SEMI-AUTOMATED CYBER THREAT INTELLIGENCE (ACT)

Training – FIRST Conference 2019 Edinburgh

Martin Eian, Geir Skjøtskift, Siri Bromander and Tom Spangebu

mnemonic

When	What
09:00 – 10:30	Introduction to ACT
10:30 – 10:45	Break
10:45 – 13:00	Assignments, case study
13:00 – 14:00	Lunch (not provided)
14:00 – 15:30	Recap, breakout (API/Graph queries)
15:30 – 15:45	Break
15:45 – 18:00	Practical work (build something)
18:30 – 19:00	Newbie reception
19:00 – 21:00	Icebreaker reception

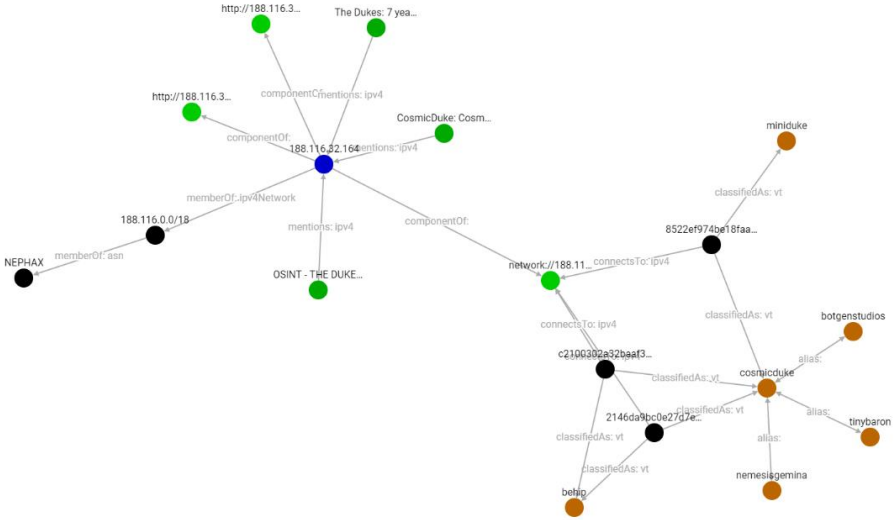
Goal

To collect and organize our knowledge of threats to make it useful

Data and Information

Data

Information



mnemonic

Semi-Automated…

- Analysis
- Enrichment
- Information Sharing
- Countermeasures

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.

NSM

Nordic Financial CERT

Kraft CERT

UiO : Universitetet i Oslo

NTNU
Norwegian University of Science and Technology

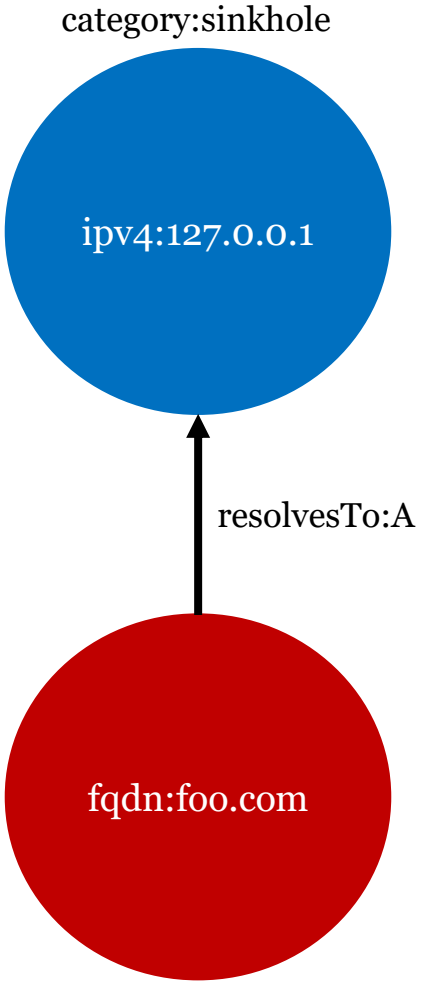
telenor

mnemonic

DATA MODEL

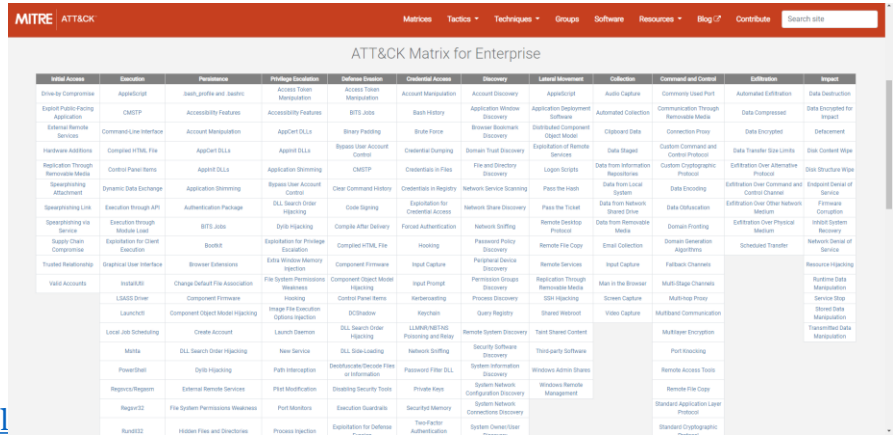
Data Model

- Objects
  - Global
  - Example: IP address
- Facts
  - Connected to one or two objects
  - Immutable
  - Timestamped
  - Owner
  - Role-based and explicit access control
  - Backed by evidence and comments



Models, Taxonomies and Vocabularies

- MITRE ATT&CK
  - https://attack.mitre.org
- MITRE PRE-ATT&CK
  - https://attack.mitre.org/pre-attack/
- STIX 2.0 vocabularies
  - https://oasis-open.github.io/cti-documentation/
- Ryan Stillions' DML model
  - http://ryanstillions.blogspot.com/2014/04/the-dml-model_21.html



Current OSINT Sources

Import:

- APTNotes
  - https://github.com/aptnotes/data
- APT & CyberCriminal Campaign Collection
  - https://github.com/CyberMonitor/APT_CyberCriminal_Campagin_Collections
- RSS Feeds
  - Infosec blogs
- MISP (circl.lu feed)
- MITRE ATT&CK

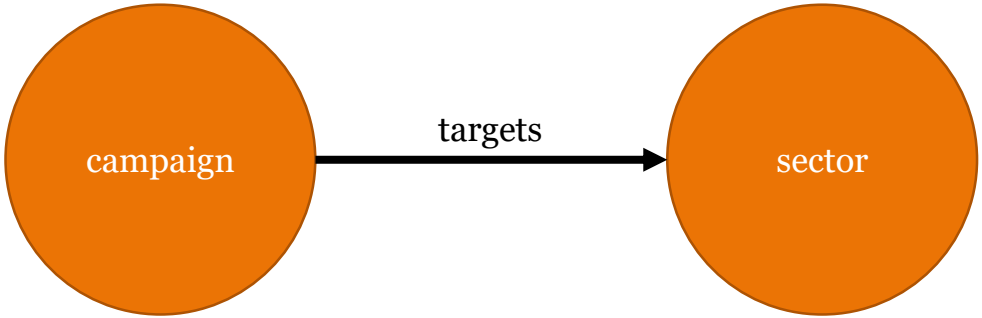
Enrich:

- mnemonic PassiveDNS
  - https://passivedns.mnemonic.no/
- Shadowserver IP-BGP
  - https://www.shadowserver.org/wiki/pmwiki.php/Services/IP-BGP
- VirusTotal

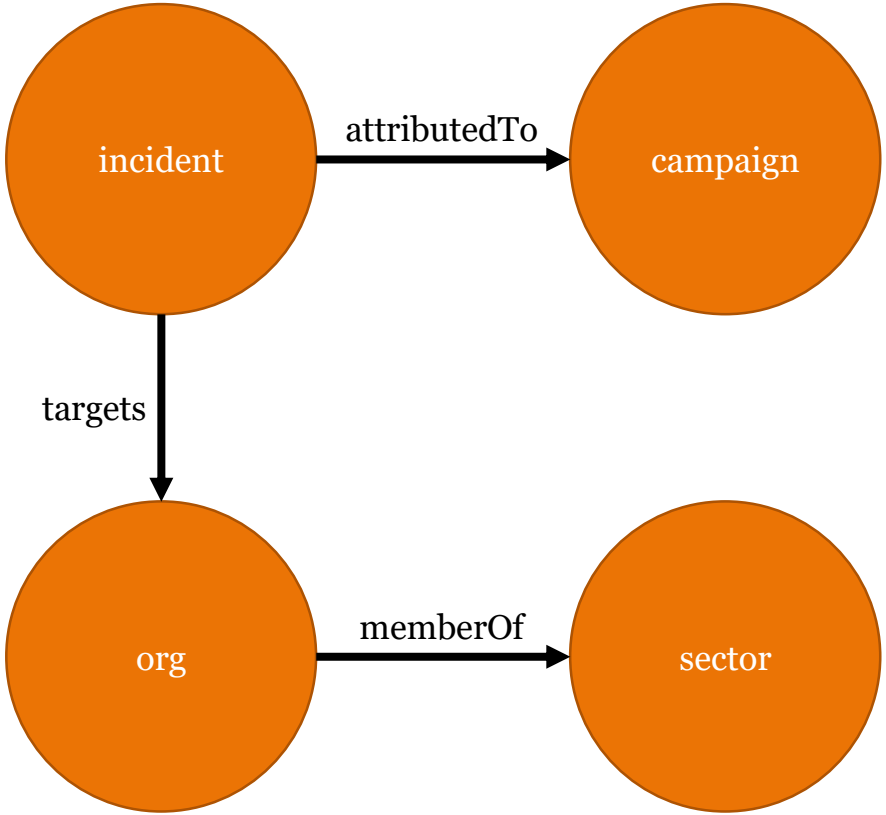
Problems

- Multiple ways to represent the same information
- Different names for the same thing
  - Threat actors
  - Malware

