SEMI-AUTOMATED CYBER THREAT INTELLIGENCE (ACT)

Training – FIRST CTI Symposium 2019 London

Martin Eian, Fredrik Borg, Geir Skjøtskift and Siri Bromander

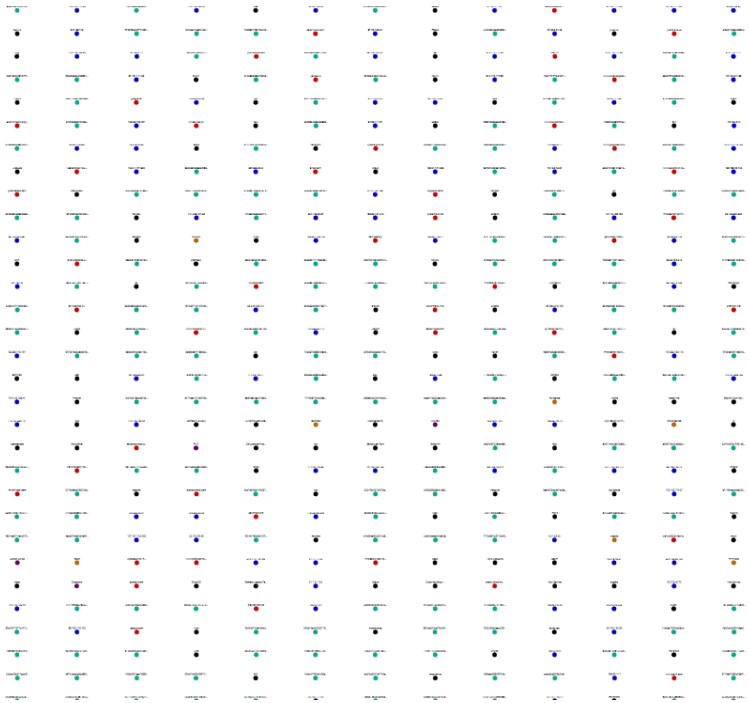
mnemonic

Goal

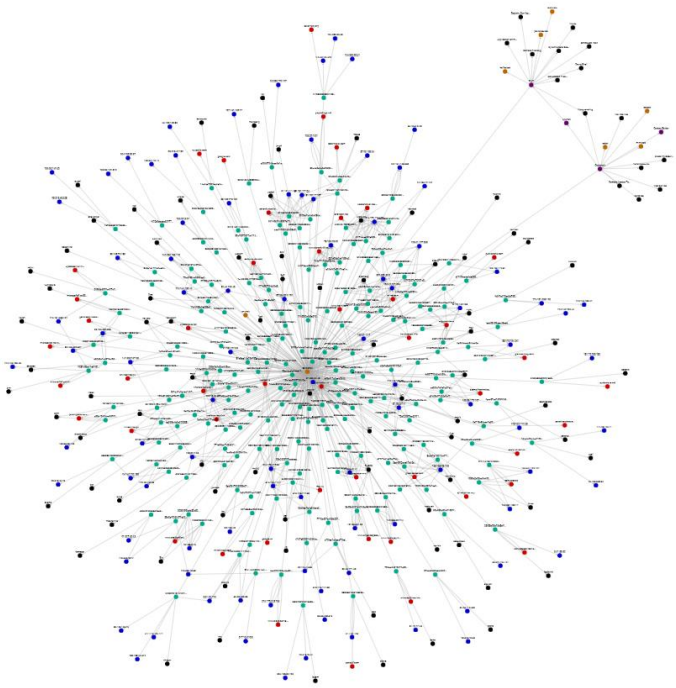
To collect and organize our knowledge of threats to make it useful

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

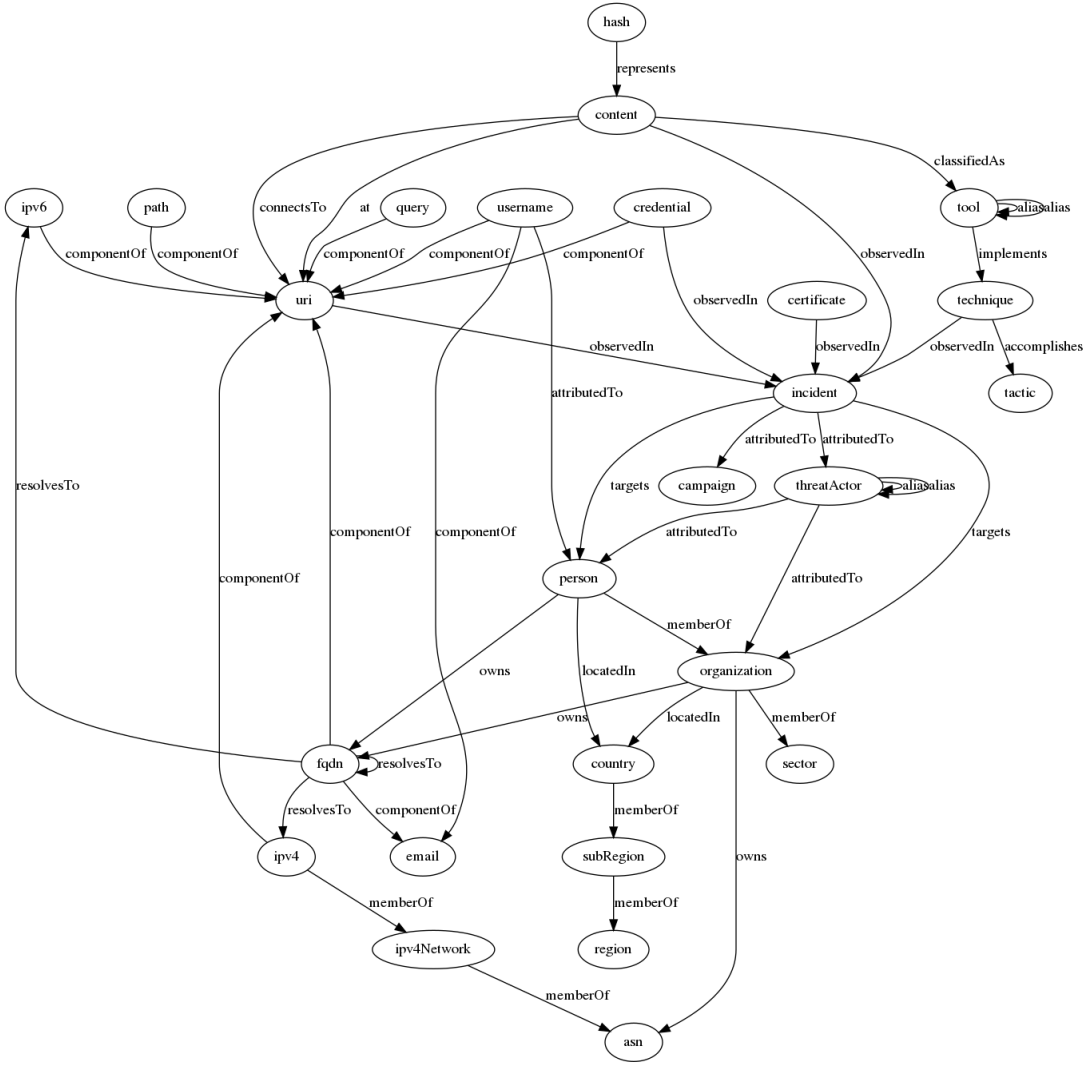
NTNU Norwegian University of Science and Technology

telenor

mnemonic

Data Model

- 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



Models, Taxonomies and Vocabularies

- MITRE ATT&CK
  - https://attack.mitre.org
- MITRE PRE-ATT&CK
  - https://attack.mitre.org/pre-attack/
- MISP galaxy
  - https://github.com/MISP/misp-galaxy
- 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

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

mnemonic

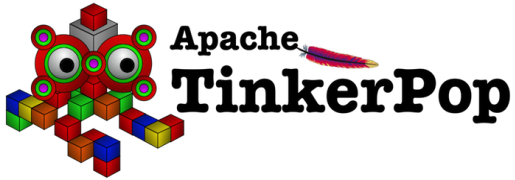
Current OSINT Sources

- APTNotes
  - https://github.com/aptnotes/data
- APT & CyberCriminal Campaign Collection
  - https://github.com/CyberMonitor/APT_CyberCriminal_Campagin_Collections
- RSS Feeds
  - Infosec blogs
- mnemonic PassiveDNS
  - https://passivedns.mnemonic.no/
- Shadowserver IP-BGP
  - https://www.shadowserver.org/wiki/pmwiki.php/Services/IP-BGP
- VirusTotal
- MISP (circl.lu)

