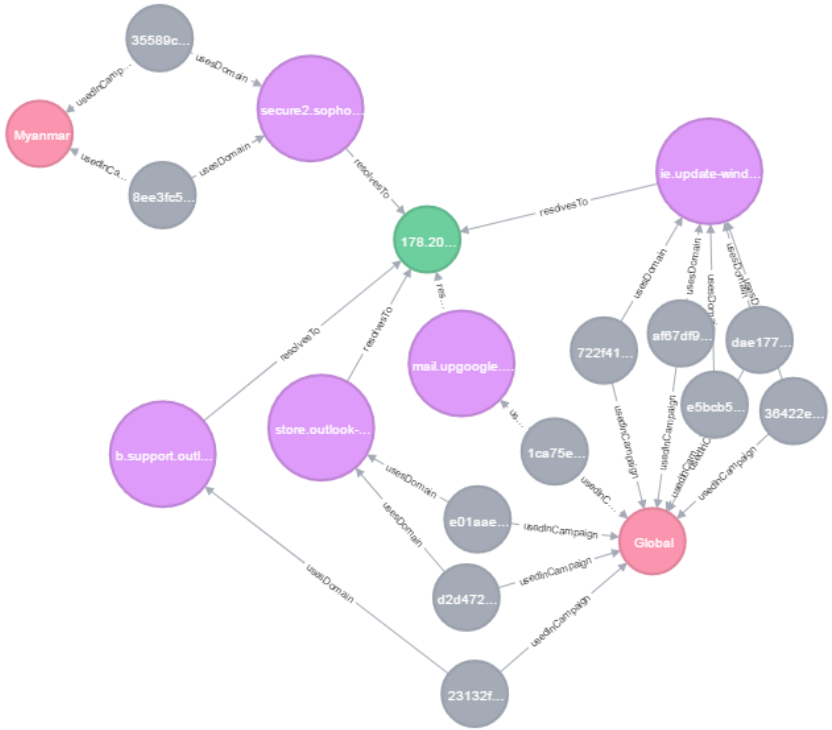
IP addresses with multiple domains



MATCH p=(n:Domain)-->(o:IP)<--(m:Domain) RETURN p

mnemonic

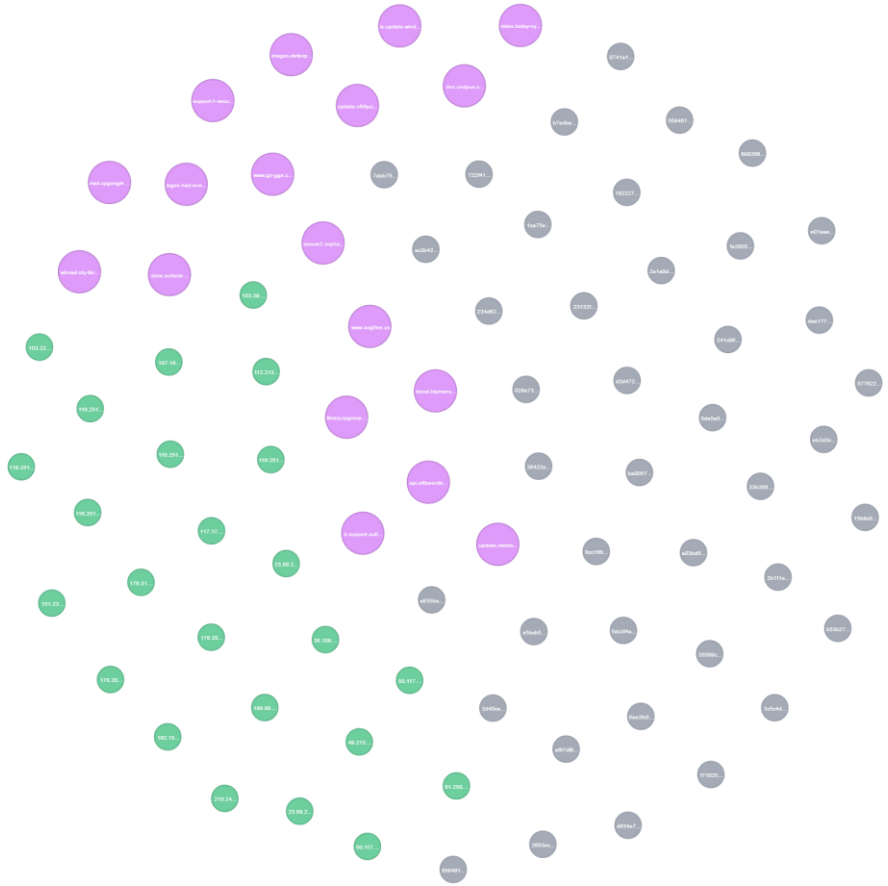
Relations to IP address




```
MATCH p=(m)<--()-[*1..3]->(n:IP {name: "178.209.52.72"}) WHERE NOT m:Malware AND
NOT m:Filename AND NOT m:Path AND NOT m:IP RETURN p
```

mnemonic

Knowledge Graph from STIX



SEMI-AUTOMATED CYBER THREAT INTELLIGENCE (ACT)

mnemonic

Semi-Automated Cyber Threat Intelligence (ACT)

The main objective of the research project is to develop a platform for cyber threat intelligence to uncover cyberattacks, cyber espionage and sabotage.

The project will result in new methods for data enrichment and data analysis to enable identification of threat agents, their motives, resources and attack methodologies.