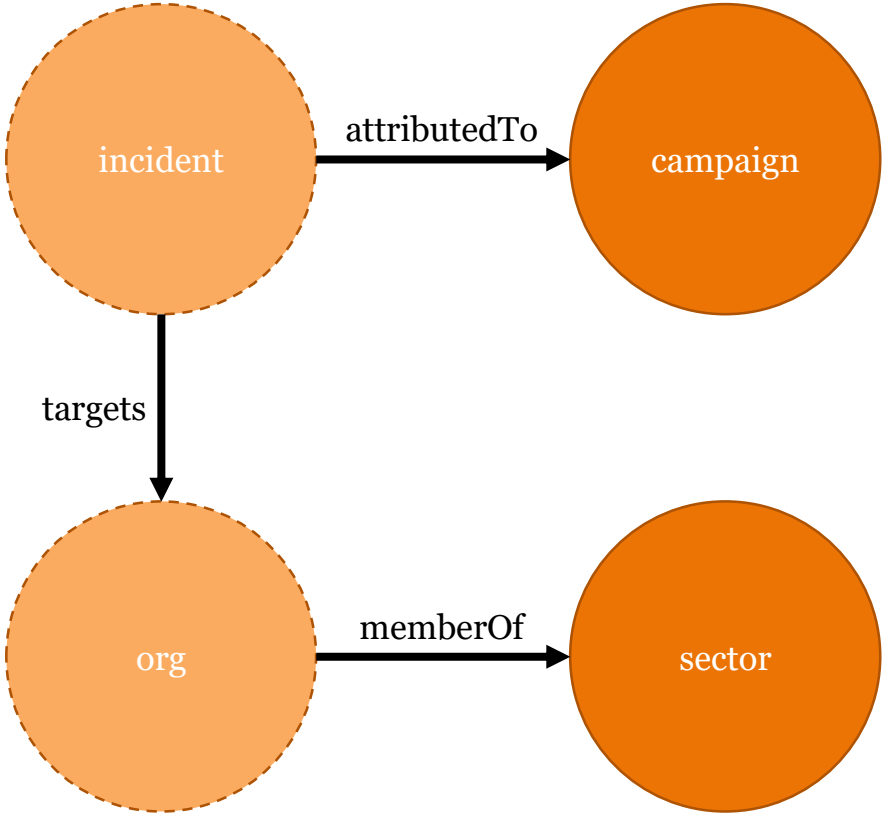
Example: Campaign targets sector



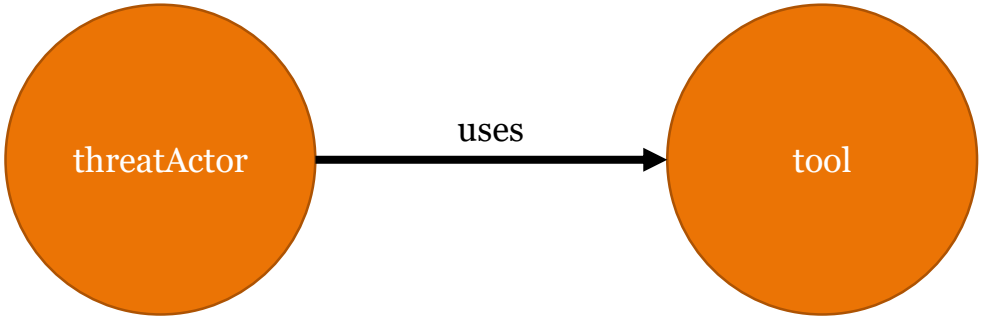
Example: Campaign targets sector



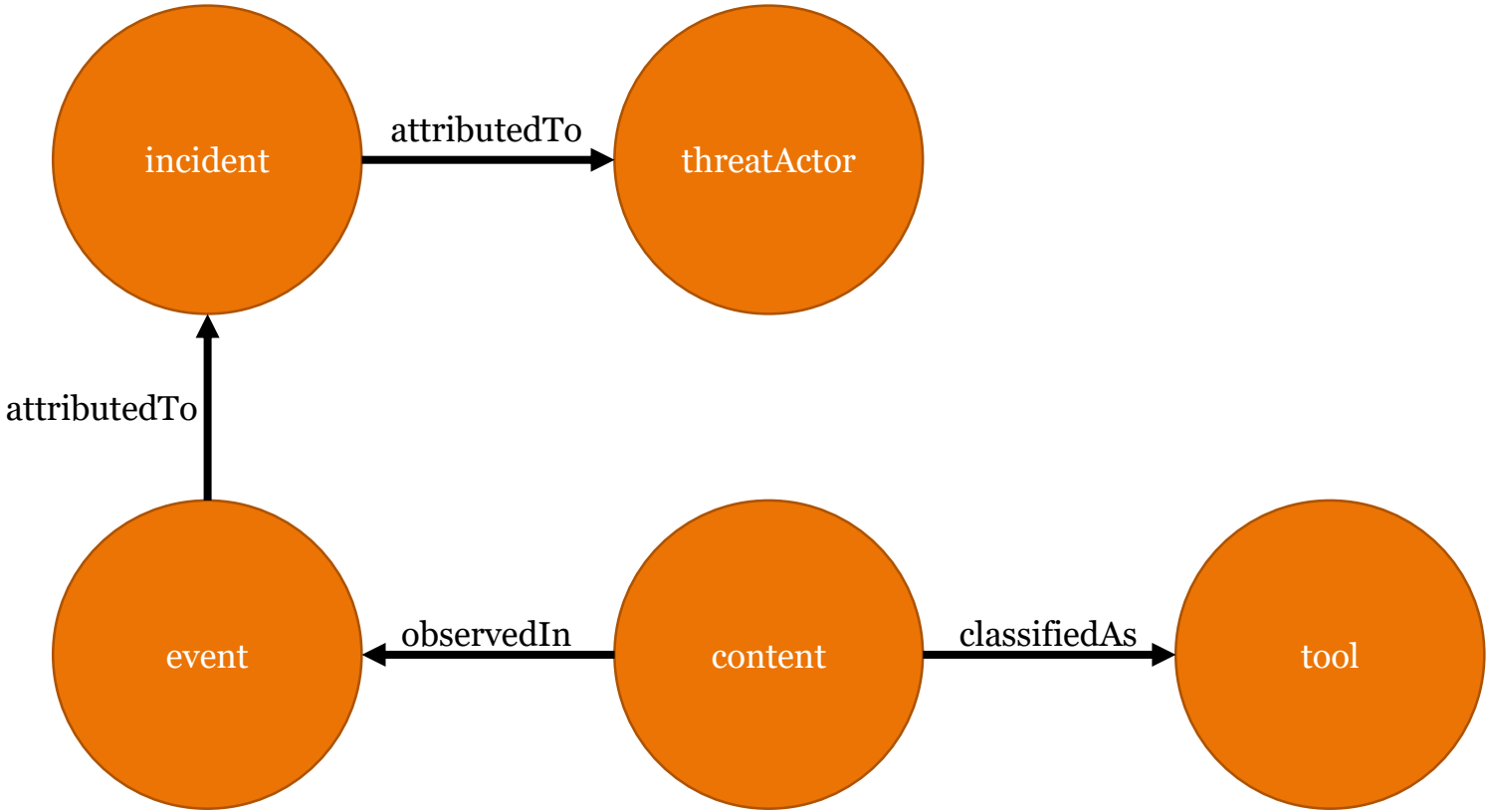
Example: Campaign targets sector



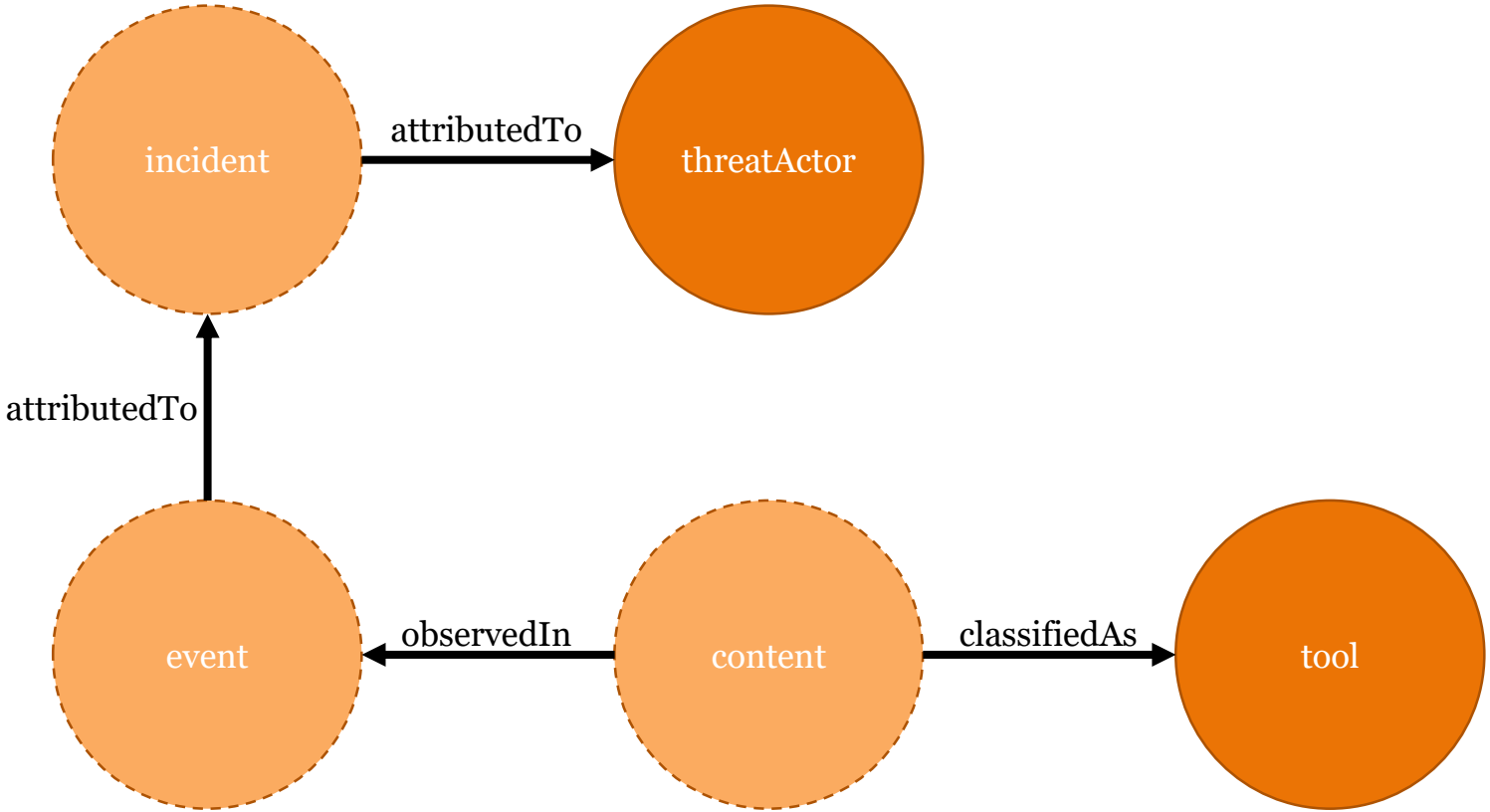
Example: Threat actor uses tool (ATT&CK)



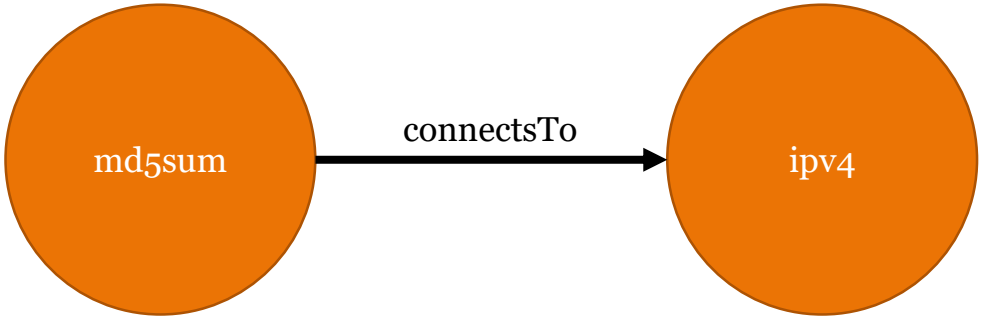
Example: Threat actor uses tool (ATT&CK)



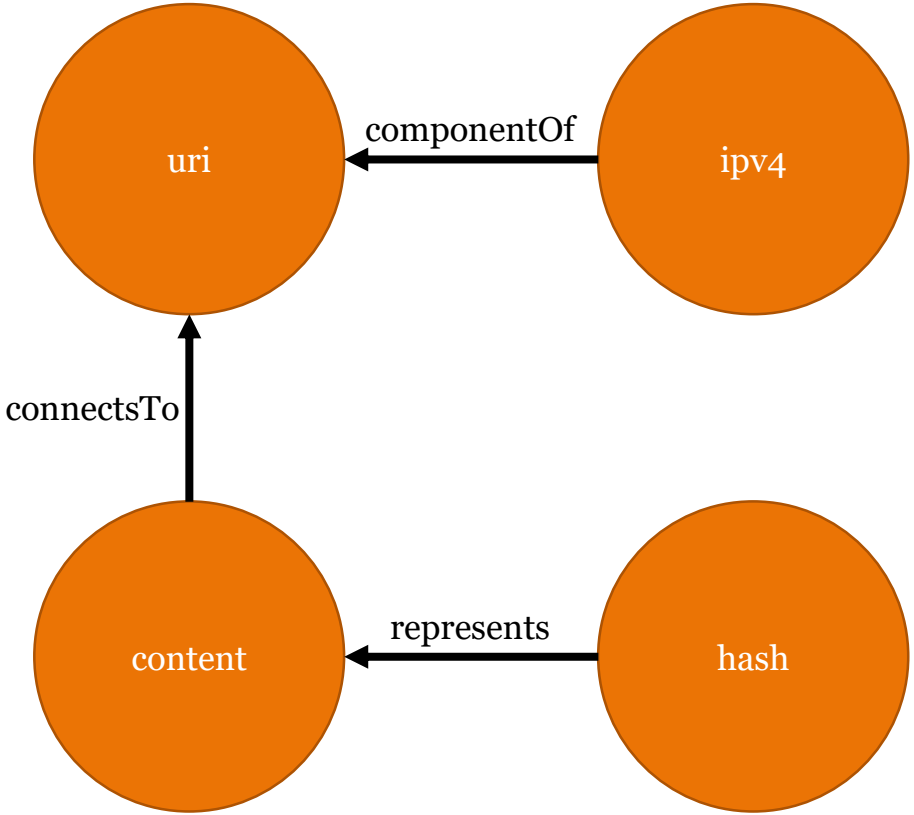
Example: Threat actor uses tool (ATT&CK)