THE ACT PLATFORM

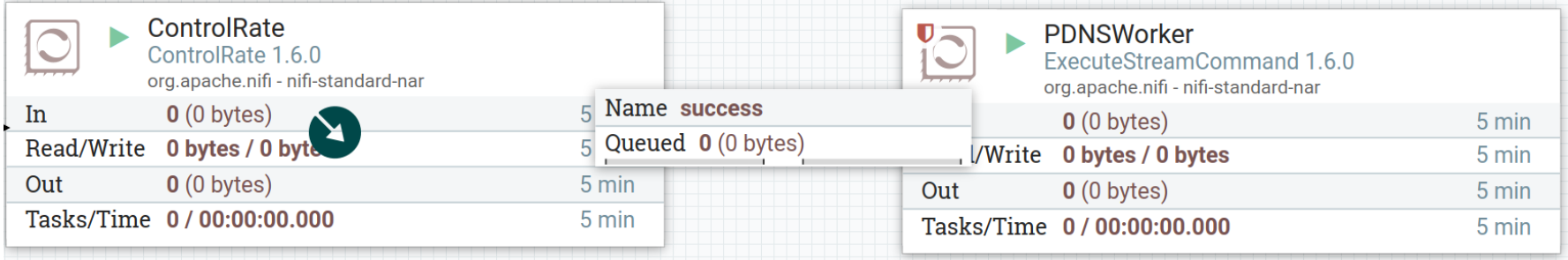
mnemonic

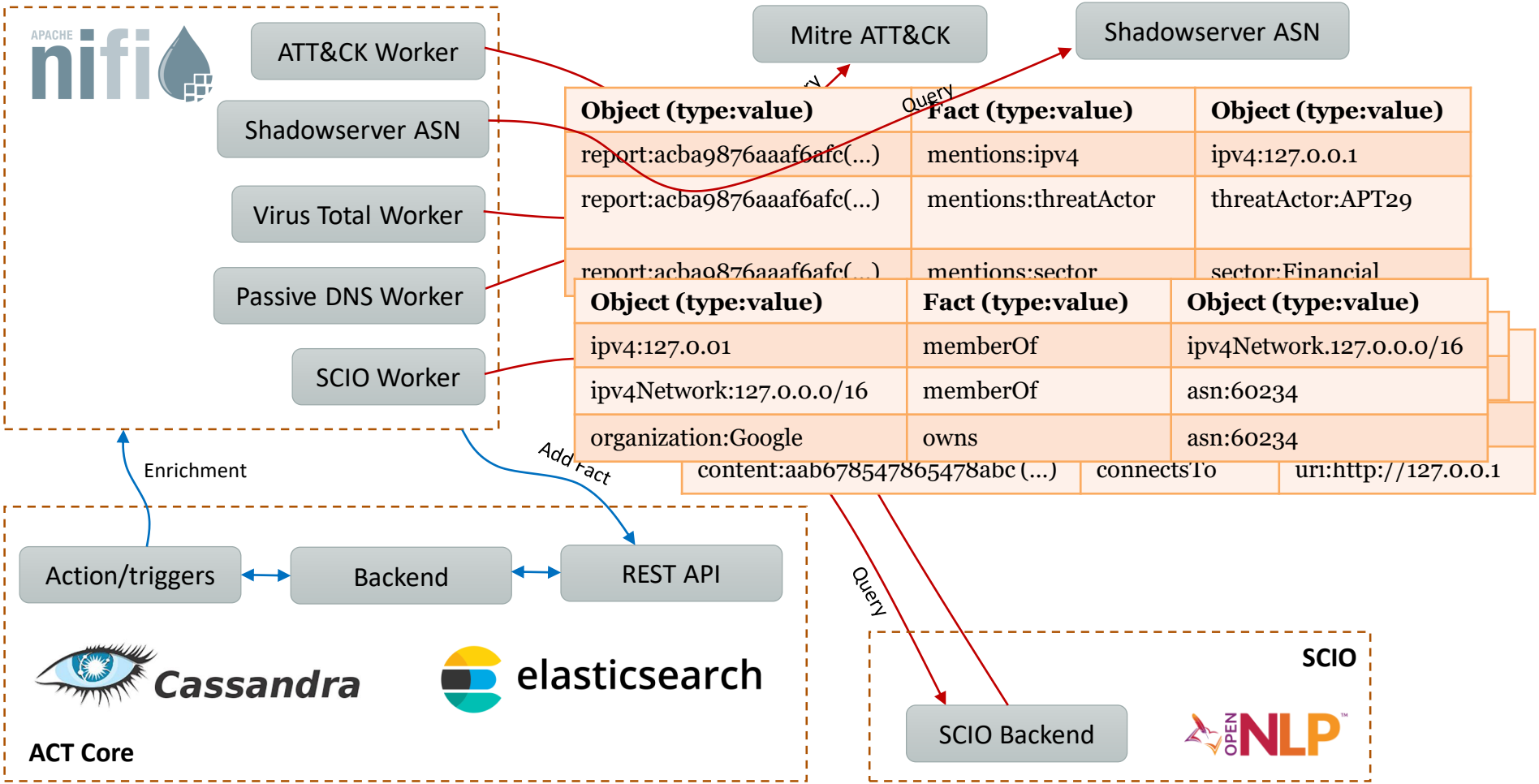
Platform Architecture – Core technologies



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



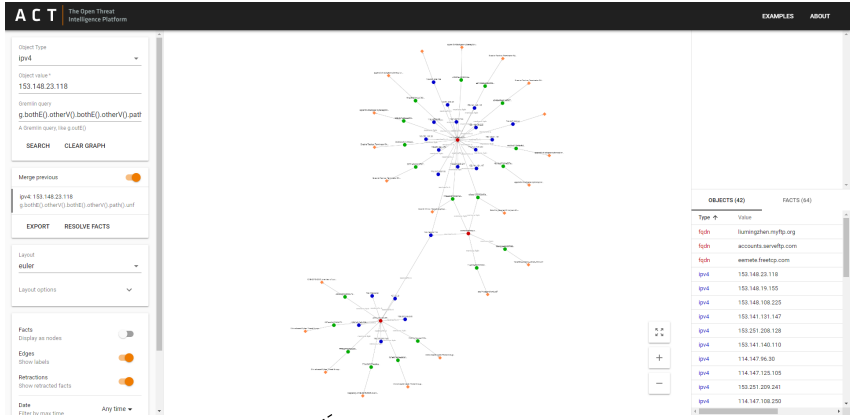

APACHE nifi
ATT&CK Worker
Shadowserver ASN
Virus Total Worker
Passive DNS Worker
SCIO Worker
Mitre ATT&CK
Shadowserver ASN
Query
Object (type:value)
Fact (type:value)
Object (type:value)
report:acba9876aaaf6afc(...)
mentions:ipv4
ipv4:127.0.0.1
report:acba9876aaaf6afc(...)
mentions:threatActor
threatActor:APT29
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.

mnemonic

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

splunk>enterprise App: Search & Reporting ▾ Administrator ▾ 2 Messages ▾ Set

Search Datasets Reports Alerts Dashboards

New Search

```
| act apt29
```

✓ 25 results (10/21/18 1:00:00.000 PM to 10/22/18 1:51:39.000 PM) No Event Sampling ▾ Job

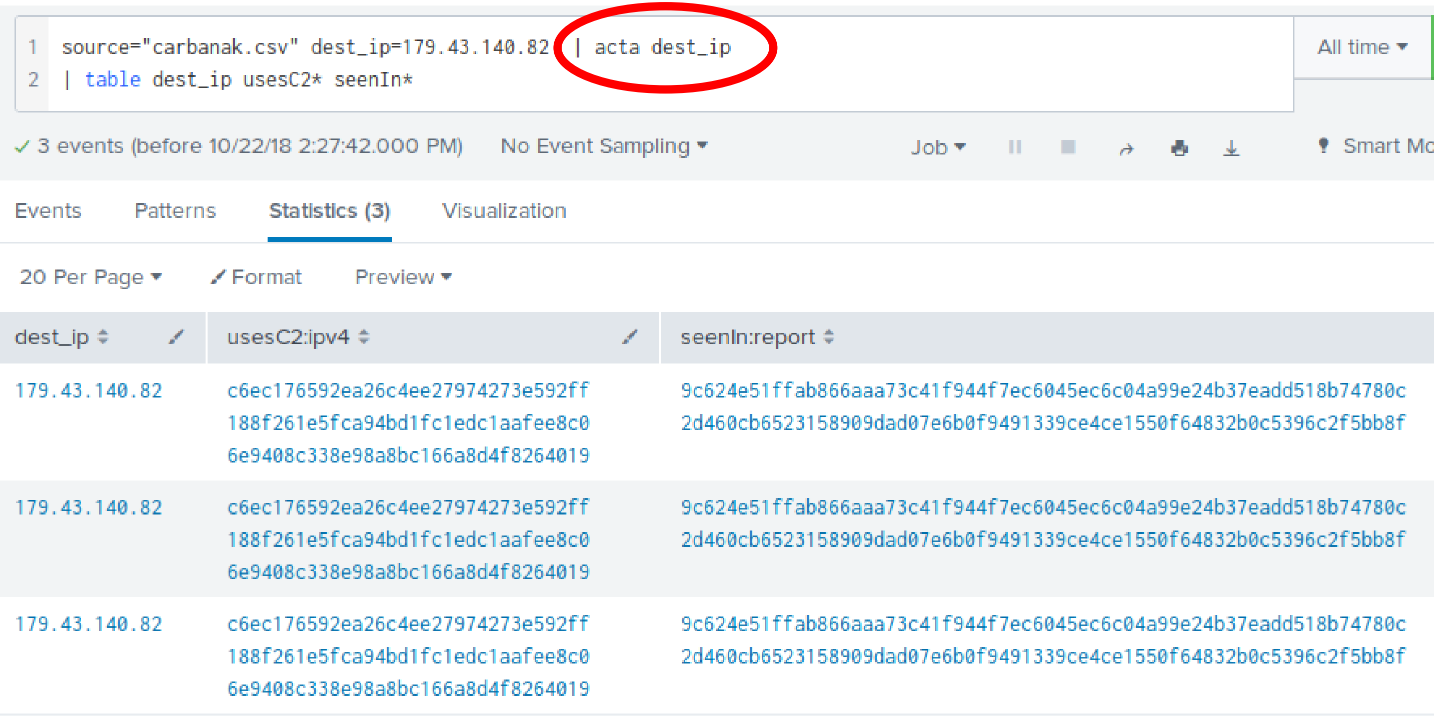
Events (0) Patterns **Statistics (25)** Visualization

20 Per Page ▾ Format Preview ▾

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

mnemonic

Splunk Add-on – Annotate search results



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

✓ 3 events (before 10/22/18 2:27:42.000 PM)