In addition, the project will develop new methods, work processes and mechanisms for the generation and distribution of threat intelligence and countermeasures, to stop ongoing and prevent future attacks.

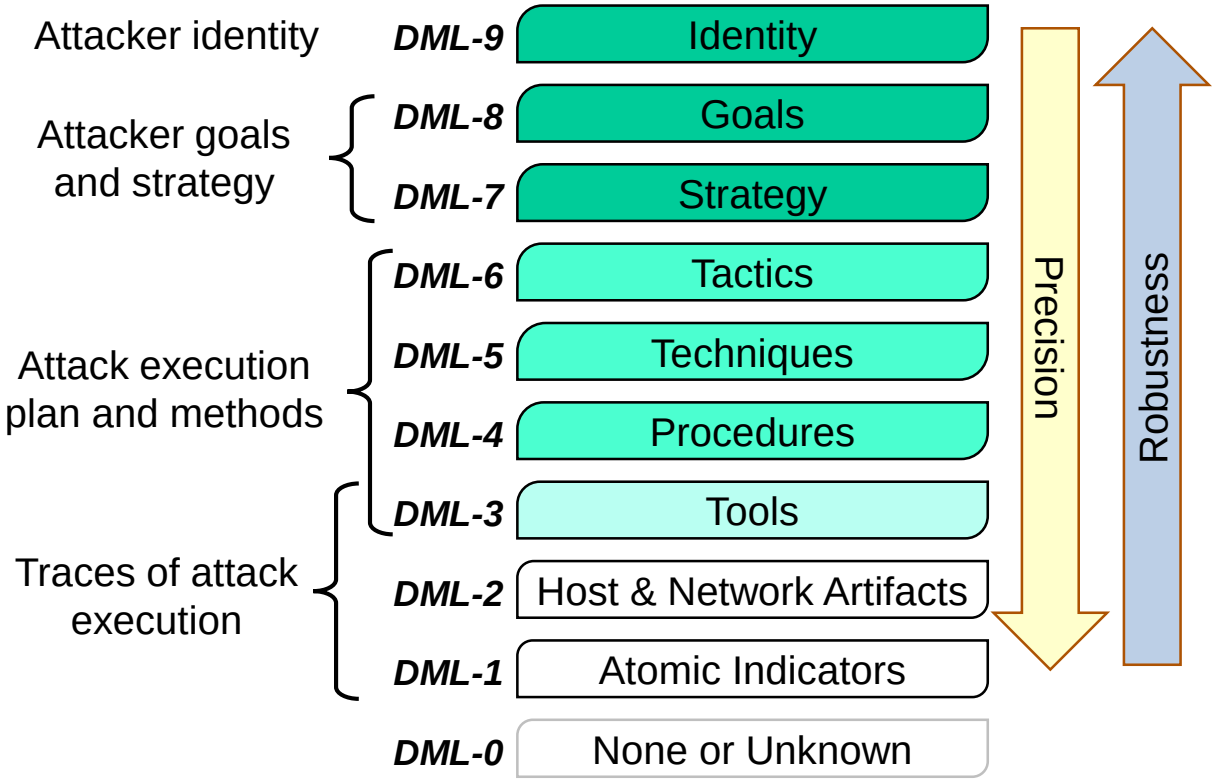
ACT Project Goals

- Holistic workspace for analysts
- Automation
  - Repetitive tasks
  - Processing of unstructured data
  - Sharing
    - Threat information
    - Countermeasures
- Advanced automated analysis
- Advanced enrichment
- Manual analysis
  - Efficiency
  - Accuracy
- Improve our knowledge of threat agents