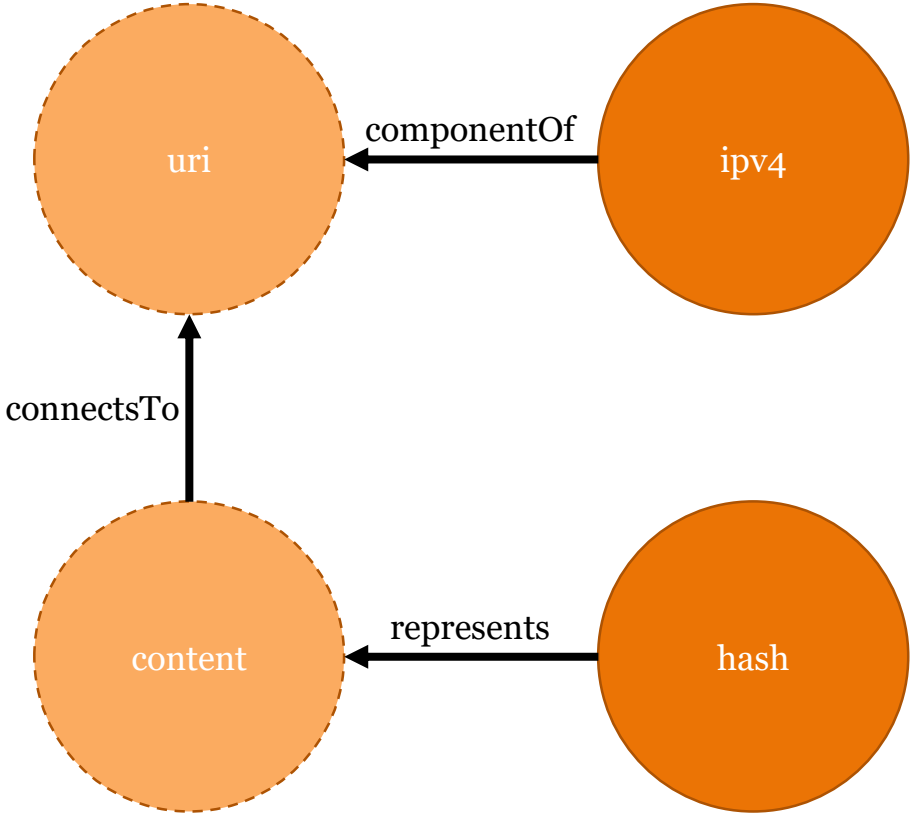
Example: md5sum connects to ipv4



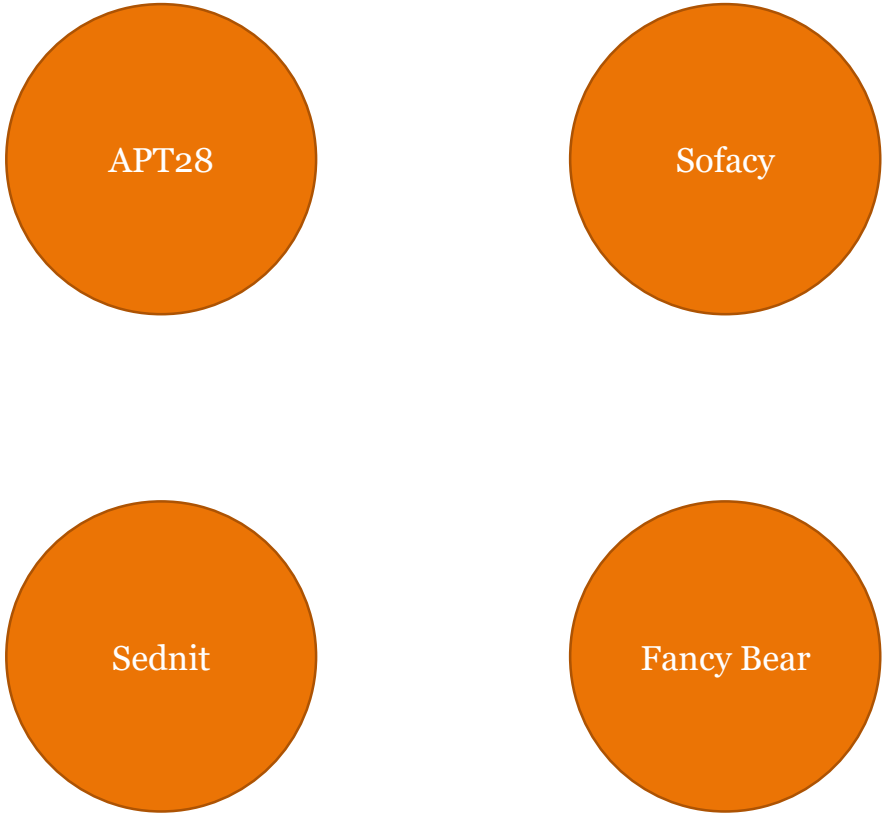
Example: md5sum connects to ipv4



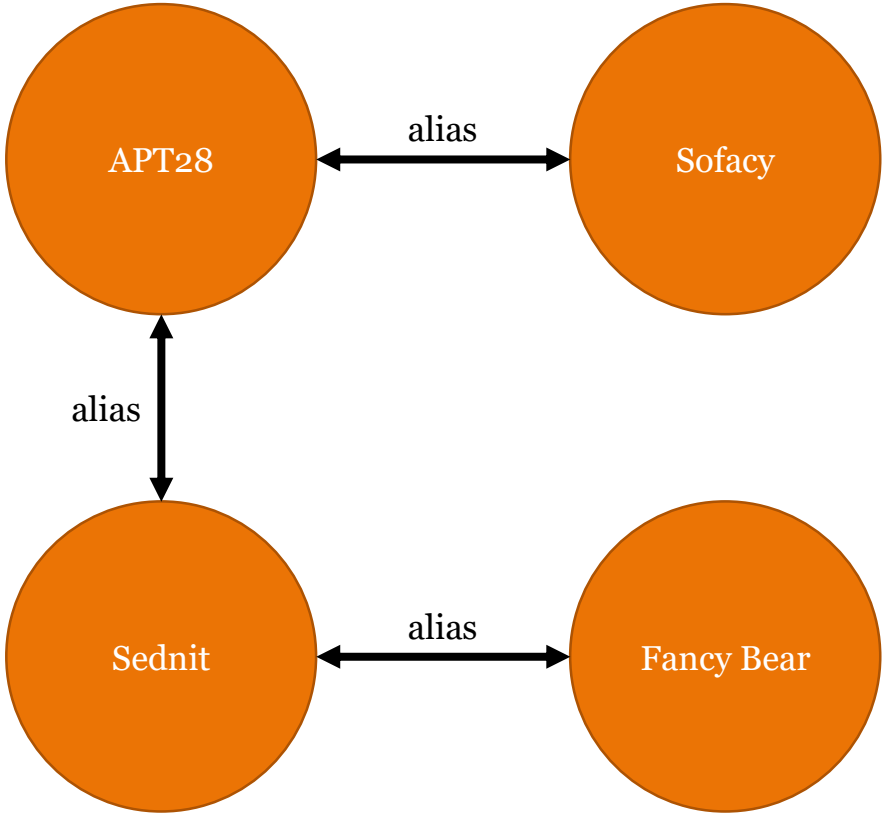
Example: md5sum connects to ipv4



Different names

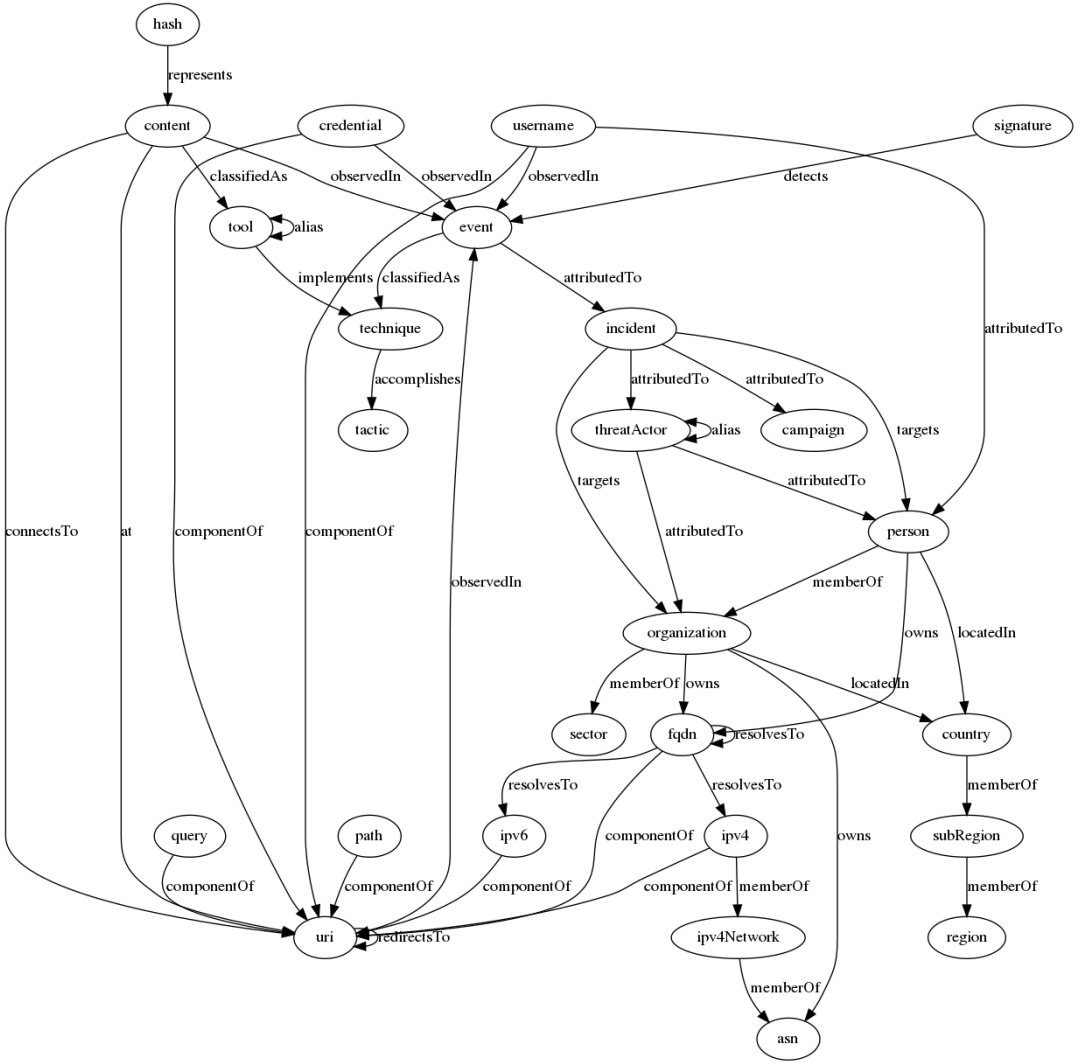


Different names



Data Model

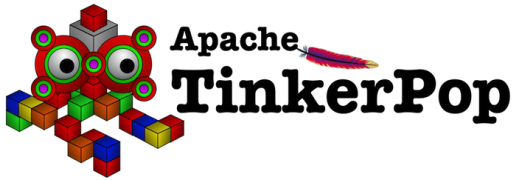
- Objects
  - Global
  - Example: IP address
- Facts
  - Connected to one or two objects
  - Immutable
  - Timestamped
  - Owner
  - Role-based and explicit access control
  - Backed by evidence and comments
- Placeholders

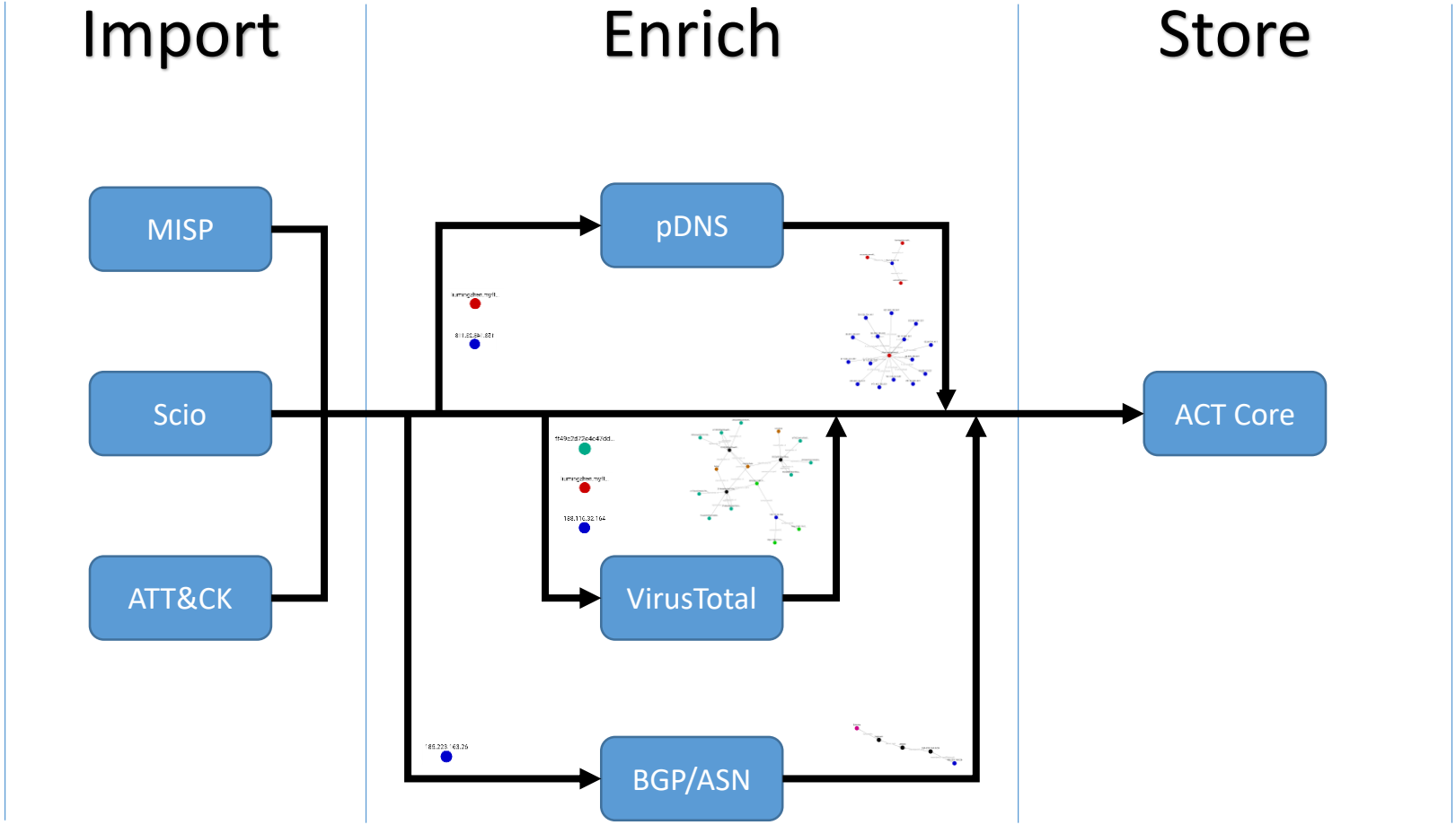


THE ACT PLATFORM

mnemonic

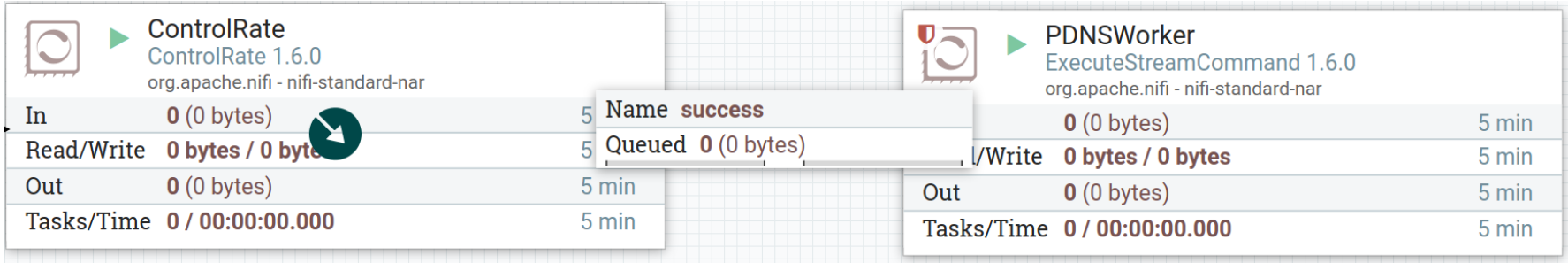
Platform Architecture – Core technologies

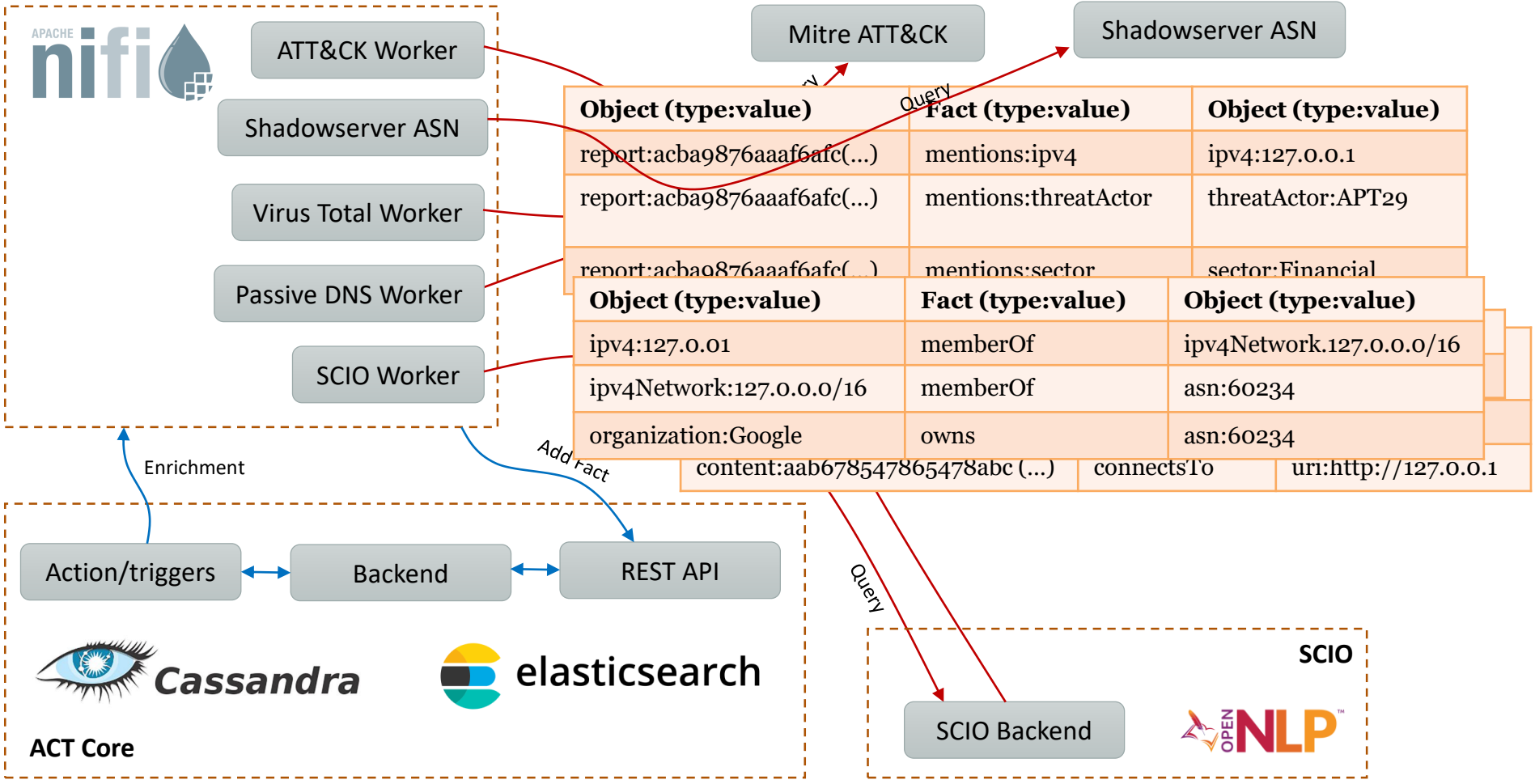



Import
Enrich
Store
MISP
Scio
ATT&CK
pDNS
VirusTotal
BGP/ASN
ACT Core

Platform Architecture – Workflow orchestration

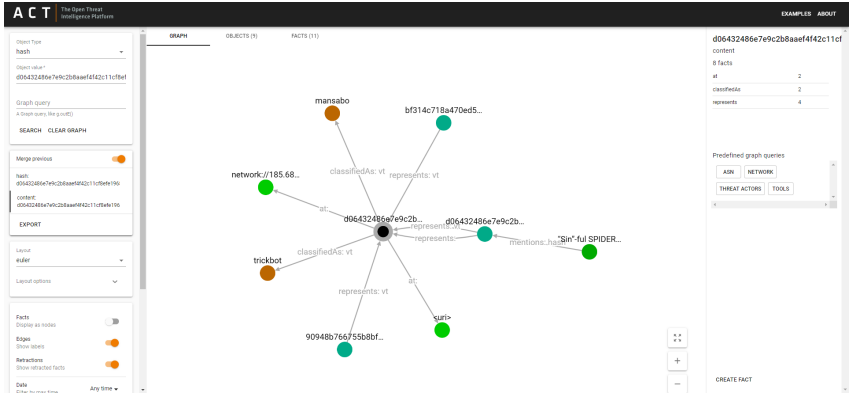
- Originally developed by NSA
- Open sourced and transferred to the Apache Foundation in 2014
- Manage flows of data supporting a large number of inputs and outputs:
  - HTTP, FTP, SCP, Kafka, Elasticsearch, JMS, Syslog, MongoDB, Hadoop, Cassandra, SMTP, POP3, etc