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

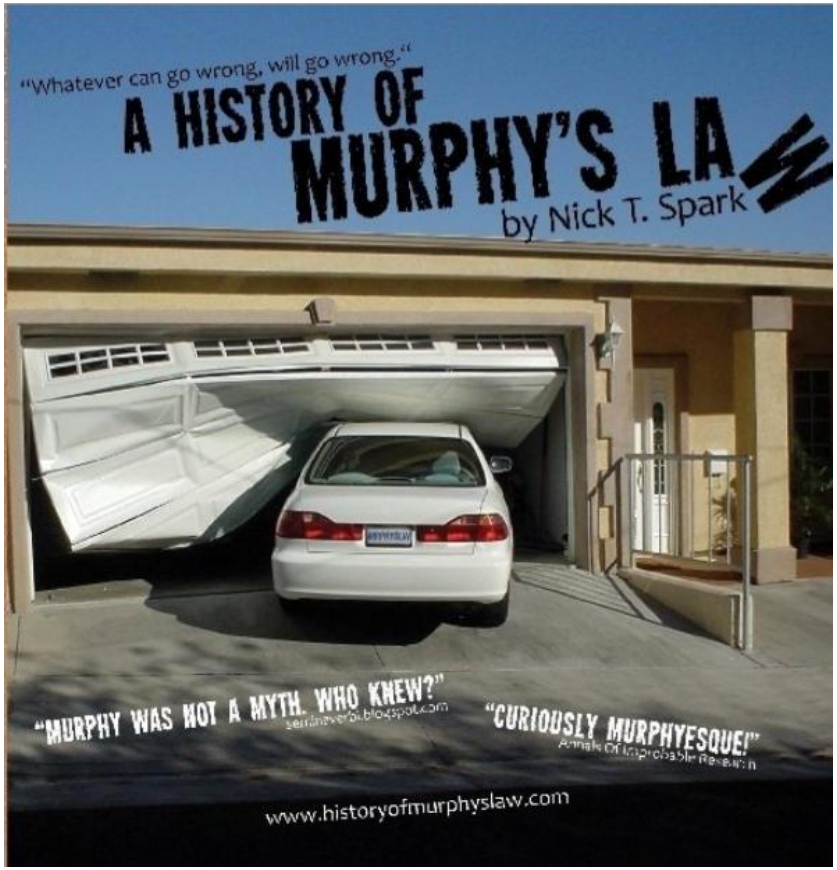
- Implemented
  - Core platform
  - API
  - GUI
  - Workflow orchestration
  - Graph queries

- Github repositories
  - 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)

WORKSHOP - INTRODUCTION

mnemonic

Before We Start



Accessing the GUI

- Read-only
- https://act-eu1.mnemonic.no
- https://act-eu2.mnemonic.no

- Tasks: /examples/
- API: /swagger/
- API-assignments: https://github.com/mnemonic-no/act-workshop-api (jupyter notebook you can test yourself if you are interested in testing the python API)

Introduction 1

mnemonic

Introduction 1 – Right Click / Left Click

mnemonic

Introduction 1 – History, Layouts and Filtering

ACT | The Open Threat Intelligence Platform

EXAMPLES ABOUT

Object Type
hash

Object value *
d06432486e7e9c2b8aaef4f42c11cf8e

Gremlin query
A Gremlin query, like g.outE()

SEARCH CLEAR GRAPH

Merge previous

hash:
d06432486e7e9c2b8aaef4f42c11cf8efe1968

content:
d06432486e7e9c2b8aaef4f42c11cf8efe1968

tool: trickbot

EXPORT RESOLVE FACTS

Layout
euler

Layout options

Facts
Display as nodes

Edges
Show labels

trickbot
tool
24 facts
alias: 2
classifiedAs: 4
implements: 18

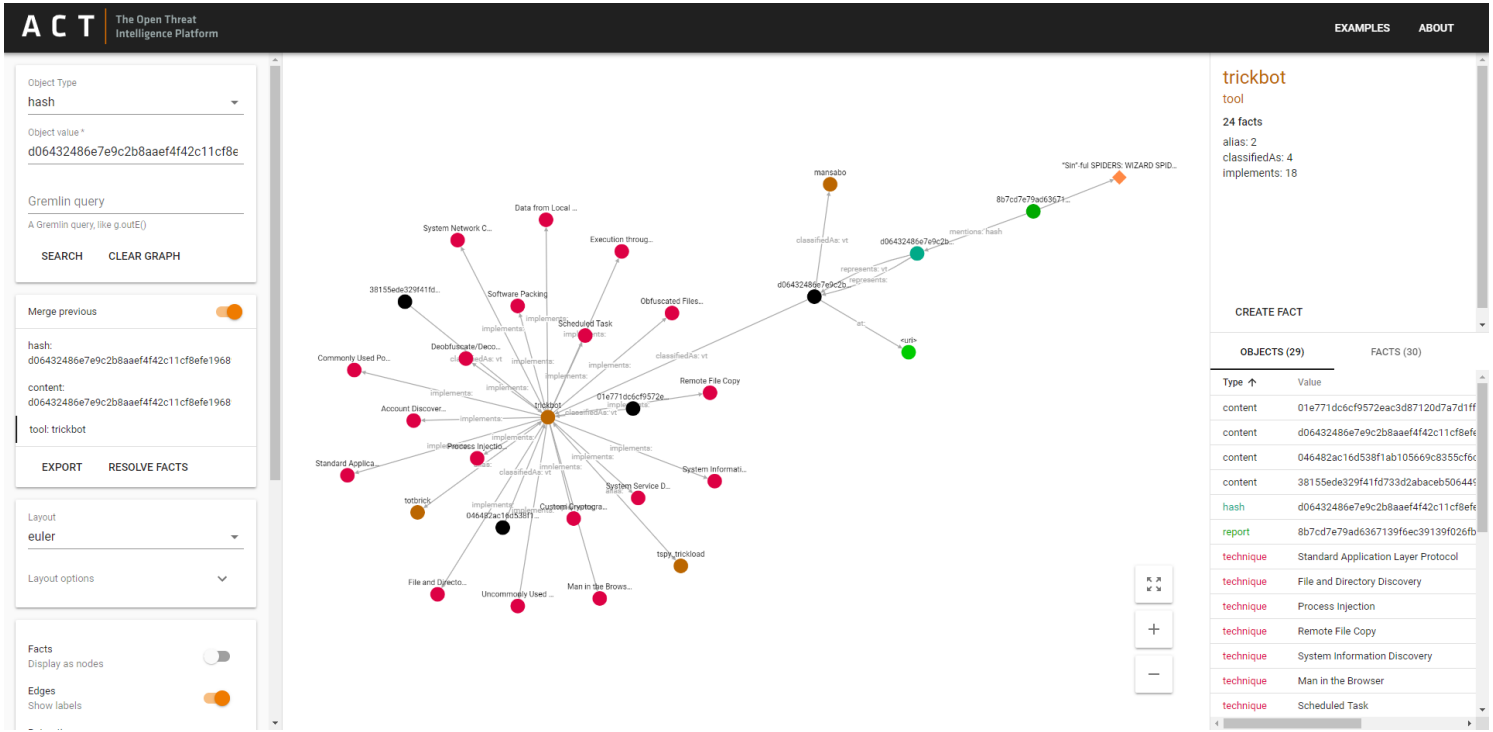
CREATE FACT

OBJECTS (29) FACTS (30)

Type	Value
content	01e771dc6cf9572eac3d87120d7a7d1ff
content	d06432486e7e9c2b8aaef4f42c11cf8efe
content	046482ac16d538f1ab105669c8355cf6c
content	38155ede329f41fd733d2abaceb506449
hash	d06432486e7e9c2b8aaef4f42c11cf8efe
report	8b7cd7e79ad6367139f6ec39139f026fb
technique	Standard Application Layer Protocol
technique	File and Directory Discovery
technique	Process Injection
technique	Remote File Copy
technique	System Information Discovery
technique	Man in the Browser
technique	Scheduled Task

mnemonic

Introduction 1 – Fact Types



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: Find the Report

https://www.crowdstrike.com/blog/sin-ful-spiders-wizard-spider-and-lunar-spider-sharing-the-same-web/