Data Model

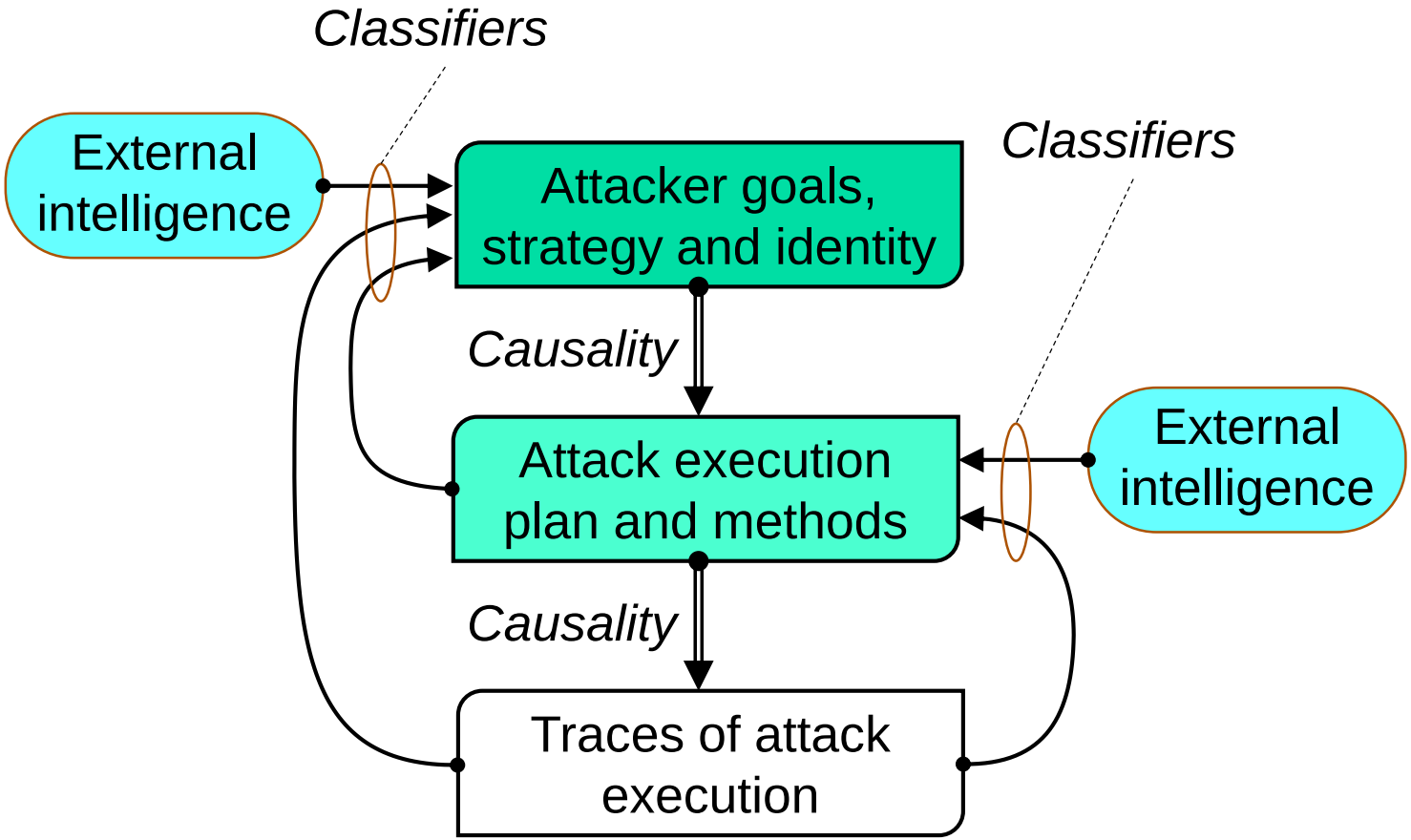
- Objects
  - Global
  - Example: IP address
- Facts
  - Connected to a single object or multiple objects (relation)
  - Immutable
  - Timestamped
  - Owner
  - Role-based and explicit access control
  - Backed by evidence and comments

The Detection Maturity Level (DML) Model



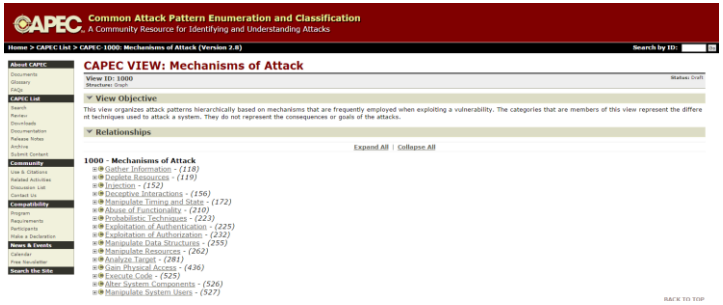
http://ryanstillions.blogspot.com/2014/04/the-dml-model_21.html

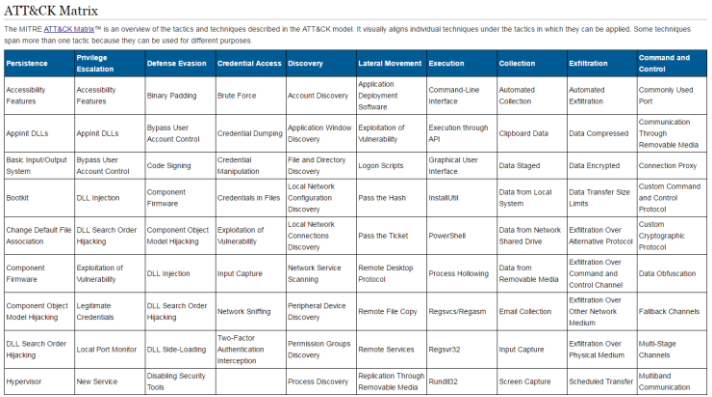
Causality and Classifiers for the DML Model



Semantic Feature Extraction

- Formal definitions of
  - -Goals
  - -Strategy
  - -Tactics
  - -Techniques
  - -Procedures
- Relevant initiatives
  - -MITRE CAPEC
    - https://capec.mitre.org
  - -MITRE ATT&CK
    - https://attack.mitre.org
  - -MITRE CAR
    - https://car.mitre.org





ATT&CK Matrix

The MITRE ATT&CK Matrix™ is an overview of the tactics and techniques described in the ATT&CK model. It visually aligns individual techniques under the tactics in which they can be applied. Some techniques span more than one tactic because they can be used for different purposes.

Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Execution	Collection	Exfiltration	Command and Control
Accessibility Features	Accessibility Features	Binary Padding	Brute Force	Account Discovery	Application Deployment Software	Command-Line Interface	Automated Collection	Automated Exfiltration	Commonly Used Port
AppInit DLLs	AppInit DLLs	Bypass User Account Control	Credential Dumping	Application Window Discovery	Exploitation of Vulnerability	Execution through API	Clipboard Data	Data Compressed	Communication Through Removable Media
Basic Input/Output System	Bypass User Account Control	Code Signing	Credential Manipulation	File and Directory Discovery	Logon Scripts	Graphical User Interface	Data Staged	Data Encrypted	Connection Proxy
Bootkit	DLL Injection	Component Firmware	Credentials in Files	Local Network Configuration Discovery	Pass the Hash	InstallUtil	Data from Local System	Data Transfer Size Limits	Custom Command and Control Protocol
Change Default File Association	DLL Search Order Hijacking	Component Object Model Hijacking	Exploitation of Vulnerability	Local Network Connections Discovery	Pass the Ticket	PowerShell	Data from Network Shared Drive	Exfiltration Over Alternative Protocol	Custom Cryptographic Protocol
Component Firmware	Exploitation of Vulnerability	DLL Injection	Input Capture	Network Service Scanning	Remote Desktop Protocol	Process Hollowing	Data from Removable Media	Exfiltration Over Command and Control Channel	Data Obfuscation
Component Object Model Hijacking	Legitimate Credentials	DLL Search Order Hijacking	Network Sniffing	Peripheral Device Discovery	Remote File Copy	Regsvcs/Regasm	Email Collection	Exfiltration Over Other Network Medium	Fallback Channels
DLL Search Order Hijacking	Local Port Monitor	DLL Side-Loading	Two-Factor Authentication Interception	Permission Groups Discovery	Remote Services	Regsvr32	Input Capture	Exfiltration Over Physical Medium	Multi-Stage Channels
Hypervisor	New Service	Disabling Security Tools		Process Discovery	Replication Through Removable Media	Rundll32	Screen Capture	Scheduled Transfer	Multiband Communication

APT Report Example: Tactics, Techniques and Procedures

Network hopping and exfiltration

Once APT10 have a foothold in victim networks, using either legitimate MSP or local domain credentials, or their sustained malware such as PlugX, RedLeaves or Quasar RAT, they will begin to identify systems of interest.

The operator will either access these systems over RDP, or browse folders using Remote Access Trojan (RAT) functionality, to identify data of interest. This data is then staged for exfiltration in multi-part archives, often placed in the Recycle Bin, using either RAR or TAR. The compression tools are often launched via a remote command execution script which is regularly named 't.vbs' and is a customised version of an open source WMI command executor which pipes the command output back to the operator.

https://www.pwc.co.uk/issues/cyber-security-data-privacy/insights/operation-cloud-hopper.html

Example Procedure: Authentication with stolen credentials

Environment: Windows cmd.exe command line

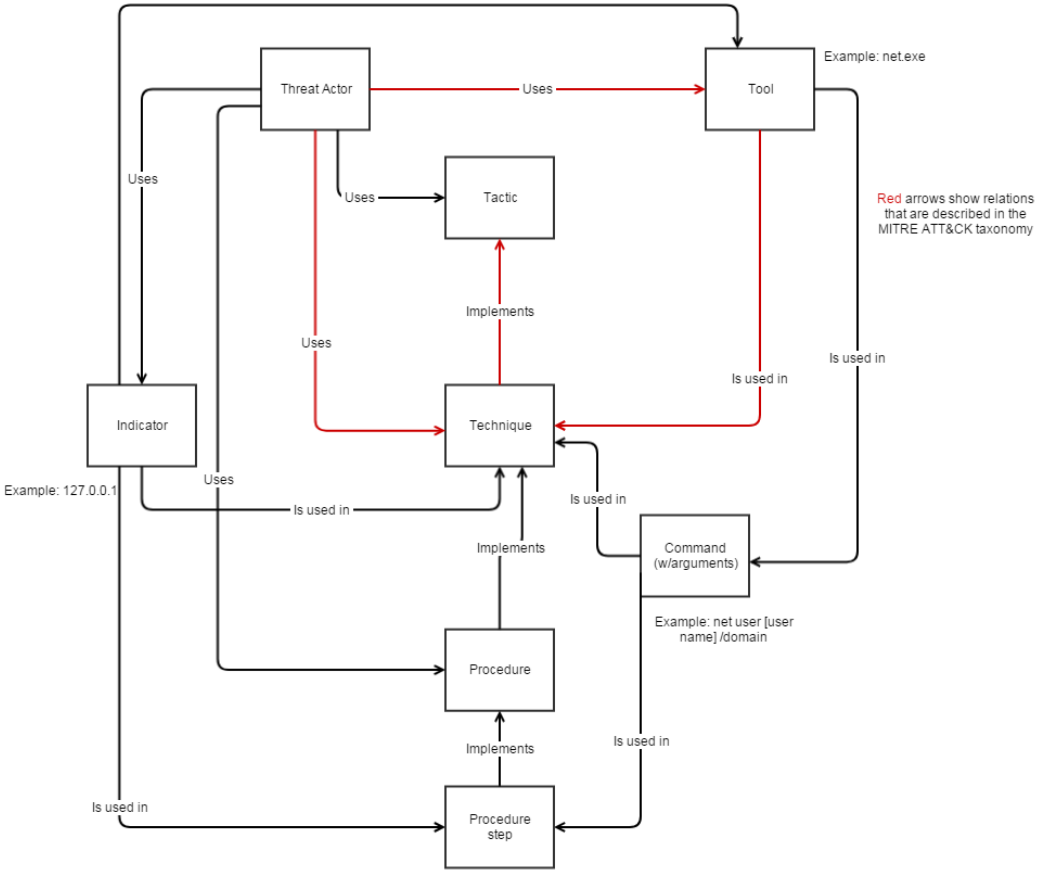
1. ping -n 1 HOSTNAME
2. net use \\HOSTNAME\ipc$ "PASSWORD" /user:"DOMAIN\USERNAME"