APACHE nifi
ATT&CK Worker
Shadowserver ASN
Virus Total Worker
Passive DNS Worker
SCIO Worker
Mitre ATT&CK
Shadowserver ASN
Query
Object (type:value)
Fact (type:value)
Object (type:value)
report:acba9876aaaf6afc(...)
mentions:ipv4
ipv4:127.0.0.1
report:acba9876aaaf6afc(...)
mentions:threatActor
threatActor:APT29
report:acba9876aaaf6afc(...)
mentions:sector
sector:Financial
Object (type:value)
Fact (type:value)
Object (type:value)
ipv4:127.0.01
memberOf
ipv4Network.127.0.0.0/16
ipv4Network:127.0.0.0/16
memberOf
asn:60234
organization:Google
owns
asn:60234
content:aab678547865478abc (...)
connectsTo
uri:http://127.0.0.1
Enrichment
Add Fact
Action/triggers
Backend
REST API
Cassandra
elasticsearch
ACT Core
Query
SCIO
SCIO Backend
OPEN NLP
mnemonic

Platform Architecture – Graph database

- Looked into existing graph databases, but they lacked proper fine granular permissions (and many of them had commercial licenses that could not be used in the research project)
- Apache Tinkerpop implemented on top of Cassandra/Elasticsearch
- Graph queries opens up a range of possibilites that is not possible on a flat data structure



Backend

REST API

GUI

Apache TinkerPop

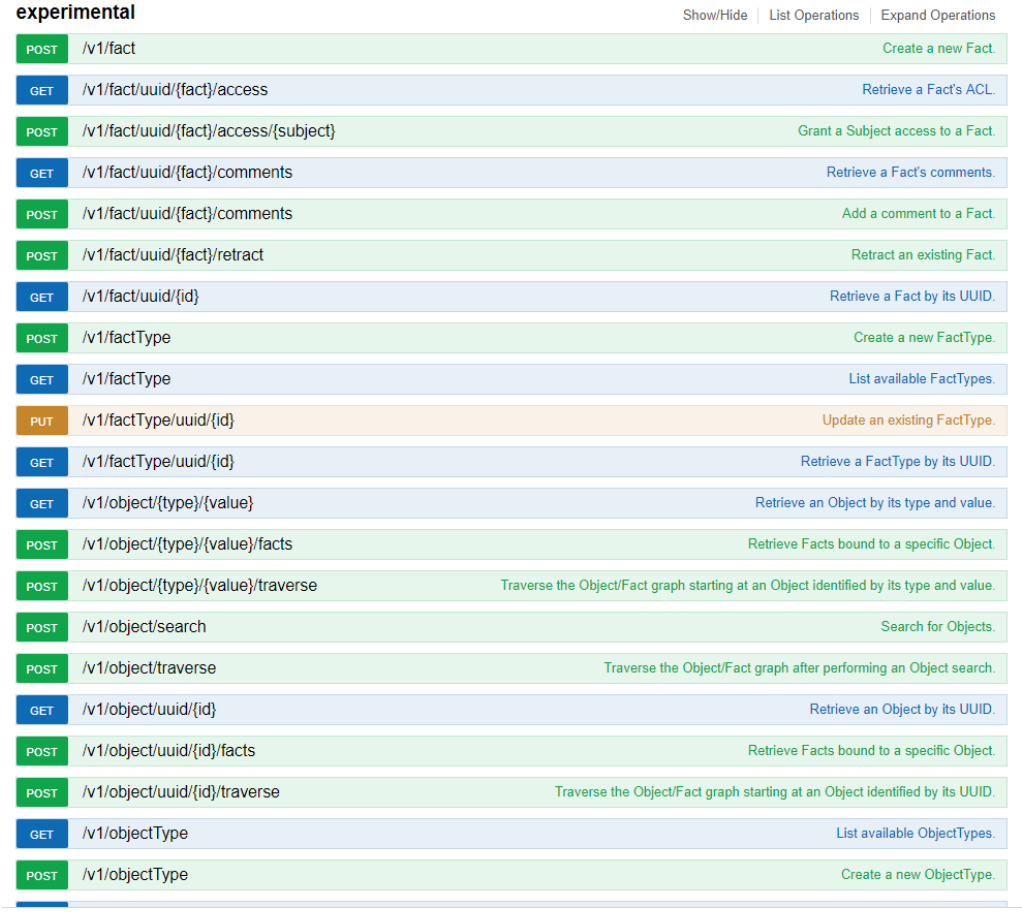
Cassandra

elasticsearch

ACT Core

mnemonic

API - Swagger



experimental

Show/Hide | List Operations | Expand Operations

POST	/v1/fact	Create a new Fact.
GET	/v1/fact/uuid/{fact}/access	Retrieve a Fact's ACL.
POST	/v1/fact/uuid/{fact}/access/{subject}	Grant a Subject access to a Fact.
GET	/v1/fact/uuid/{fact}/comments	Retrieve a Fact's comments.
POST	/v1/fact/uuid/{fact}/comments	Add a comment to a Fact.
POST	/v1/fact/uuid/{fact}/retract	Retract an existing Fact.
GET	/v1/fact/uuid/{id}	Retrieve a Fact by its UUID.
POST	/v1/factType	Create a new FactType.
GET	/v1/factType	List available FactTypes.
PUT	/v1/factType/uuid/{id}	Update an existing FactType.
GET	/v1/factType/uuid/{id}	Retrieve a FactType by its UUID.
GET	/v1/object/{type}/{value}	Retrieve an Object by its type and value.
POST	/v1/object/{type}/{value}/facts	Retrieve Facts bound to a specific Object.
POST	/v1/object/{type}/{value}/traverse	Traverse the Object/Fact graph starting at an Object identified by its type and value.
POST	/v1/object/search	Search for Objects.
POST	/v1/object/traverse	Traverse the Object/Fact graph after performing an Object search.
GET	/v1/object/uuid/{id}	Retrieve an Object by its UUID.
POST	/v1/object/uuid/{id}/facts	Retrieve Facts bound to a specific Object.
POST	/v1/object/uuid/{id}/traverse	Traverse the Object/Fact graph starting at an Object identified by its UUID.
GET	/v1/objectType	List available ObjectTypes.
POST	/v1/objectType	Create a new ObjectType.

API – Python library (act-api on pypi)

Navigation

Project description

Release history

Download files

Project links

Homepage

Statistics

View statistics for this project via Libraries.io, or by using Google BigQuery

Meta

License: ISC License (ISCL) (MIT)

Author: mnemonic AS

ACT, mnemonic

Project description

python-act

python-act is a library used to connect to the ACT platform.

The platform has a REST api, and the goal of this library is to expose all functionality in the API.

Objects and Facts

The act platform is built on two basic types, the object and fact.

Objects are universal elements that can be referenced uniquely by its value. An example of an object can be an IP address.

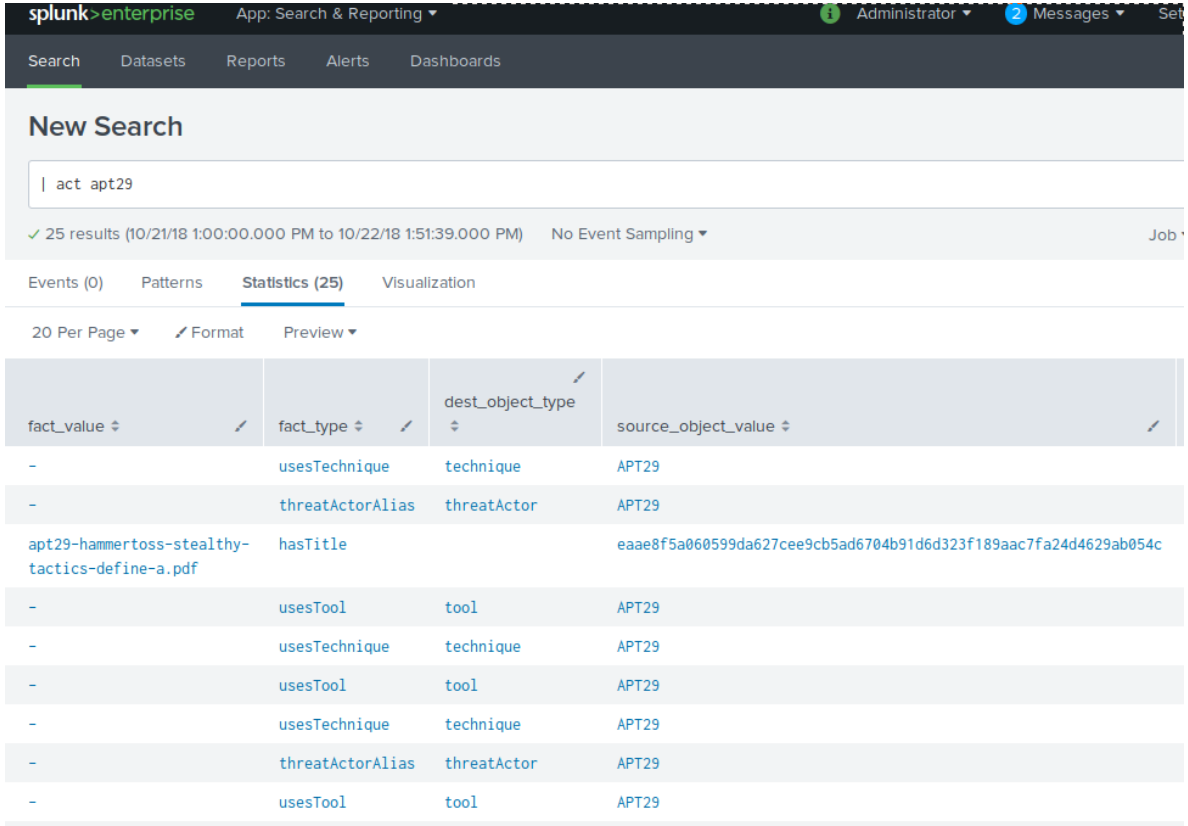
Facts are assertions or obsersvations that ties objects together. A fact may or may not have a value desribing further the fact.

Facts can be linked on or more objects. Below, the seenIn fact is linked to both an ipv4 object and report object, but the hasTitle fact is only linked to a report.

Object type	Object value	Fact type	Fact value	Object type	Object value
ipv4	127.0.0.1	seenIn	report	report	cbc80bb5c0c0f8944bf73(...)
report	cbc80bb5c0c0f8944bf73(...)	hasTitle	Threat Intel Summary	*n/a*	*n/a*

mnemonic

Splunk Add-on - Queries



fact_value	fact_type	dest_object_type	source_object_value
-	usesTechnique	technique	APT29
-	threatActorAlias	threatActor	APT29
apt29-hammertoss-stealthy-tactics-define-a.pdf	hasTitle		eaae8f5a060599da627cee9cb5ad6704b91d6d323f189aac7fa24d4629ab054c
-	usesTool	tool	APT29
-	usesTechnique	technique	APT29
-	usesTool	tool	APT29
-	usesTechnique	technique	APT29
-	threatActorAlias	threatActor	APT29
-	usesTool	tool	APT29

Splunk Add-on – Annotate search results

```
1  source="carbanak.csv" dest_ip=179.43.140.82 | acta dest_ip
2  | table dest_ip usesC2* seenIn*
```

All time ▾

✓ 3 events (before 10/22/18 2:27:42.000 PM) No Event Sampling ▾ Job ▾ Smart Mc

Events Patterns **Statistics (3)** Visualization

20 Per Page ▾ Format Preview ▾

dest_ip	usesC2:ipv4	seenIn:report
179.43.140.82	c6ec176592ea26c4ee27974273e592ff 188f261e5fca94bd1fc1edc1aafee8c0 6e9408c338e98a8bc166a8d4f8264019	9c624e51ffab866aaa73c41f944f7ec6045ec6c04a99e24b37eadd518b74780c 2d460cb6523158909dad07e6b0f9491339ce4ce1550f64832b0c5396c2f5bb8f
179.43.140.82	c6ec176592ea26c4ee27974273e592ff 188f261e5fca94bd1fc1edc1aafee8c0 6e9408c338e98a8bc166a8d4f8264019	9c624e51ffab866aaa73c41f944f7ec6045ec6c04a99e24b37eadd518b74780c 2d460cb6523158909dad07e6b0f9491339ce4ce1550f64832b0c5396c2f5bb8f
179.43.140.82	c6ec176592ea26c4ee27974273e592ff 188f261e5fca94bd1fc1edc1aafee8c0 6e9408c338e98a8bc166a8d4f8264019	9c624e51ffab866aaa73c41f944f7ec6045ec6c04a99e24b37eadd518b74780c 2d460cb6523158909dad07e6b0f9491339ce4ce1550f64832b0c5396c2f5bb8f