Task 3

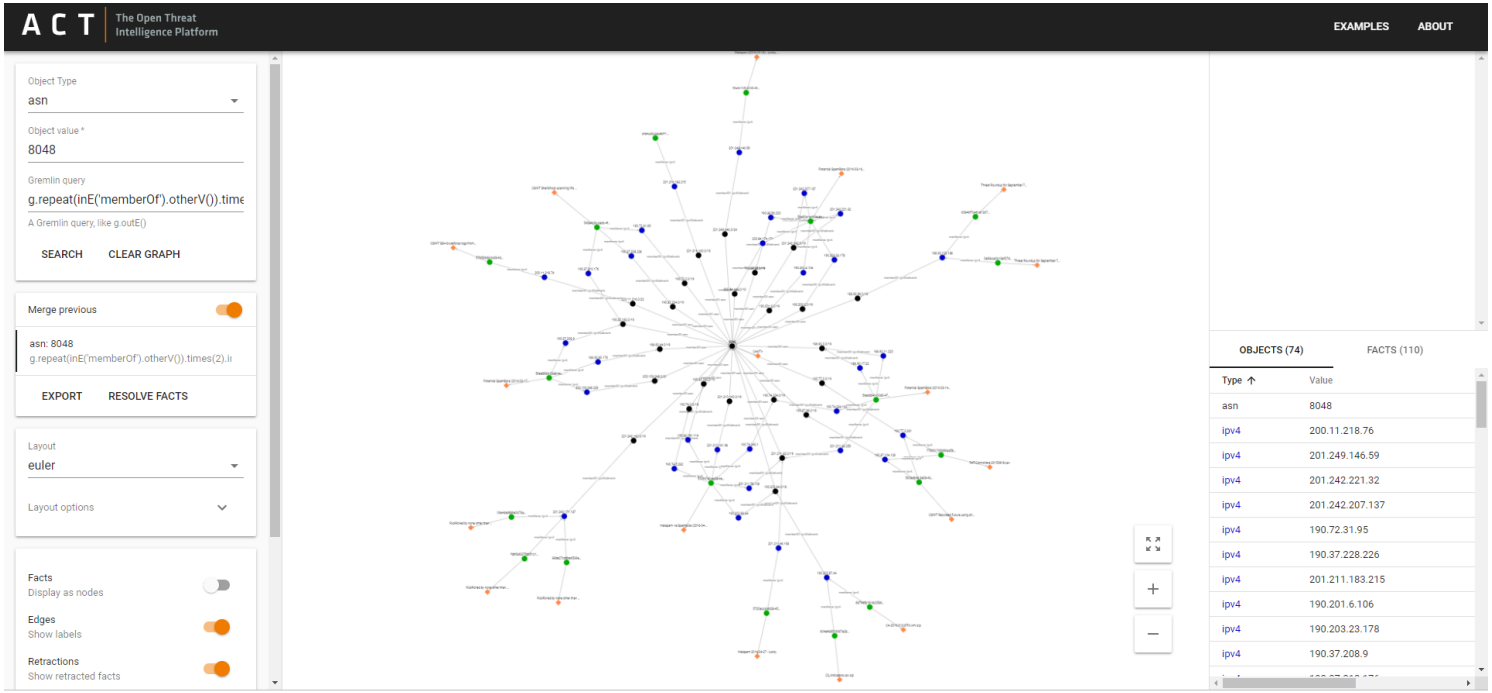
Explore Autonomous System Number 8048

- asn: 8048

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

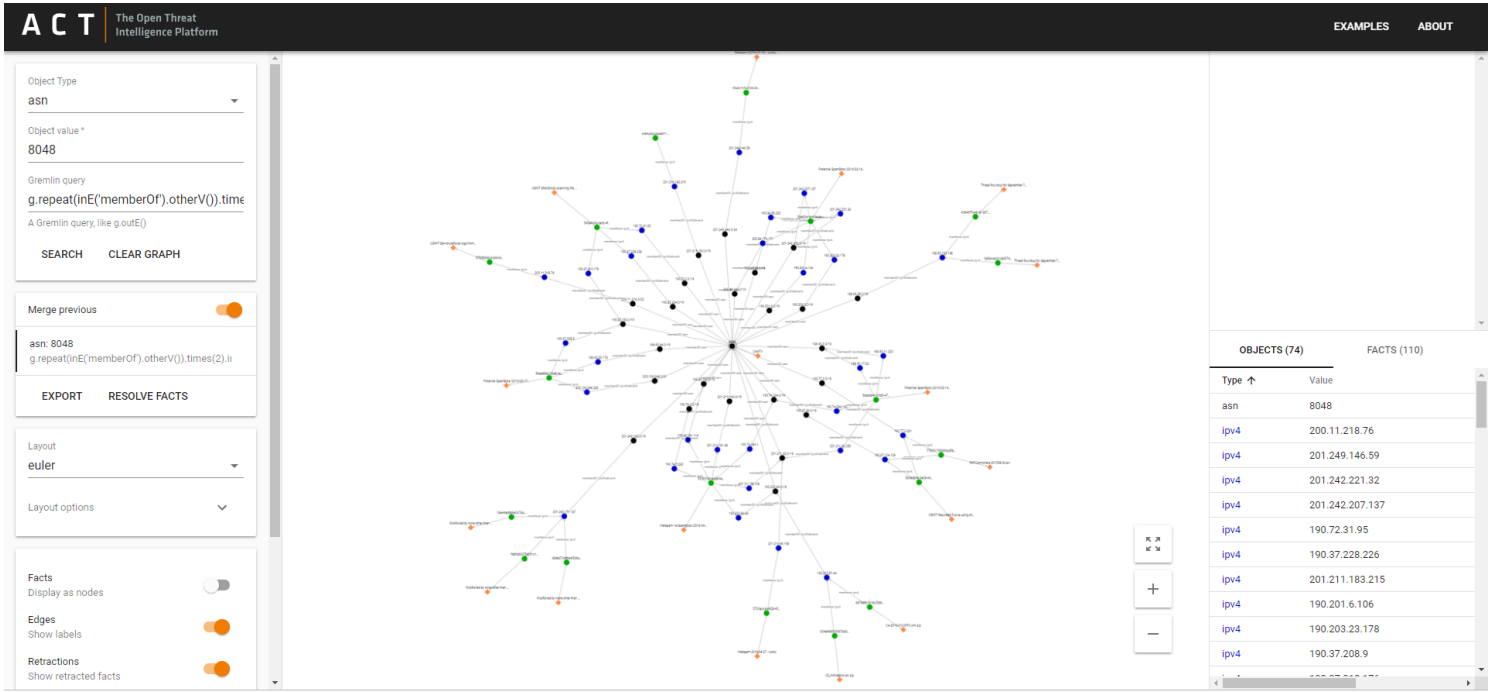
Where is the organization that owns AS8048 located?

Introduction 3 – Graph Query



mnemonic

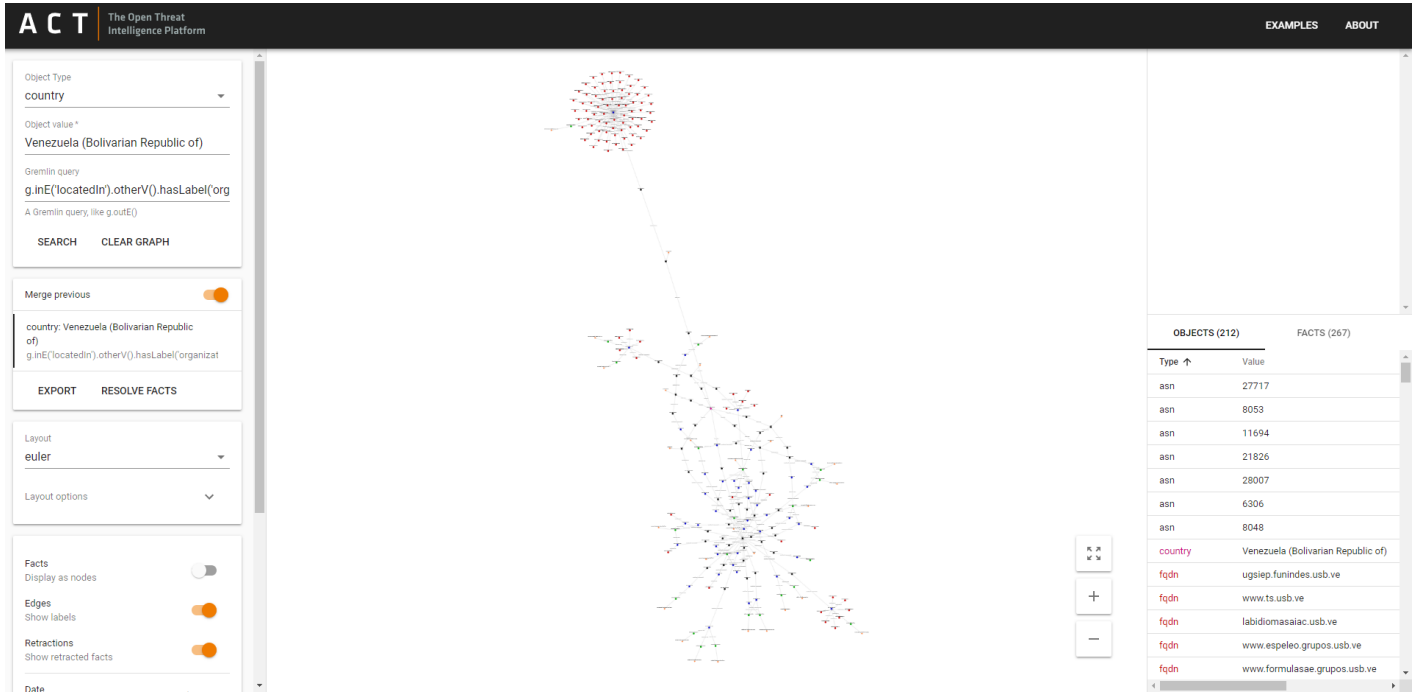
Introduction 3 – Graph Query



Try to replace ‘mentions’ with ‘resolvesTo’ in the graph query (you can edit the URL).