Example Procedure Detection

Prerequisite: logging of cmd.exe command line (e.g. Sysmon)

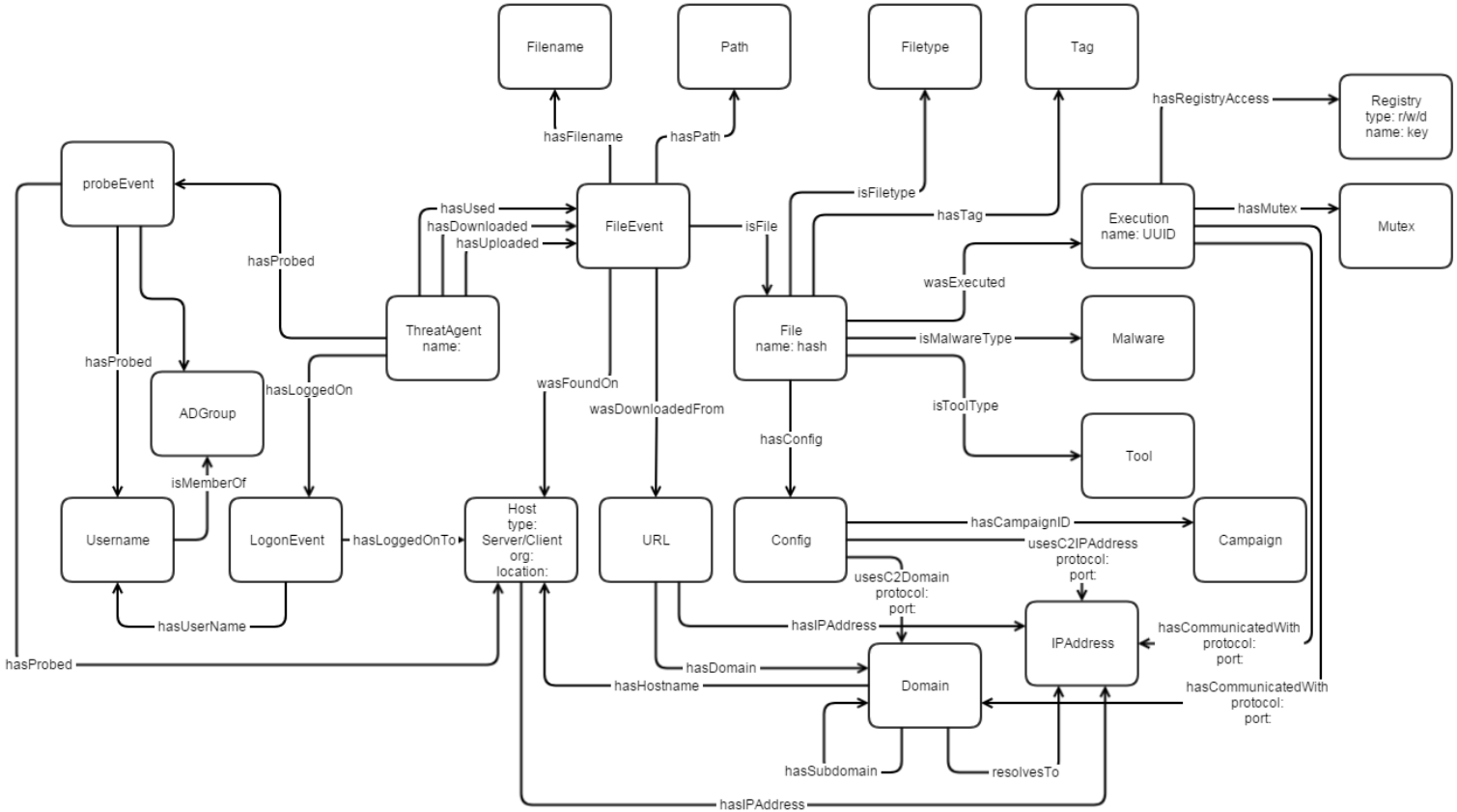
for each COMMANDLINE in cmd.exe process:
    if COMMANDLINE matches ‘ping -n 1 HOSTNAME’:
        if next COMMANDLINE starts with ‘net use \\HOSTNAME\ipc$’:
            Trigger alarm

TTP Knowledge Graph



mnemonic

Traces Knowledge Graph



mnemonic

Unstructured Data – Natural Language Processing

- No corpus for the cyber security domain
- *Snowball*: Extracting Relations from Large Plain-Text Collections [1]
- Test case: APTNotes (https://github.com/aptnotes/data)