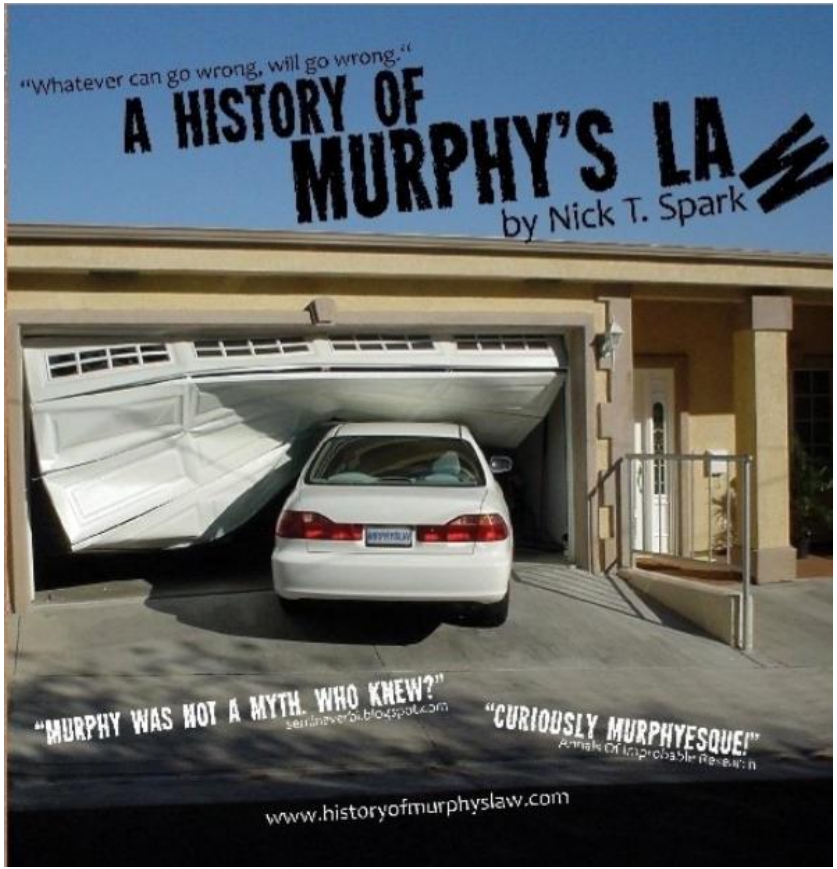
Threat Intelligence Platform - Summary

- Github repositories
  - https://github.com/mnemonic-no/act (installation docs)
  - https://github.com/mnemonic-no/act-api-python
  - https://github.com/mnemonic-no/act-bootstrap
  - https://github.com/mnemonic-no/act-frontend
  - https://github.com/mnemonic-no/act-platform
  - https://github.com/mnemonic-no/act-scio
  - https://github.com/mnemonic-no/act-splunk
  - https://github.com/mnemonic-no/act-triggers
  - https://github.com/mnemonic-no/act-workers
  - License: ISC (BSD compatible)
- Public AWS instance
  - https://act-eu1.mnemonic.no

TRAINING - INTRODUCTION

mnemonic

Before We Start



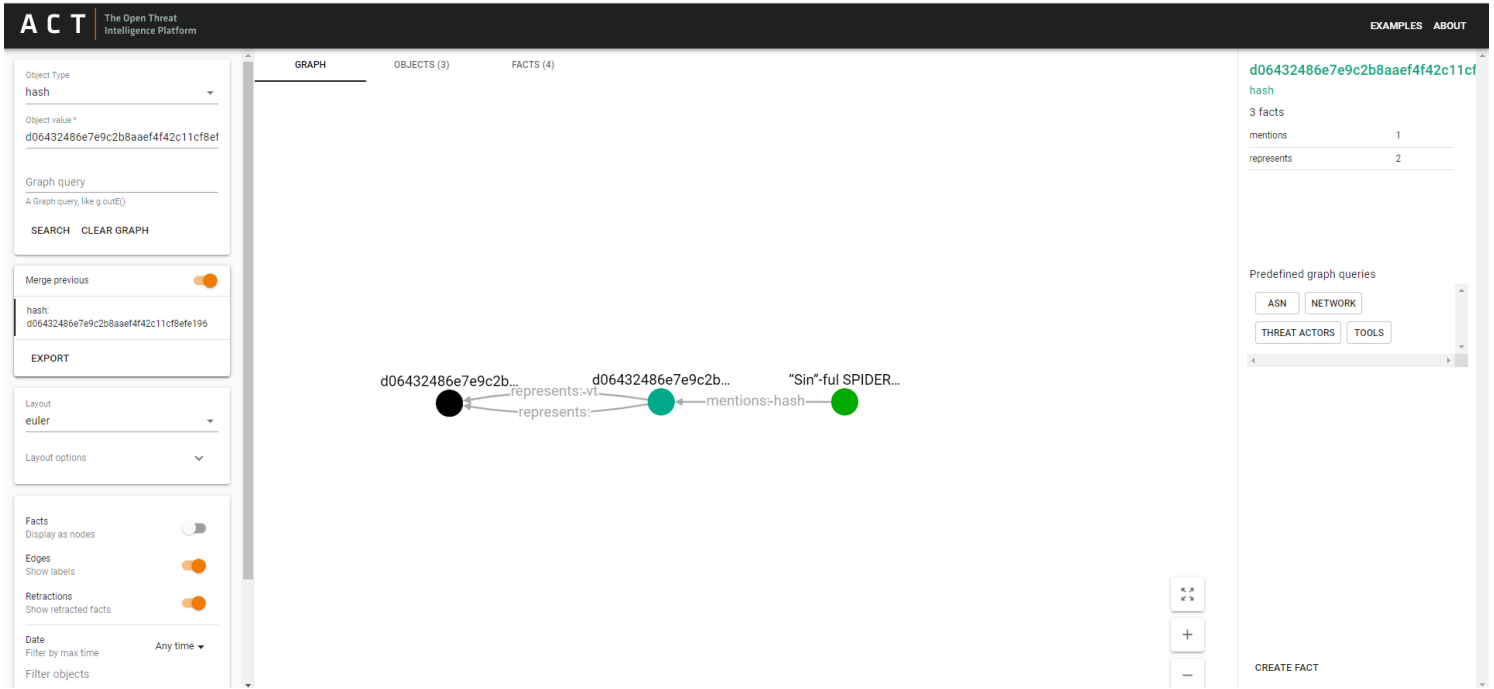
Accessing the read-only AWS instance

GUI:
https://act-eu1.mnemonic.no

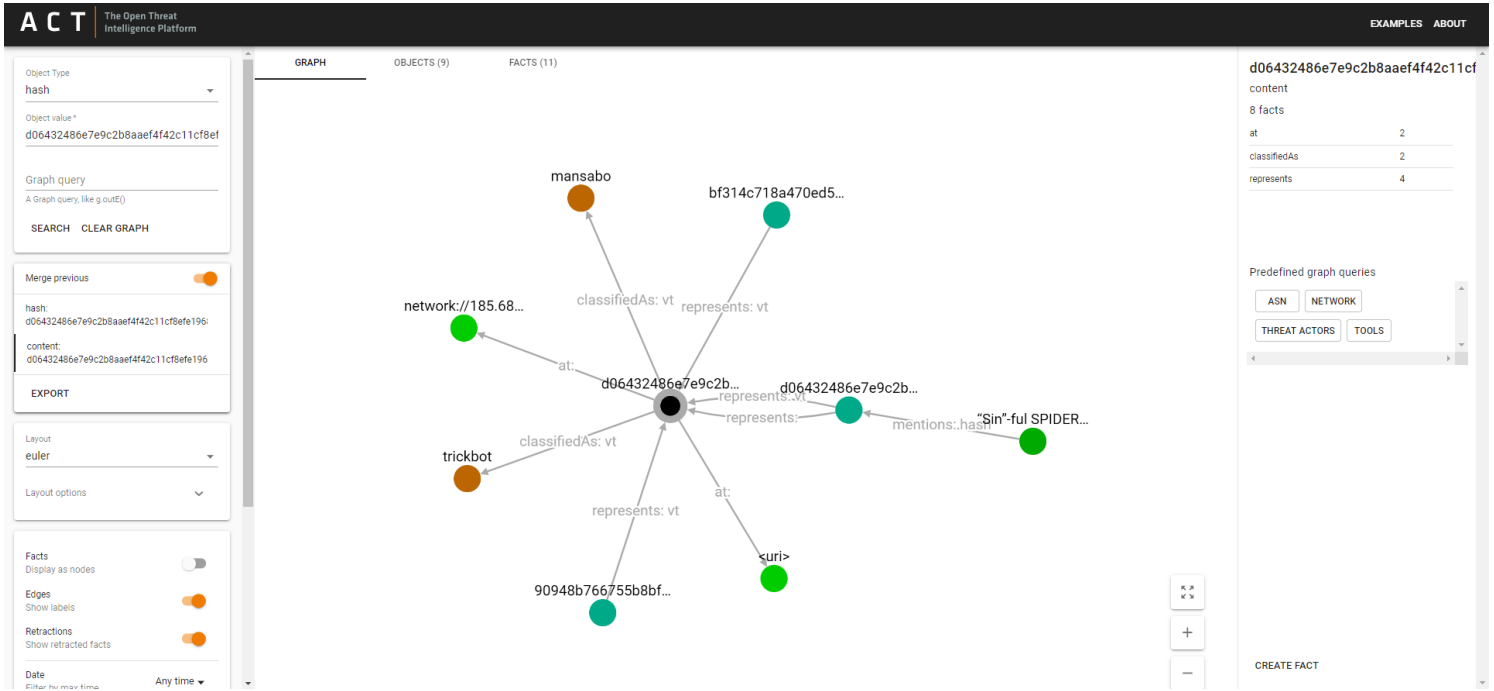
Tasks:
https://act-eu1.mnemonic.no/examples/