mnemonic

Introduction 4 – Extended Graph Query



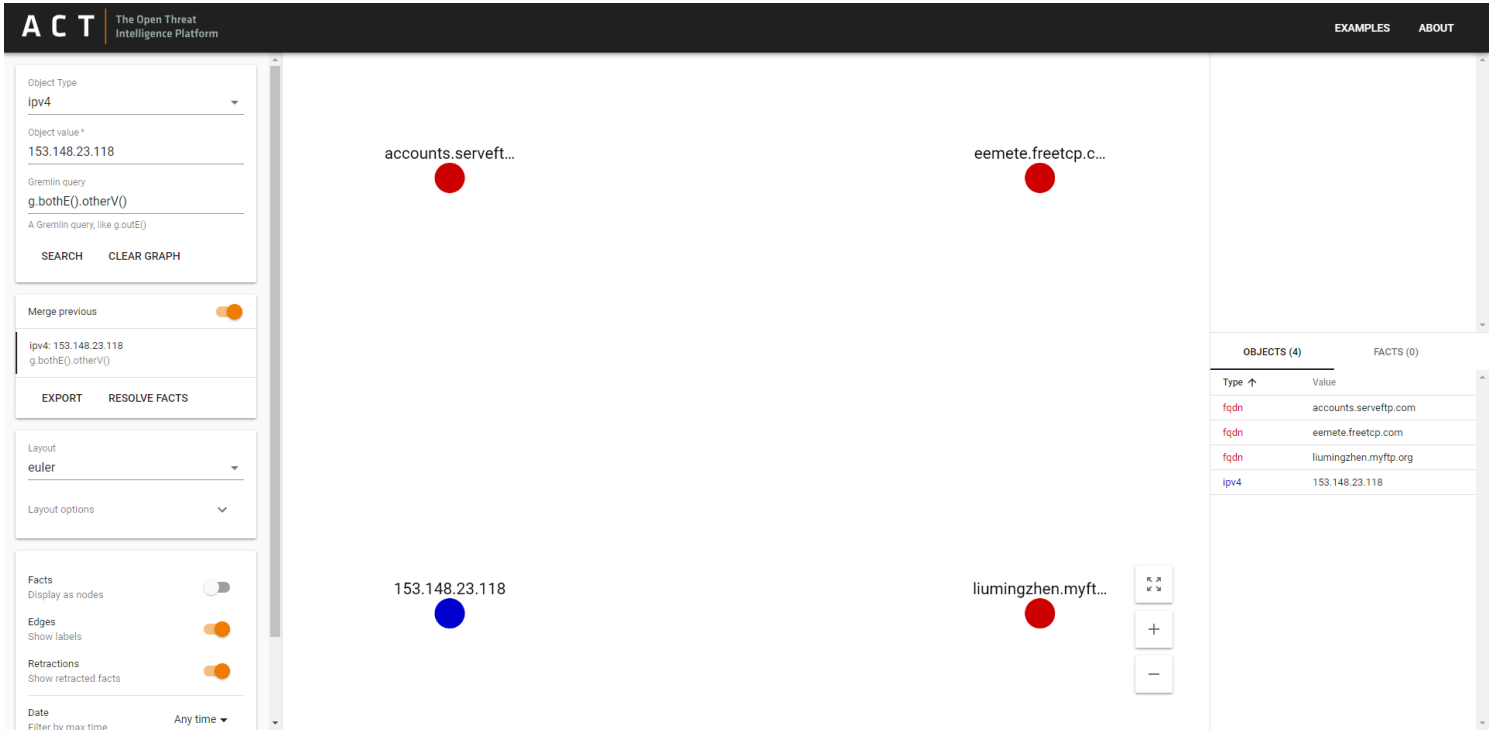
mnemonic

WORKSHOP – GRAPH QUERIES

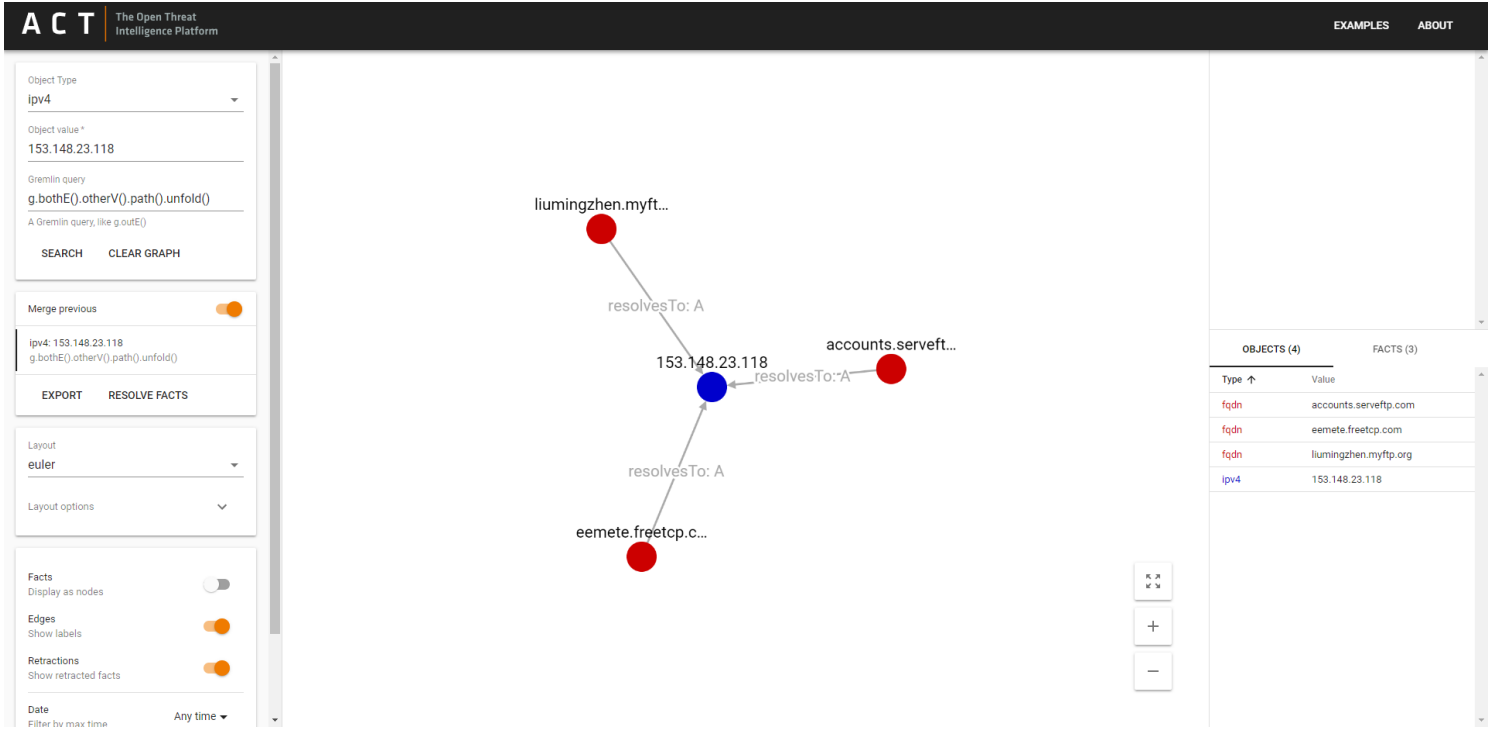
With Great Power Comes Great Responsibility

mnemonic

Graph Query 1

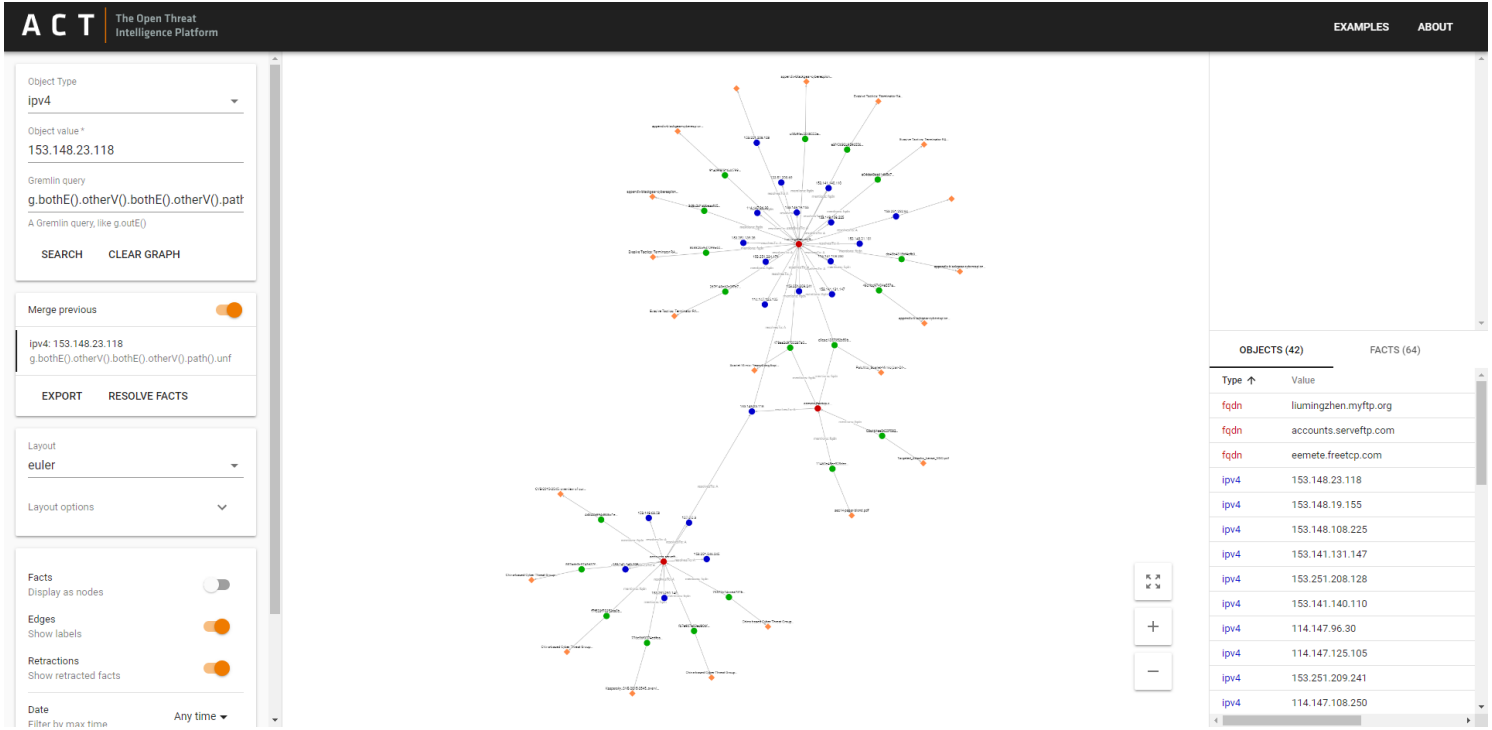


Graph Query 2 – Show Edges



mnemonic

Graph Query 3 – 2 hops



mnemonic

Graph Query 4 – Filter Edges (Facts)

ACT The Open Threat Intelligence Platform

EXAMPLES ABOUT

Object Type
ipv4

Object value *
153.148.23.118

Gremlin query
g.bothE().otherV().bothE('resolvesTo').o

A Gremlin query, like g.outE()