1: http://www.cs.columbia.edu/~gravano/Papers/2000/dl00.pdf

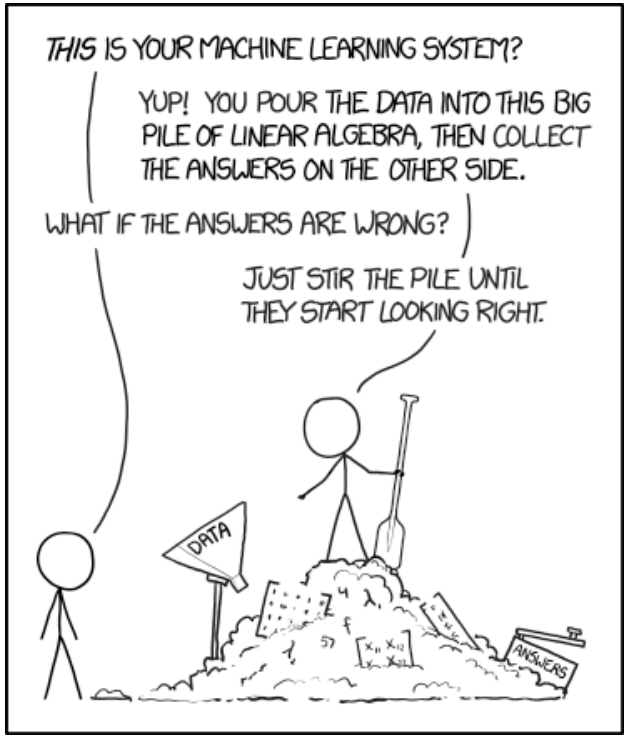
APTNotes NLP processing

t	description	The report focuses on a targeted threat group that we call APT 28 (Advanced Persistent Threat group 28) and details ongo
t	indicators.cve	
t	indicators.email	nato_smtp@mail.ru, nato_pop@mail.ru, lisa.cuddy@wind0ws.kz, dr.house@wind0ws.kz
t	indicators.fqdn	www.nytimes.com, www.nytimes.com, police.ge, mia.ge.gov, uropa.eu, www.freedomhouse.org, www.upi.com, voiceofrussia.com gov.hu, q0v.pl, mail.q0v.pl, gov.pl, mail.gov.pl, poczta.mon.q0v.pl, poczta.mon.gov.pl, www.mil.ee, www.kam.lt, nato.ns om, windous.kz, wind0ws.kz, dr.house, wind0ws.kz
t	indicators.ipv4	
t	indicators.ipv6	
t	indicators.md5	272f0fde35dbdfccbca1e33373b3570d, 8b92fe86c5b7a9e34f433a6fbac8bc3a, 9eebfebe3987fec3c395594dc57a0c4c, da2a657dc69d7320f2 48656a93f9ba39410763a2196aabc67f, 8c4fa713c5e2b009114adda758adc445
t	indicators.msid	ms12-27
t	indicators.sha1	
t	indicators.sha256	
🕐	meta:creation-date	October 27th 2014, 23:28:50.000
🕐	meta:save-date	October 27th 2014, 23:29:32.000
🕐	modified	October 27th 2014, 23:29:32.000
t	nlp.dates	12 August 2008, 2008, August 2010, 2010, 2007, 2008, 2008, six years, 2007, 2013, Over 96, Monday, Friday, 2007 to Sept 009, June 2014, June 2014, June 2014, 8251, 2014, 2014, June 2014, Today, June 2014, 2010, February, 2010, 2012 16, 2014 5285, 2196, April 2013, April 2013, 2004, 2004, 2004, 2004, April 2013, 2007, 2007, 2007, 2014, 0419, 0409, 0000, 0800 y October December January July August October November December 0 1 2 3 4, 5 6, 7 8 9, 0400, Over 96, Monday, Friday, 5 5, 5555, last four, This, This, 63 6, 0001, 0001, 1001, 1101, 1301, 0002, 1002, 1102, 1302, 1602, 0001, 0002, This, 1102
t	nlp.locations	Russia, Caucasus, Caucasus, Caucasus, Moscow, St . Petersburg, Russia, Georgia, Russia, Moscow, St . Petersburg, Moscow Armenia, Azerbaijan, Moscow, Caucasus, Georgian, Georgia, Georgian, Caucasus, Georgia, Georgia, Georgian, Tbilisi, Georg ia, Georgia, U . S ., Caucasus, Caucasus, Georgia, Russia, China, Iran, Caucasus, Moscow, Caucasus, Russia, Sweden, Russ nia, Baltic, Russia, Russia, Soviet Union, Russia, Russia, Europe, Europe, Russia, U . S ., London, Turkey, Russia, Euro a, Russia, Moscow, St . Petersburg, Moscow, Russia, Caucasus, Eastern Europe, Baltic Host, Cau Russia, Russia, HKU, Russia, Cisco, Russia, Russia, Russia, SSzFKlR 4 l 0, Russia, Russia, Russia, Russia, Russia, 267(
t	nlp.moneys	69 b 5 6, 46 36 66 35 61 68, 10 41 70 41 10 42 33, 3, $. 1, 87 88 35, 56 34 4, 1 " versus
t	nlp.organizations	Georgian Ministry of Internal Affairs, MIA, Georgian Ministry of Defense, NATO, Defense Exhibitions, Nation State Inter aty Organization, NATO, Georgian Ministry of Internal Affairs, Ministry of Defense, RUSSIA Chechnya GEORGIA Abkhazia TU rgian Ministry of Internal Affairs, Georgian Ministry of Internal Affairs, MIA, Georgian Ministry of Defense APT, MOD, M ouse Press Release, Kremlin, NATO, Eastern, European Ministry of Foreign Affairs, Malaysia Airlines, Eastern, Baltic Hos Defense, NATO, NATO, NATO, Russian Federation, NATO, NATO, Organization for Security and Cooperation, NATO, NATO, NATO ration, Attache Corps, AMAC, Defense Exhibitions, SMi Group, Energy, SMi Group, Army, Navy, Defense Contractors, Europe