API:
https://act-eu1.mnemonic.no/swagger/

Introduction 1

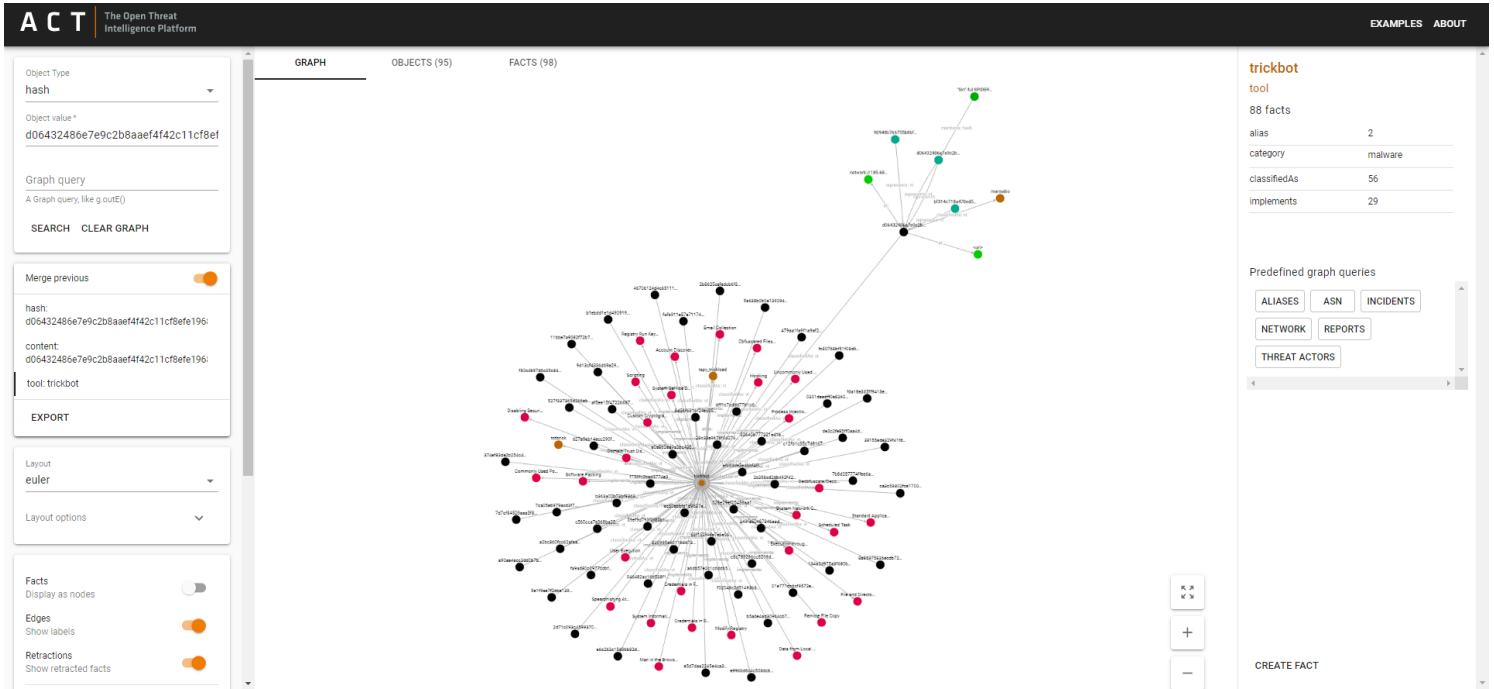


Introduction 1 – Click and Double-click

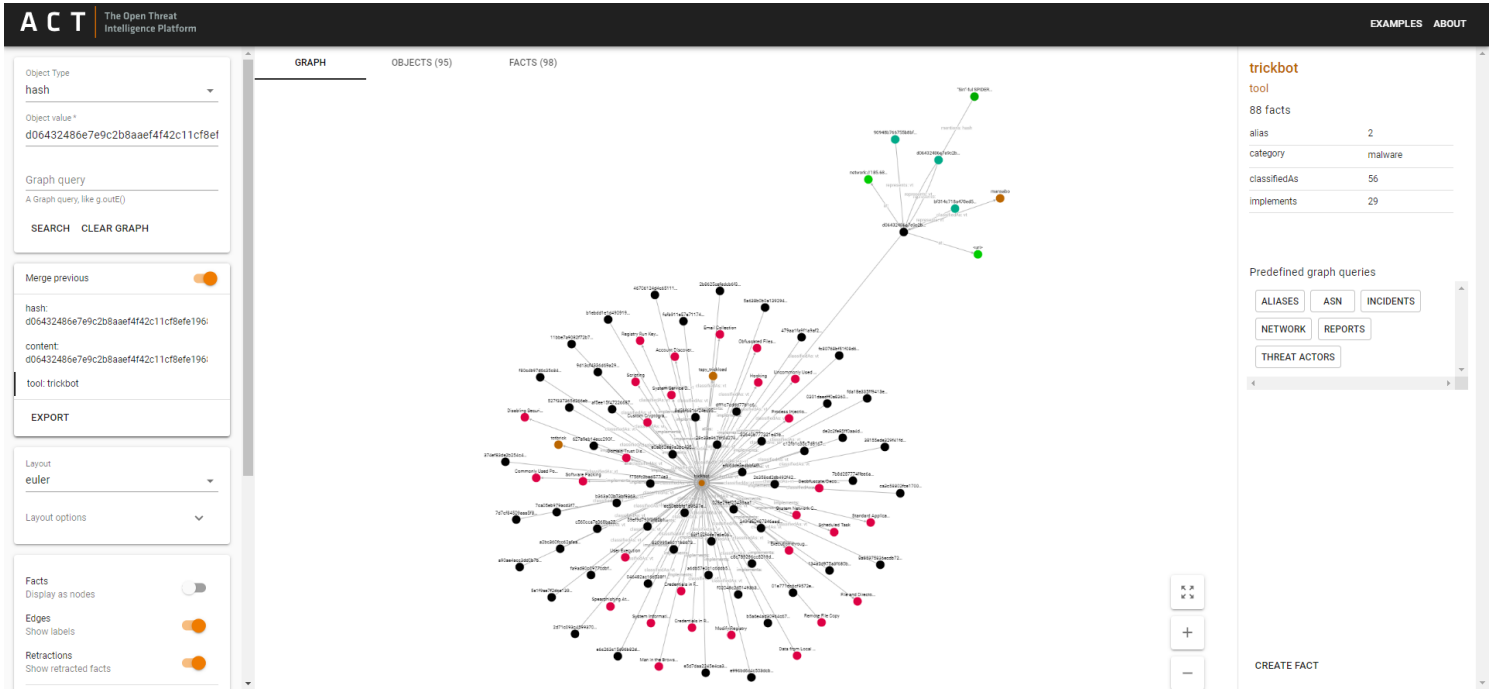


mnemonic

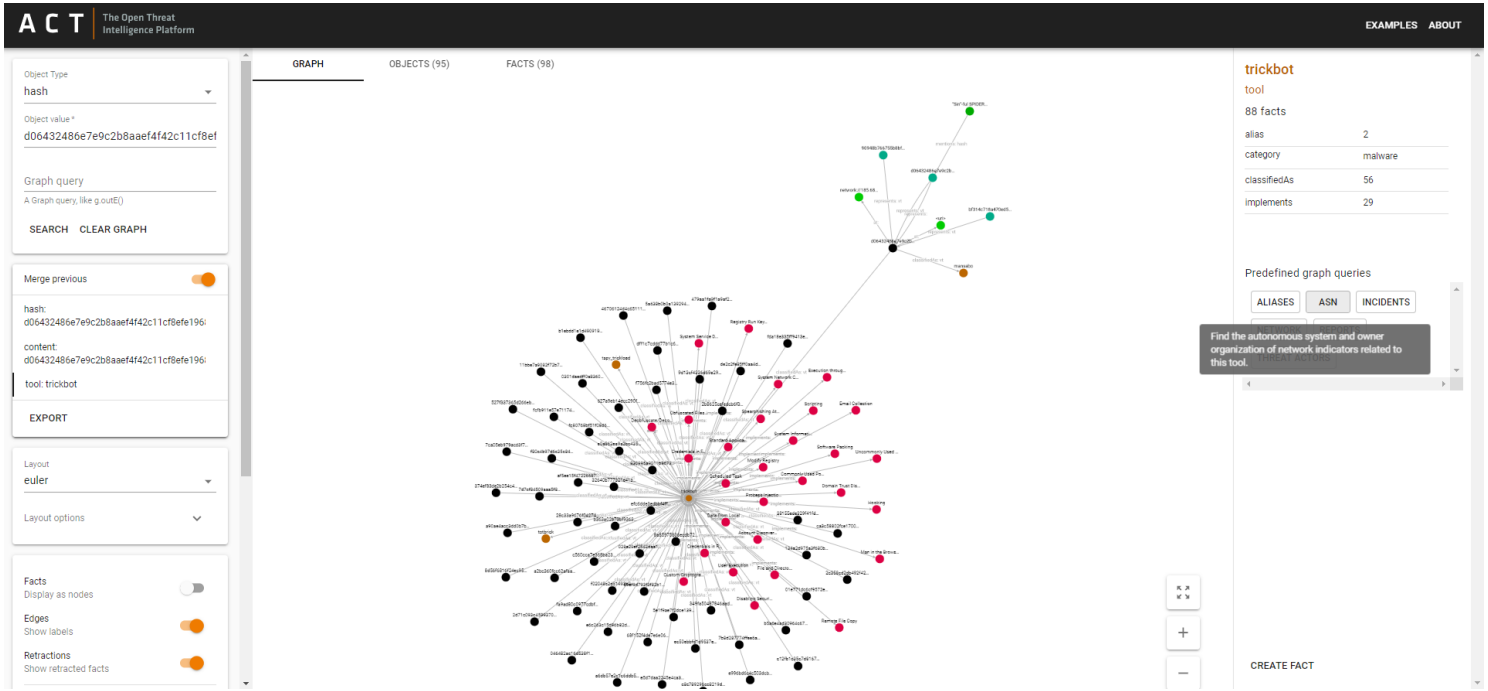
Introduction 1 – History, Layouts and Filtering



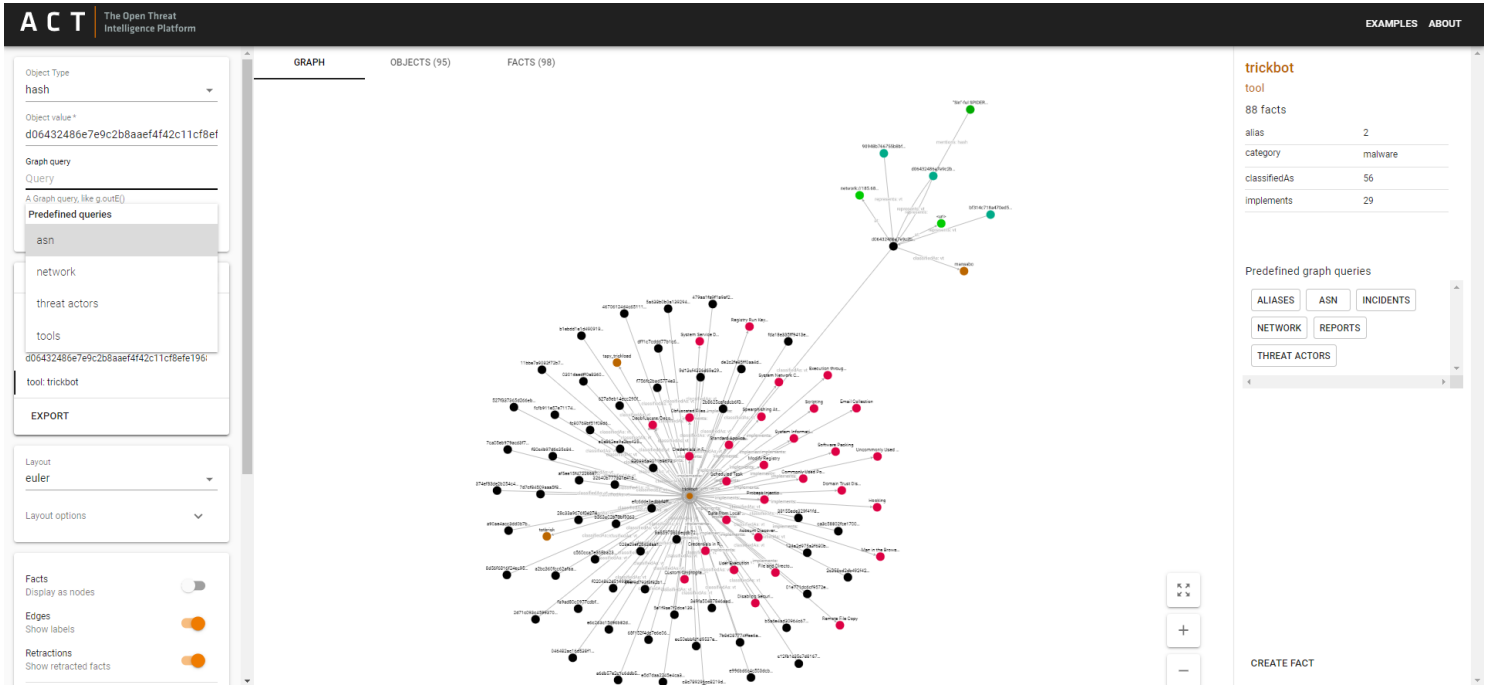
Introduction 1 – Fact Types



Introduction 1 – Graph Queries



Introduction 1 – Graph Queries



mnemonic

Introduction 2

Try the following object queries and explore the graph:

- threatActor: APT3
- tactic: lateral-movement
- tool: foosace
- ipv4: 153.148.23[.]118

Task 1

Try the following object query:

tool: remsec

Which threat actor is associated with this tool?
Which techniques are associated with this threat actor?
Can you find any reports that mention file hashes classified as remsec?

Task 2

Try the following object query:

ipv4: 188.116.32[.]164

Try to find reports, threat actors, tools and any other information related to this IP address.

Task 3

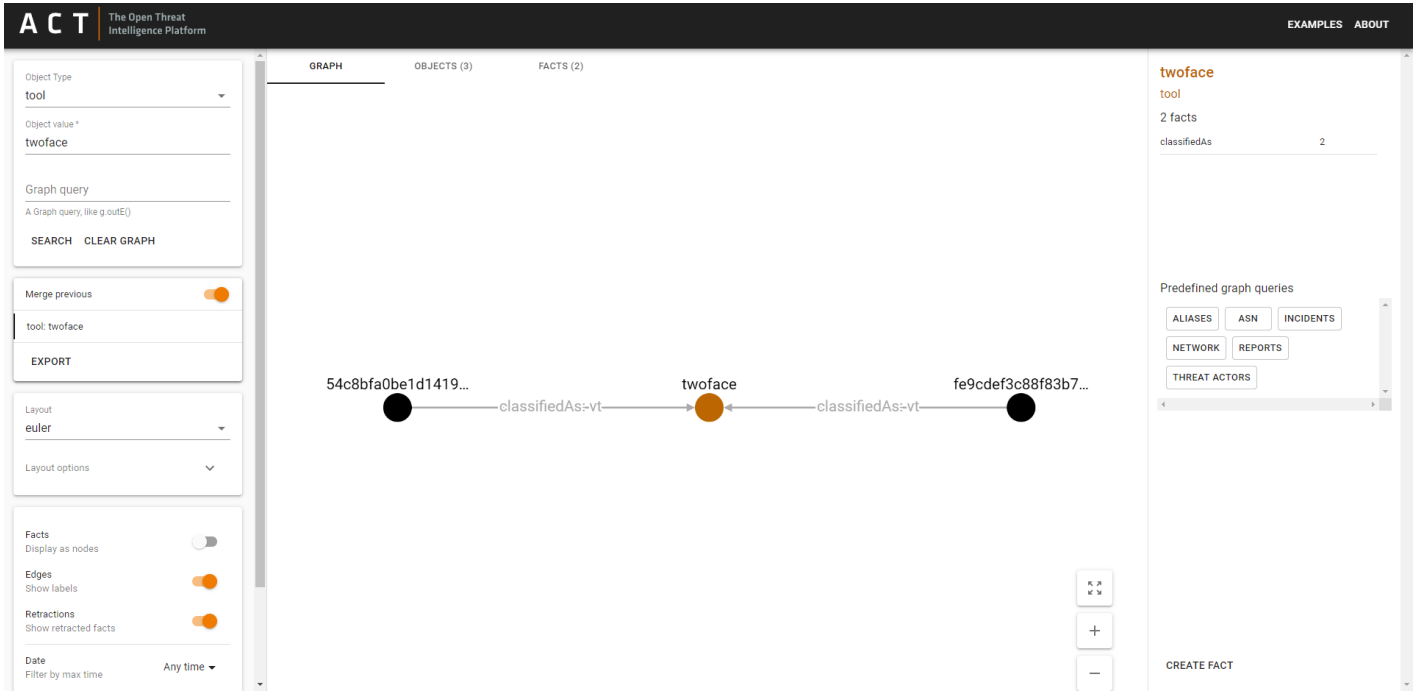
Explore Autonomous System Number 8048

- asn: 8048

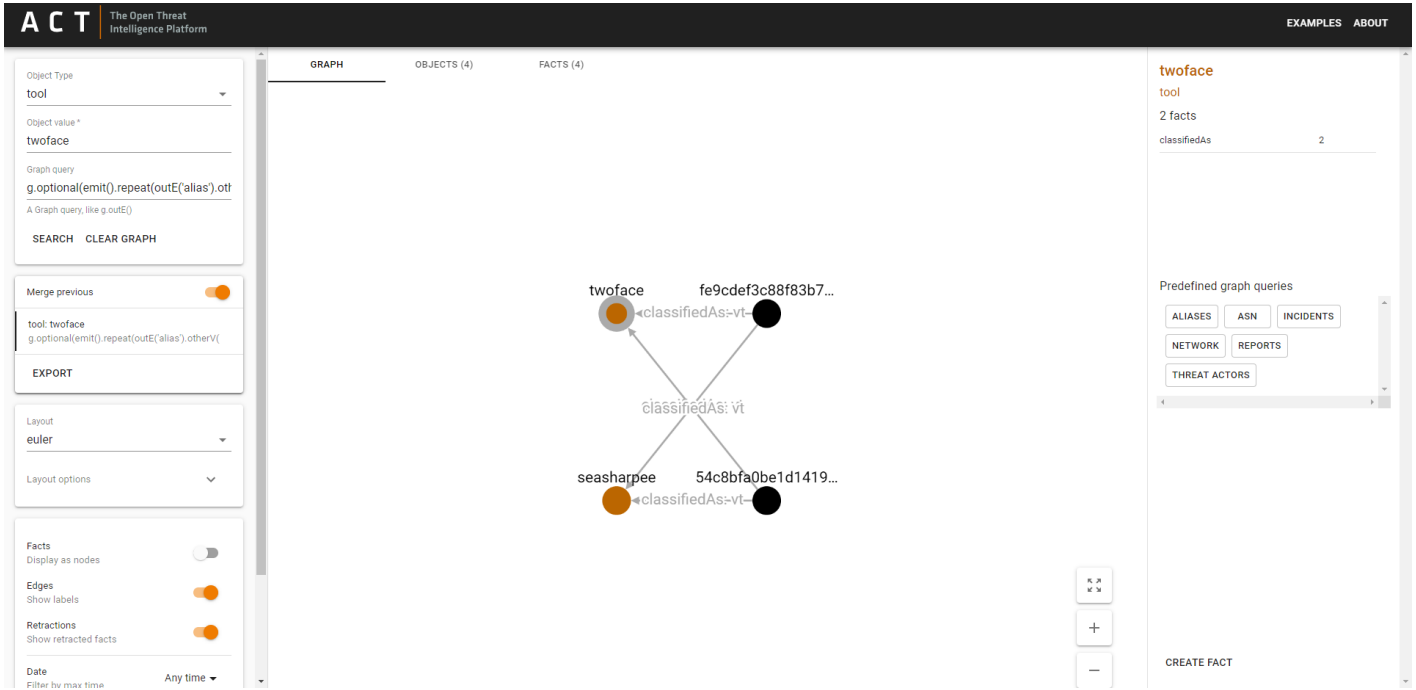
What kind of malicious behaviour has been observed from this AS?

Where is the organization that owns AS8048 located?