SEARCH CLEAR GRAPH

Merge previous

ipv4: 153.148.23.118
g.bothE().otherV().bothE('resolvesTo').otherV(

EXPORT RESOLVE FACTS

Layout
euler

Layout options

Facts
Display as nodes

Edges
Show labels

Retractions
Show retracted facts

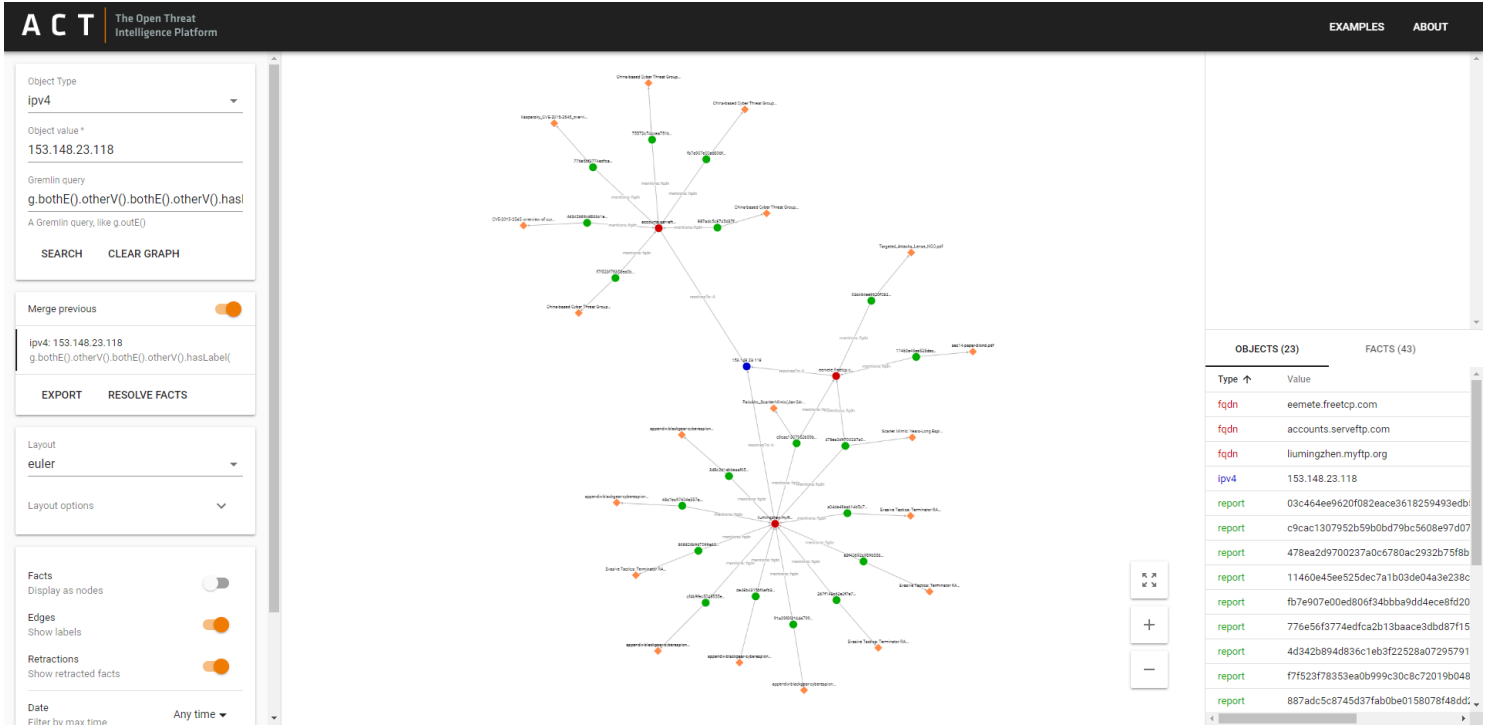
Date
Filter by max time
Any time

OBJECTS (23) FACTS (24)

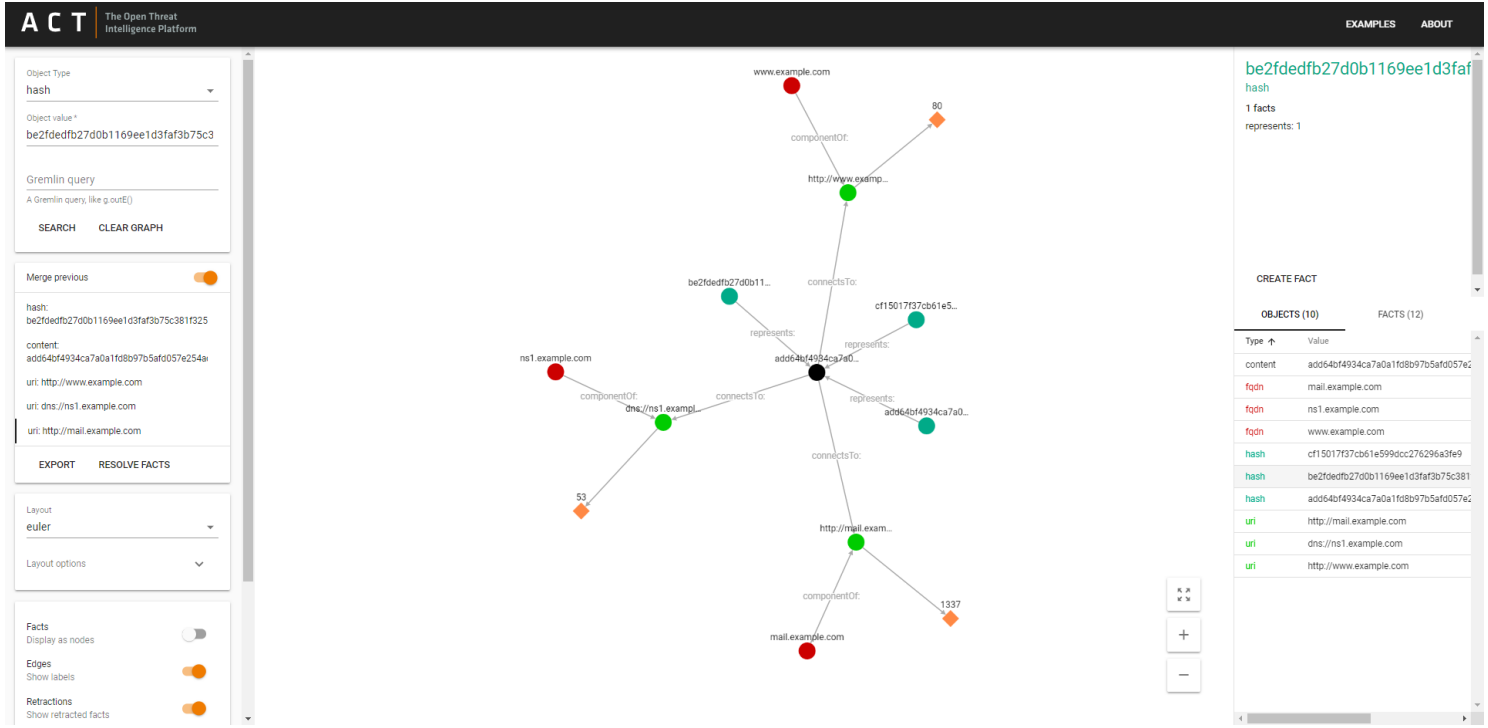
Type	Value
fqdn	accounts.serveftp.com
fqdn	eemete.freetcp.com
fqdn	liumingzhen.myftp.org
ipv4	153.148.23.118
ipv4	153.251.246.245
ipv4	153.251.250.140
ipv4	153.148.63.58
ipv4	127.0.0.3
ipv4	153.141.140.208
ipv4	153.141.131.147
ipv4	123.51.208.69
ipv4	153.148.108.225
ipv4	153.251.252.64

mnemonic

Graph Query 5 – Filter Nodes (Objects)