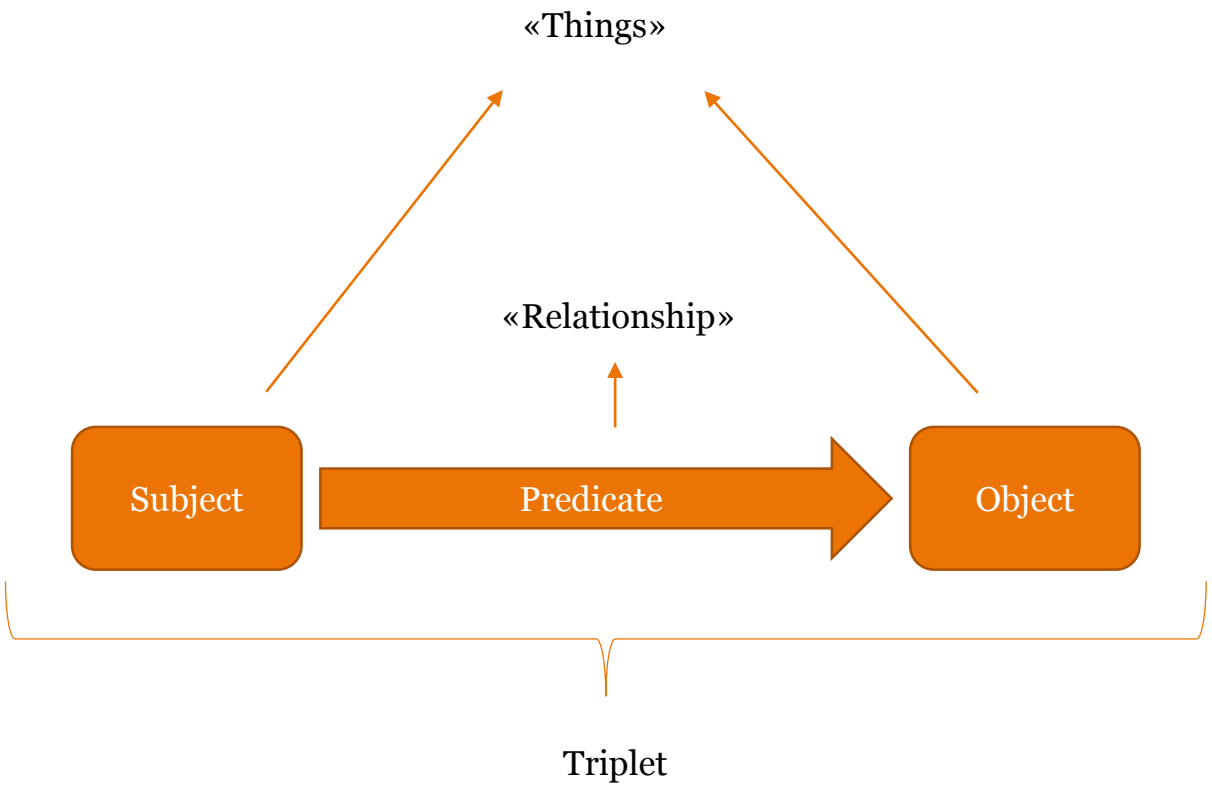
ADVANCED AUTOMATED ANALYSIS

mnemonic

Machine Learning

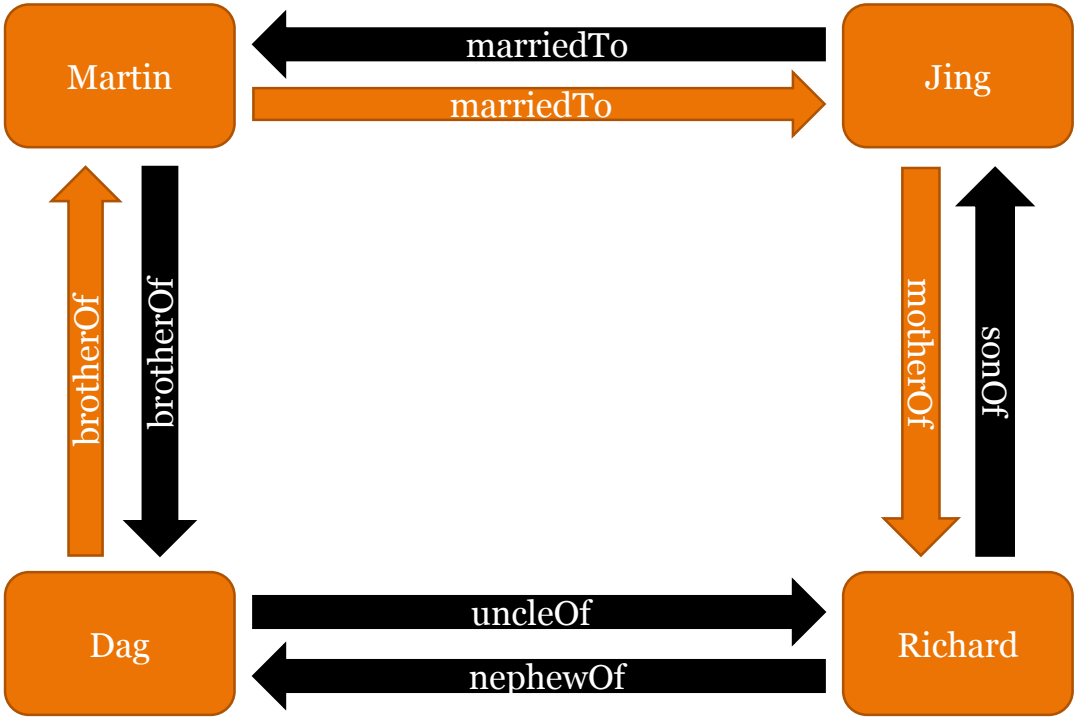


Triplets and semantic reasoning

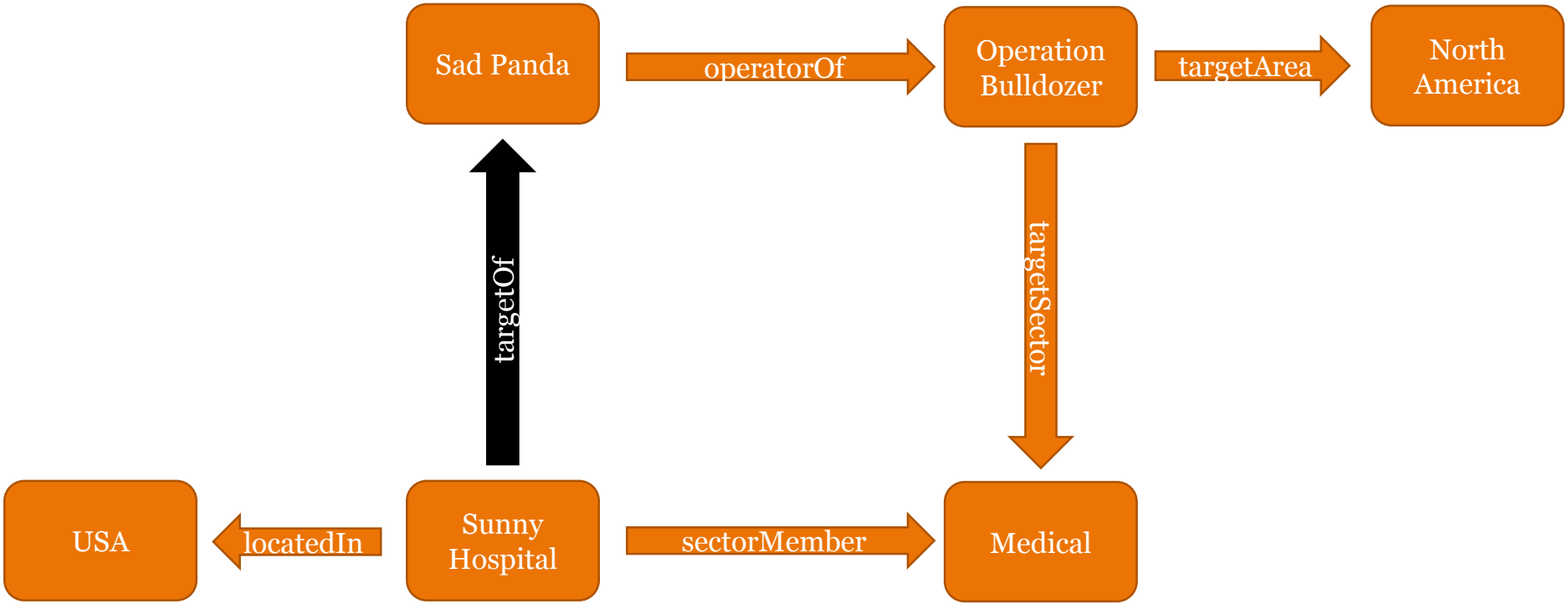


Triplets and semantic reasoning

Triplets and semantic reasoning



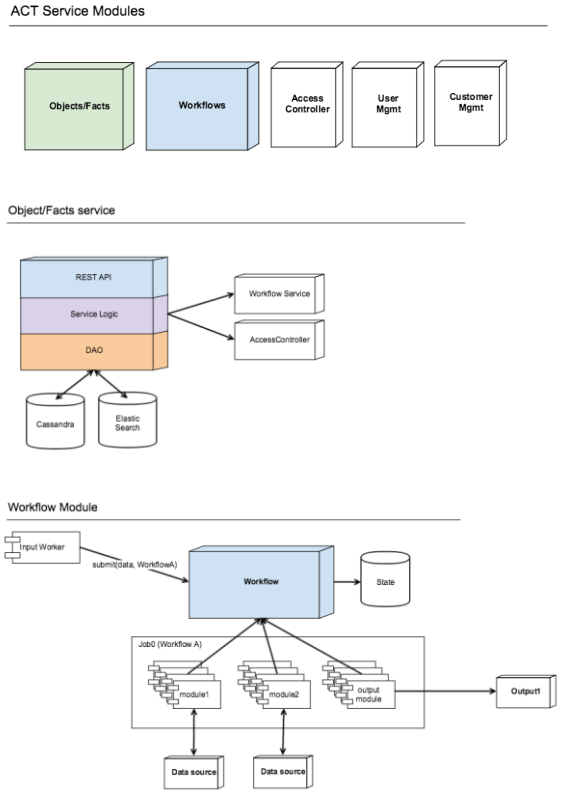
Triplets and semantic reasoning



SUMMARY

Threat Intelligence Platform

- Data model and architecture done
  - Objects and immutable facts (relations/predicates)
  - ACL on facts
  - Queues and workers
- Platform core, API and GUI under development and testing
- Github project
  - https://github.com/mnemonic-no
- Ongoing research:
  - Threat ontologies
  - Analysis techniques
  - Enrichment techniques
  - Sharing and Countermeasures
  - Workflow orchestration



mnemonic

Feedback and ideas

- Useful, formal definitions of TTPs
- Examples of predicates («marriedTo») for Threat Intelligence
- Experiences, use cases
- Any other clever ideas