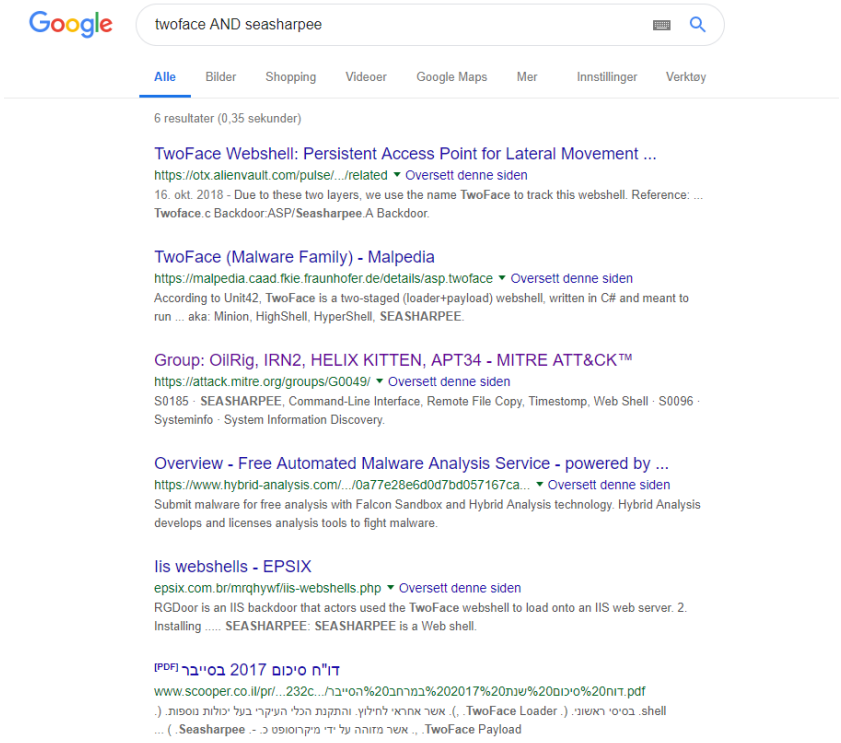
Introduction 3 – Aliases



Introduction 3 – Aliases

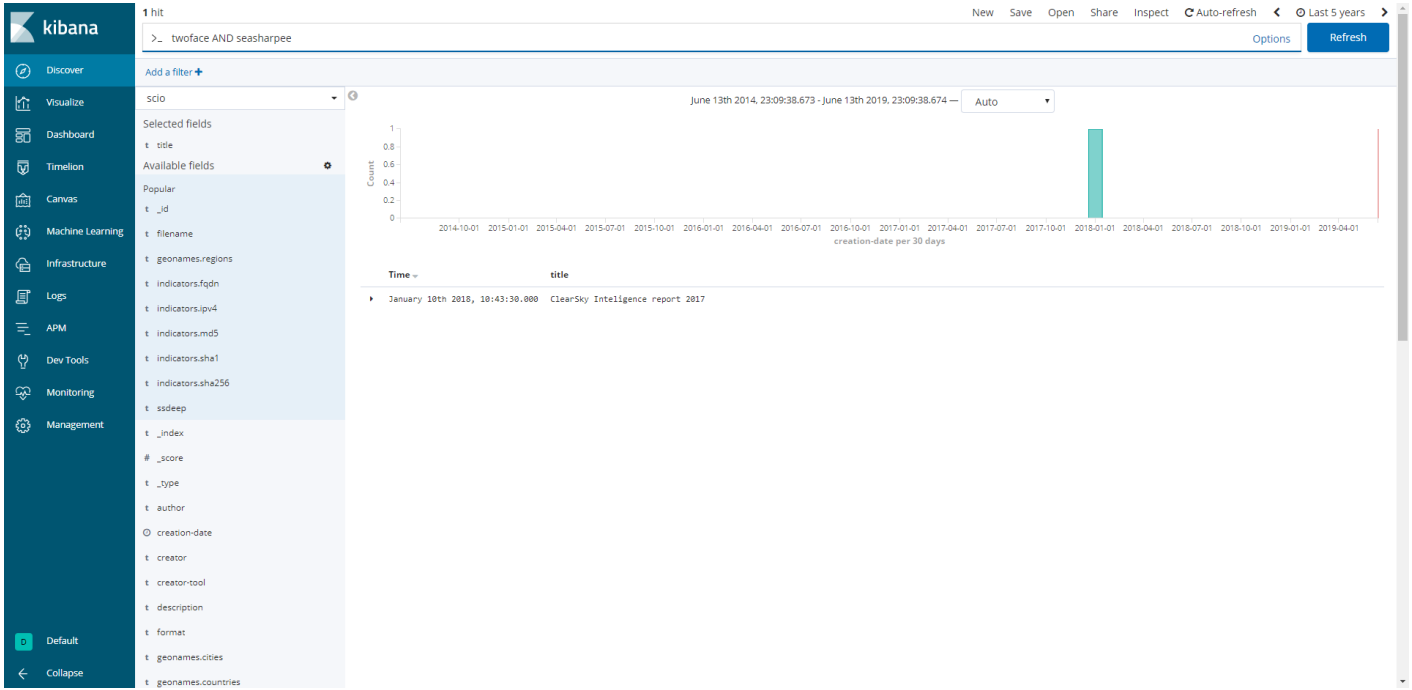


Introduction 3 – Aliases



mnemonic

Introduction 3 – Aliases

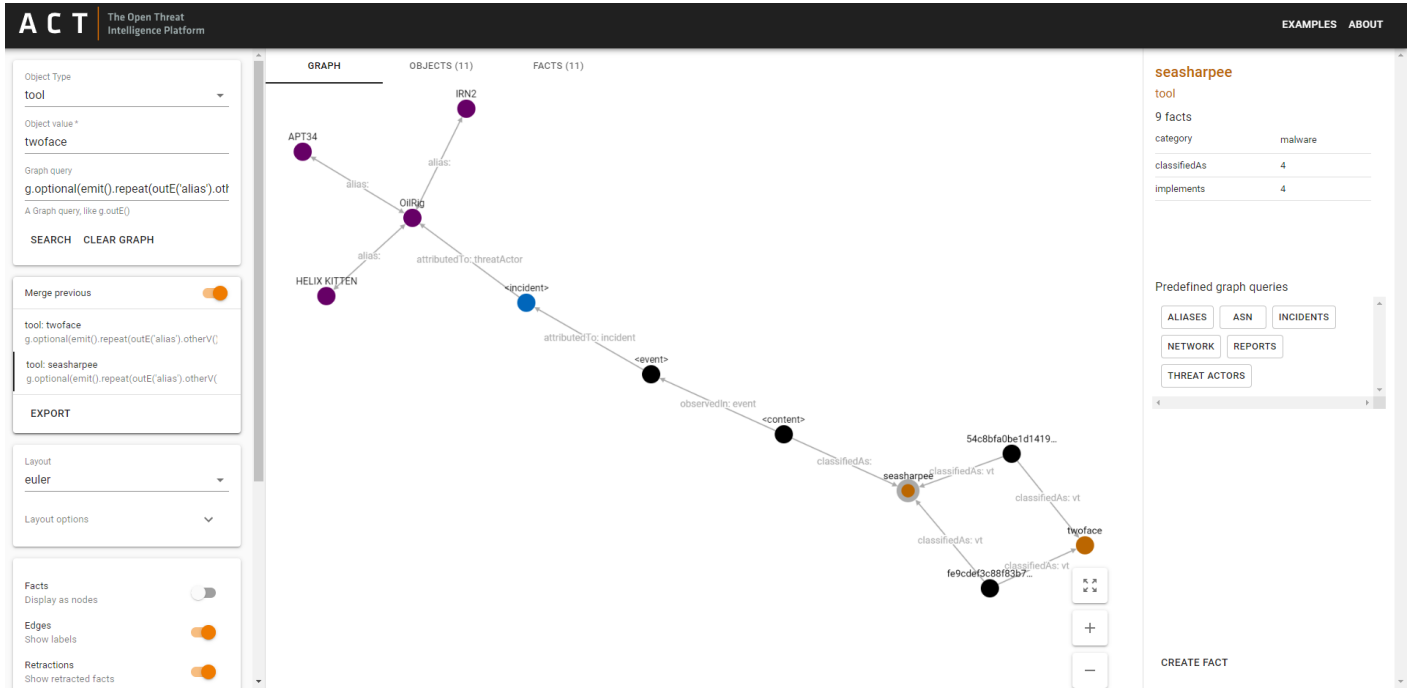


Introduction 3 – Aliases

https://www.clearskysec.com/wp-content/uploads/2018/01/ClearSky_cyber_intelligence_report_2017.pdf

“The Webshell is named **TwoFace** as it is comprised by two components. The first is named TwoFace Loader, a basic and preliminary shell that extracts and installs the second component, a more advances tool named TwoFace Payload (identified by Microsoft as **Seasharpee**). These tools are written in #C, and run on Webservers that support ASP.NET.”

Introduction 3 – Aliases



Task 4

Try to find an alias for the tool ‘gulpix’. Then try to find a publically available, credible source that confirms your findings.

ASSIGNMENTS

CASE STUDY

BREAKOUT: API, GRAPH QUERIES, EXPLORATION

mnemonic

Breakout: API/workers, graph queries, exploration

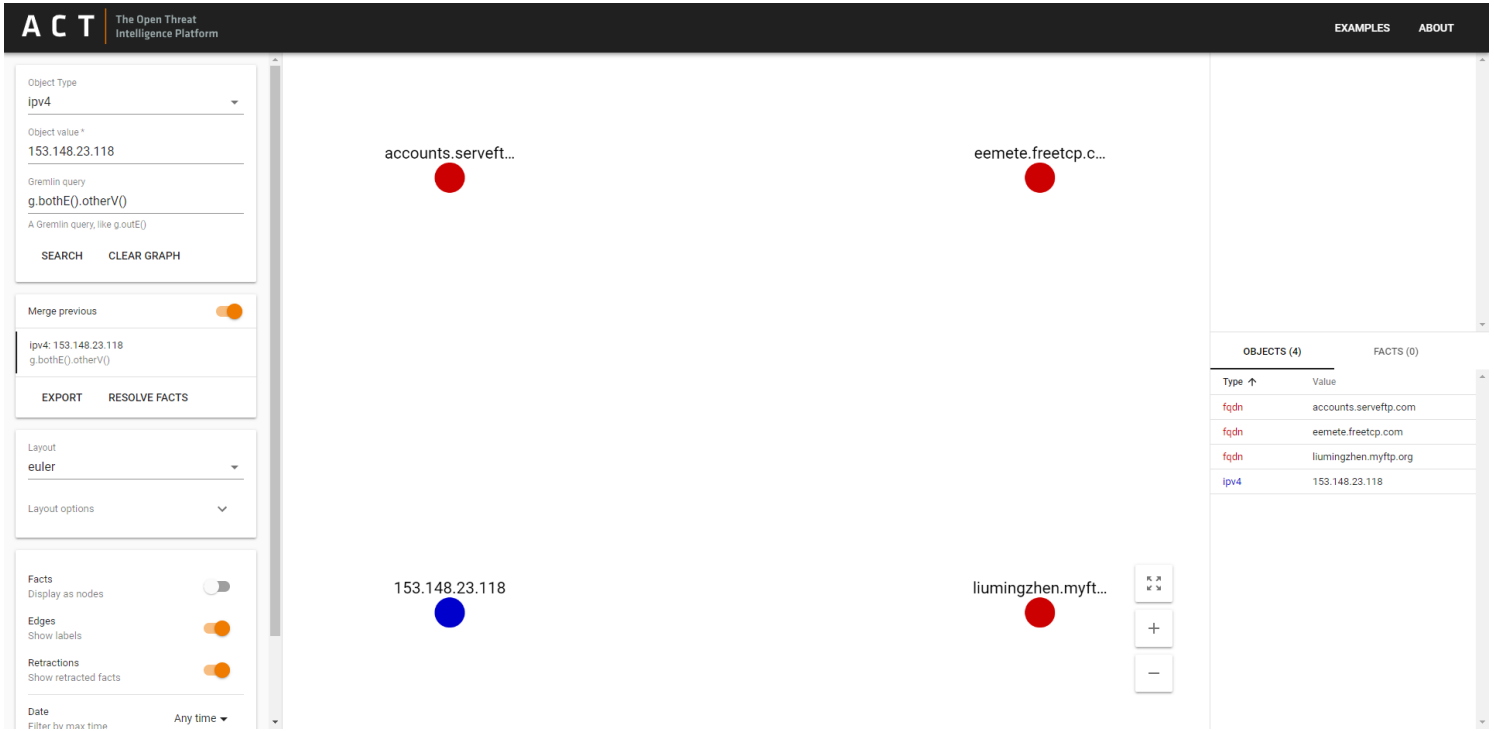
- API/workers - Geir
  - https://github.com/mnemonic-no/act-workshop-api
  - https://github.com/mnemonic-no/act-api-python
  - https://github.com/mnemonic-no/act-workers
- Graph queries - Martin
  - http://tinkerpop.apache.org/docs/current/reference/
  - https://github.com/mnemonic-no/act-frontend/blob/master/src/config.json
- Exploration

GRAPH QUERIES

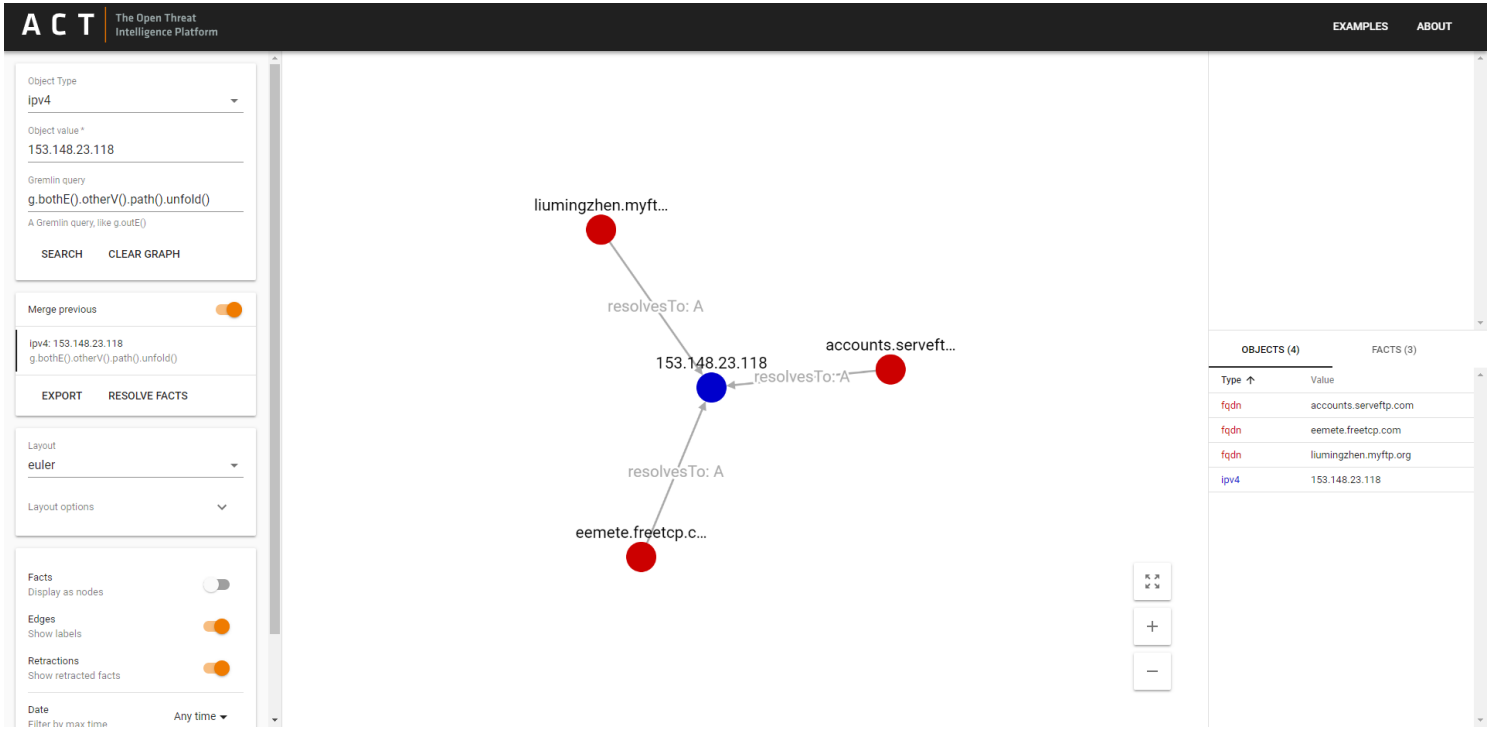
With Great Power Comes Great Responsibility