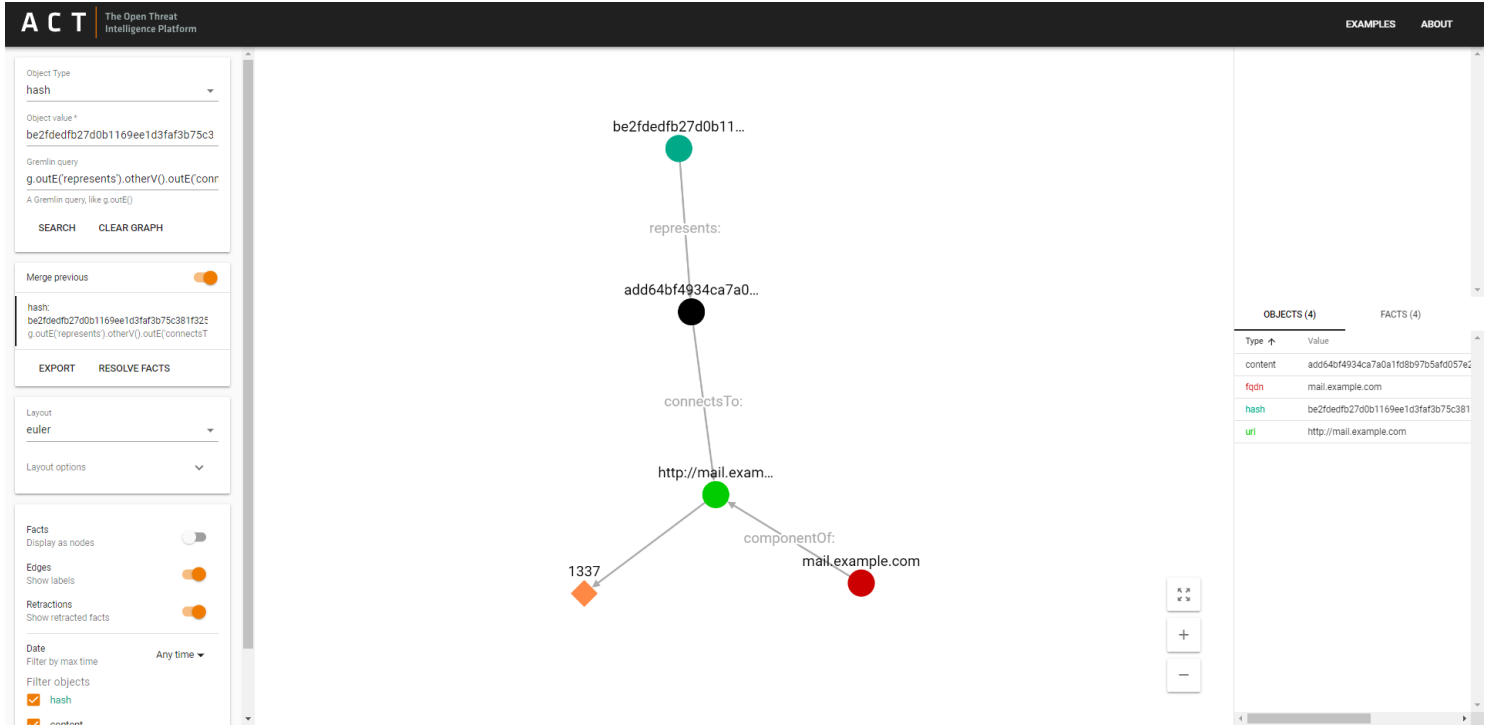


Task 4 - Subgraph



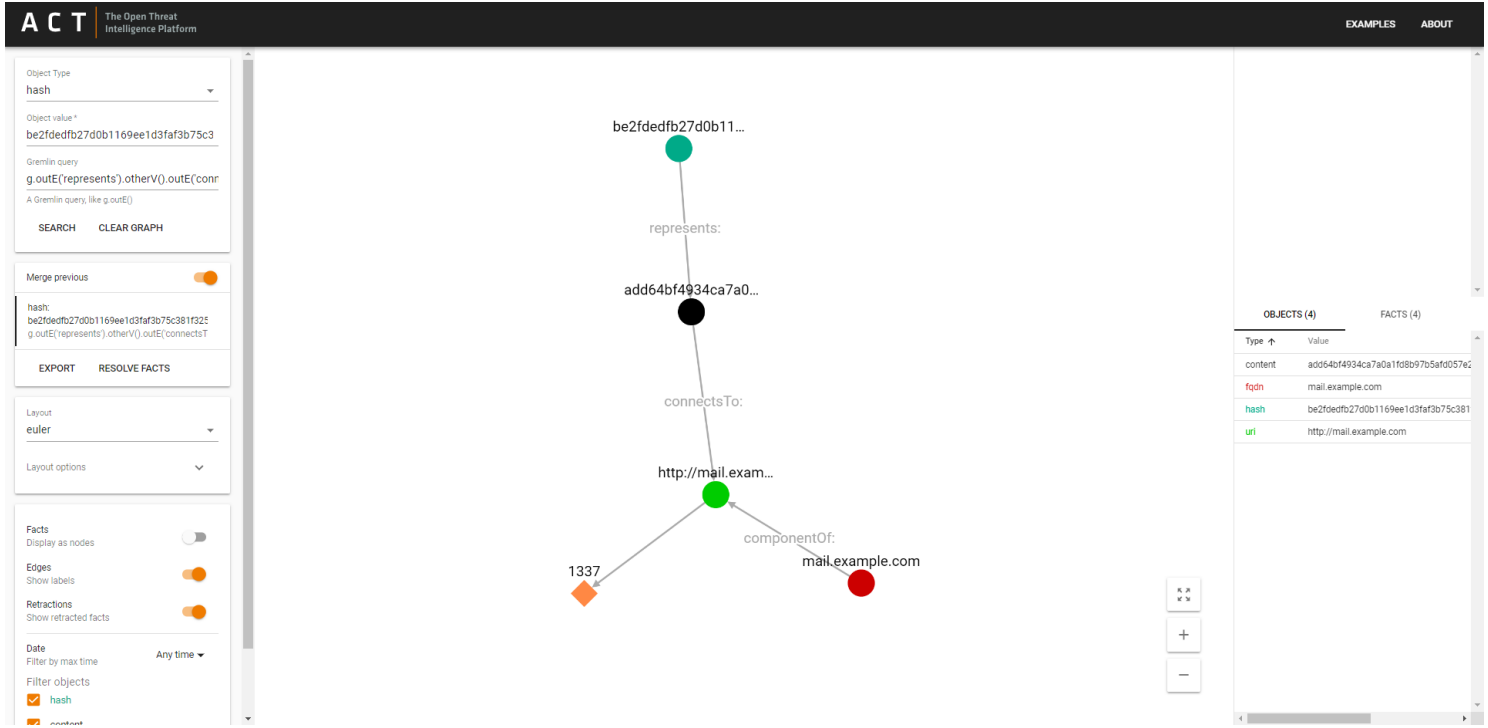
mnemonic

hash → content → uri with port number 1337 ← fqdn



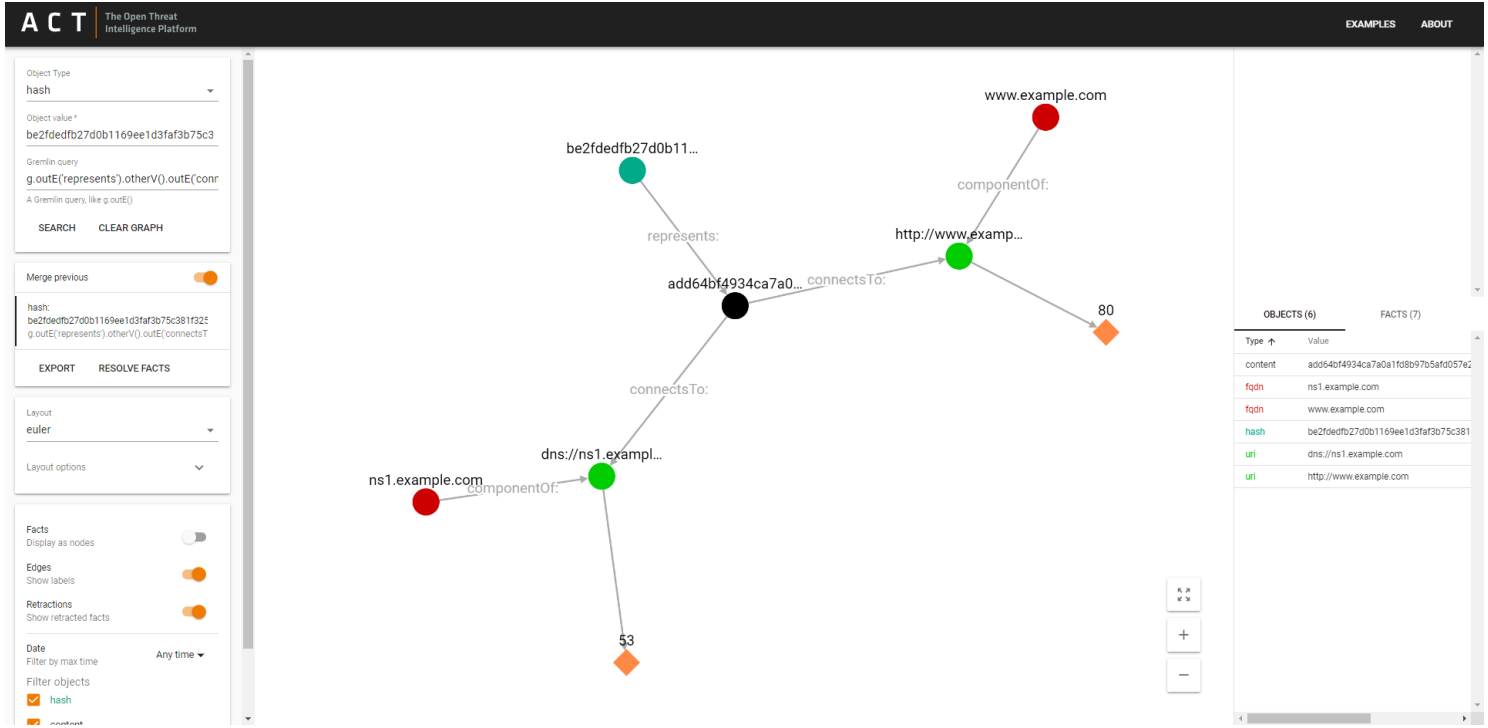
mnemonic

g.outE('represents').otherV().outE('connectsTo').otherV().where(outE().has('value','1337')).inE('componentOf').otherV().path().unfold()



mnemonic

g.outE('represents').otherV().outE('connectsTo').otherV().not(where(outE().has('value','1337'))).inE('componentOf').otherV().path().unfold()



not(where()), not where(not())