mnemonic

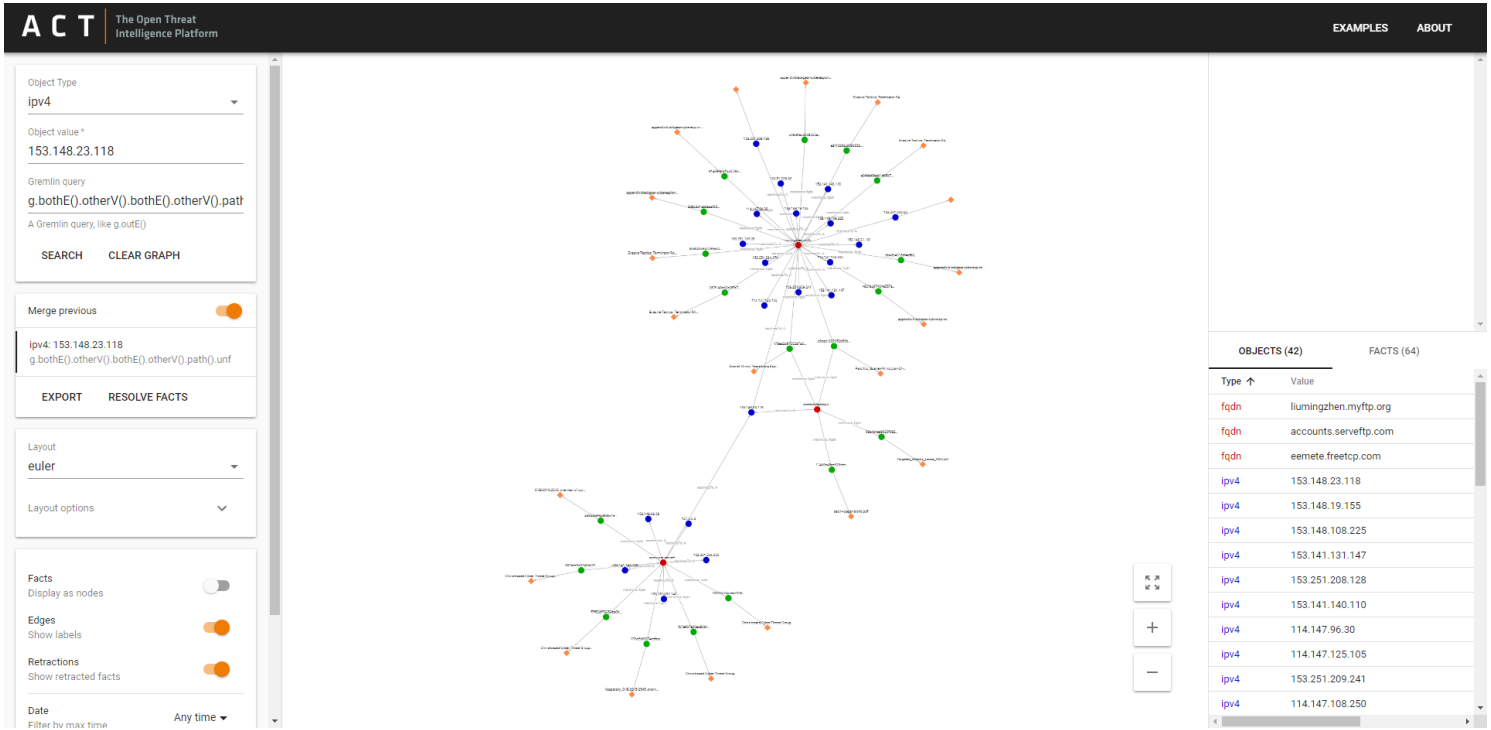
Graph Query 1



Graph Query 2 – Show Edges

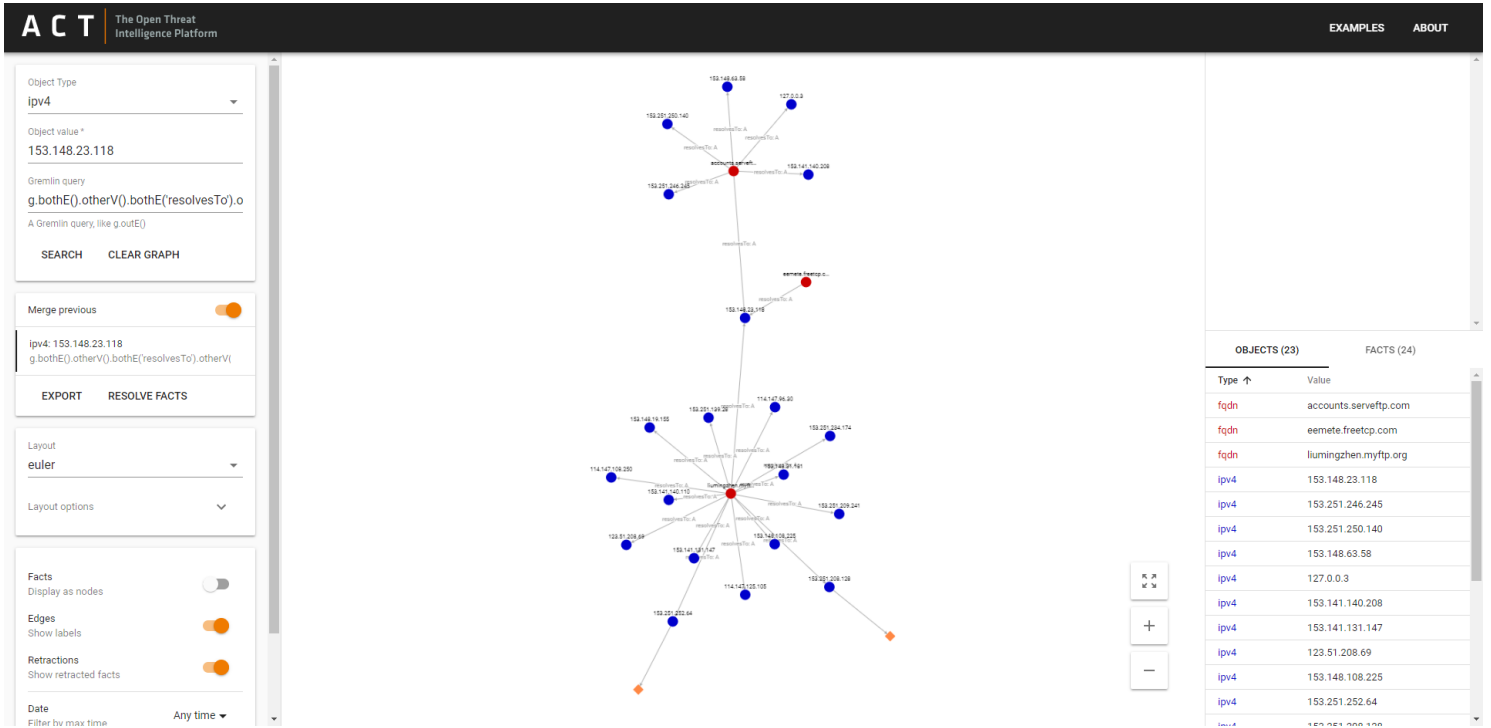


Graph Query 3 – 2 hops



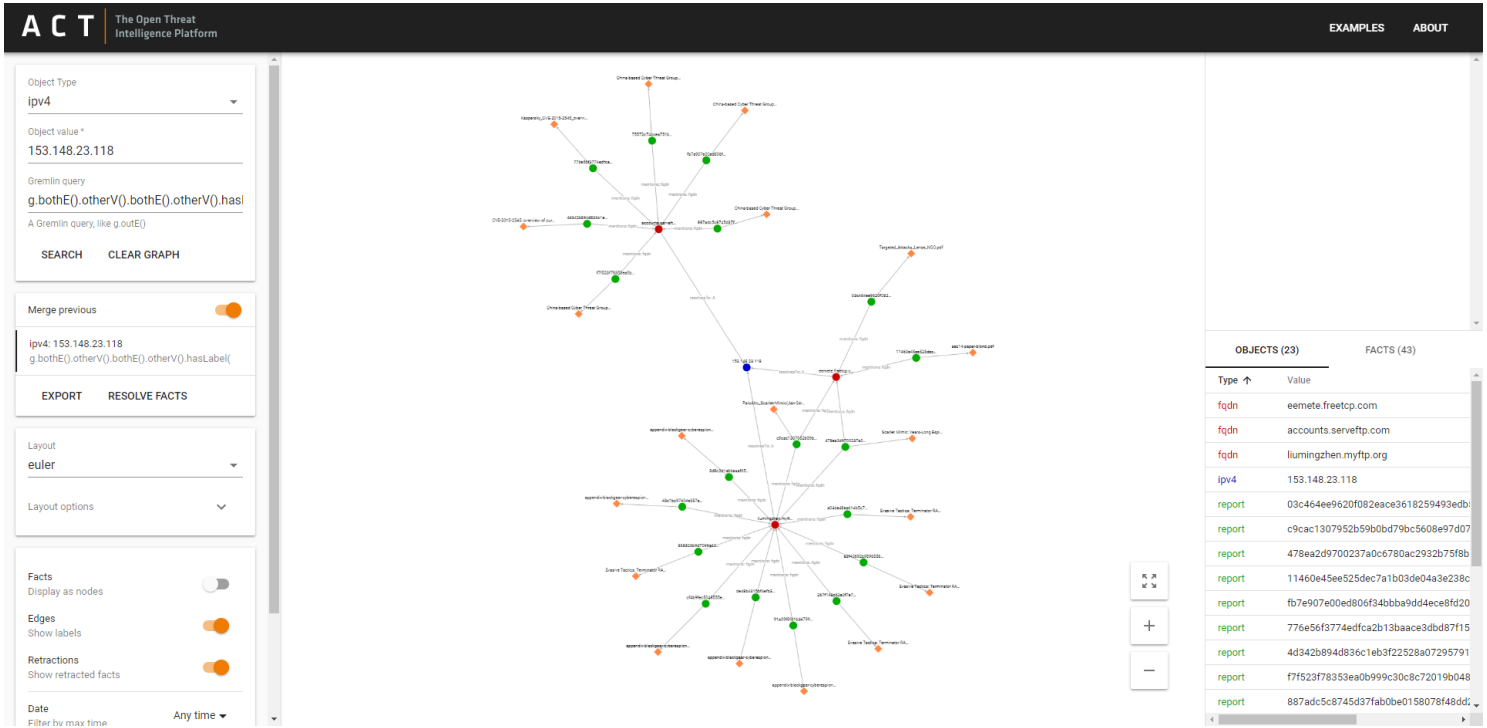
mnemonic

Graph Query 4 – Filter Edges (Facts)



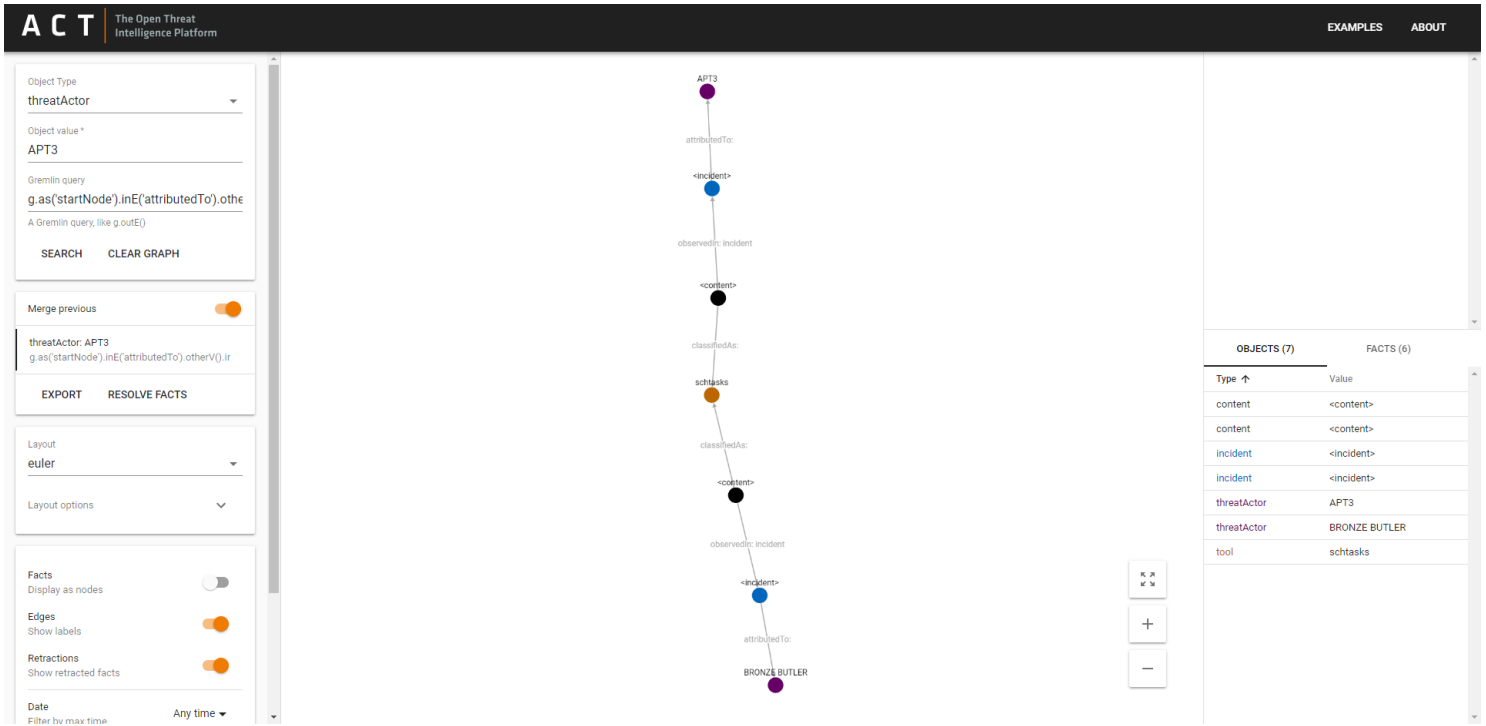
mnemonic

Graph Query 5 – Filter Nodes (Objects)



mnemonic

Graph Query 6 – Unique Tool Usage



mnemonic

Public Read-Only ACT Instance

https://act-eu1.mnemonic.no/examples/

FURTHER WORK

New Information Sources

- Security events
- Incidents
- Reputation lists
- Malware analysis systems
- STIX feeds
- …

Information Sharing

- Mechanism for sharing schema
- Format (STIX?)
- Trust models

Trust and Confidence

- Trust (source)
- Confidence (fact)
- Subjective Logic (quantify uncertainty)

GUI Improvements

- Timelines
- Share workspace
- Prune graph

GUI Improvements