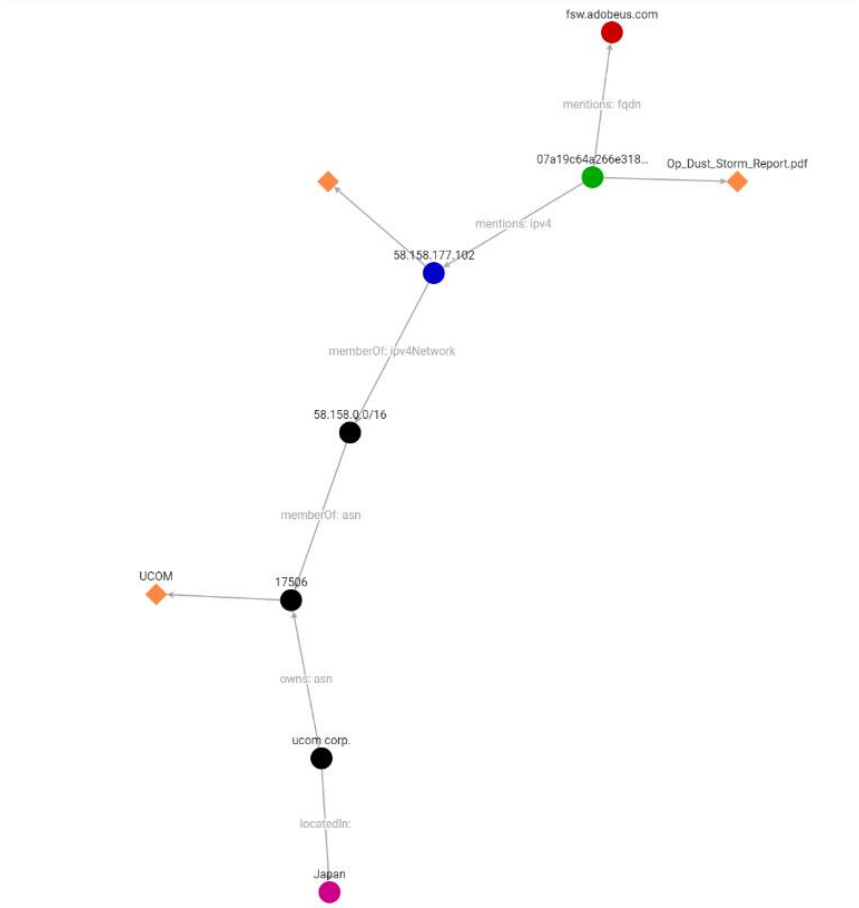
mnemonic

Task 5: Find the IP Address Owner

The fqdn fsw.adobeus[.]com is mentioned in one report. A sinkhole IPv4 address is also mentioned in the same report. Which organization owns that sinkhole IPv4 address, and which country is it located in?

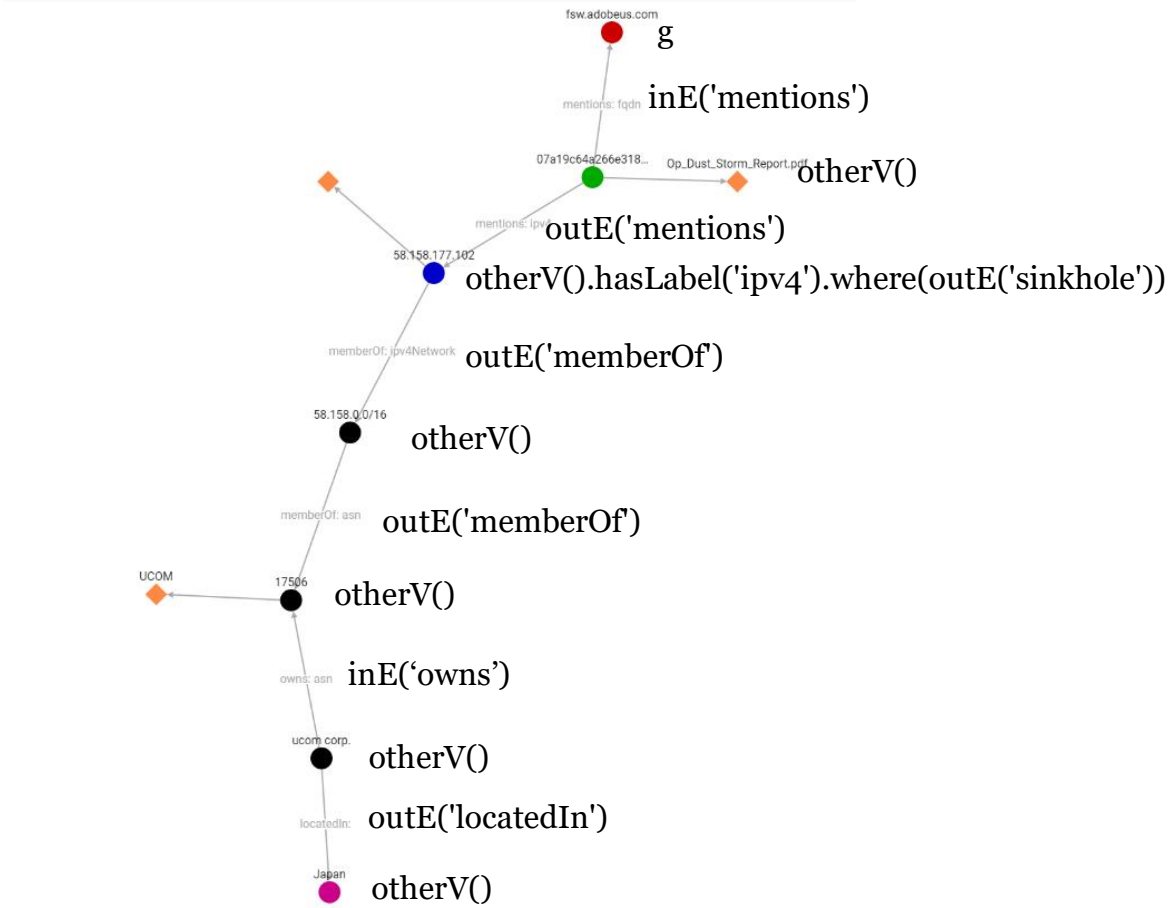
Hint: Fact Type ‘mentions’ and ‘memberOf’

Task 5 Solution



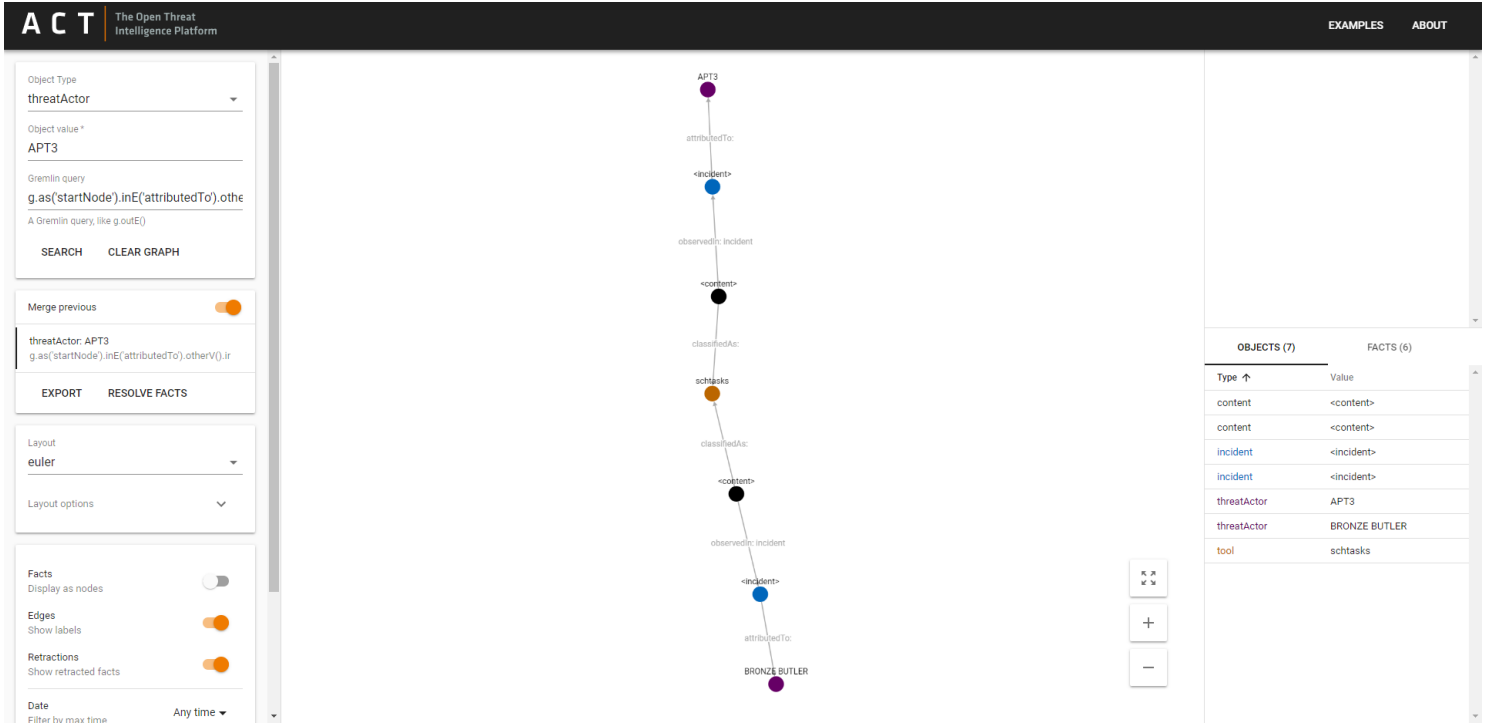
mnemonic

g.inE('mentions').otherV().outE('mentions').otherV().hasLabel('ipv4').where(outE('sinkhole')).outE('memberOf').otherV().outE('memberOf').otherV().inE('owns').otherV().outE('locatedIn').otherV().path().unfold()



mnemonic

Graph Query 6 – Unique Tool Usage



mnemonic

ASSIGNMENTS

mnemonic

CASE STUDY

mnemonic

Public Read-Only ACT Instance

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

FURTHER WORK

mnemonic

New Information Sources

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

Graph Analytics

- Post. doc. @ UiO
- Post. doc. @ NTNU

Information Sharing

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

Trust and Confidence

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

GUI Improvements

- Context menu
  - Pre-defined graph queries
  - Download report
  - …
- Timelines
- Share workspace
- Prune graph