2024
FIRST
Cyber Threat
Intelligence
Conference

Berlin, Germany
April 15-17, 2024

Solving CTI Sector Incoherence in a Living Growing OpenCTI Repository – Extend STIX 2.1 FTW

Philippe Lin & Morton Swimmer
Trend Micro Research, 16 April 2024

We know it's hard to acquire intelligence

Proofpoint / Chinese APT TA413 resumes targeting Tibet following COVID 19 themed economic …

This campaign targeted European **diplomatic** and **legislative** bodies, **non-profit** **policy research organizations**, and global **organizations dealing with economic affairs**.

[…] as part of the TA413 March 2020 campaign targeting European **economic entities**. […] The re-emergence of well-known Tibetan themed sender addresses and graphically didactic PowerPoint attachments in later July again tie TA413 to its emblematic targeting of the **Tibetan community**.

The report indicated this email address had been used in phishing campaigns from May 2012 through June 2013 to target Tibetan members of **civil society**.

It's even harder across vendors

Talos / Breaking the silence recent Truebot activity …

While we don't have enough information to say that there is a specific focus on a sector, we noticed a number of compromised **education** sector organizations.

Mandiant / Stomp 2 Dis Brilliance in the Visual Basics …

The campaigns primarily targeted **financial** services organizations in the United States, [...] At least one campaign targeted South Korean organizations, including a **marketing agency**.

Proofpoint / Servhelper and Flawedgrace New Malware Introduced TA505 …

TA505 appears to be actively targeting **banks**, **retail** businesses, and **restaurants** as they distribute these malware families.

Trend Micro / Shifting tactics breaking down TA505 groups use of HTML RATs and …

They had also targeted **retail** brands and even different **financial** companies across the world.

And when a group attacks too much …

Mandiant / Game-over detecting and stopping an APT41 operation

A China-nexus dual espionage and financially-focused group, APT41 targets industries such as **gaming**, **healthcare**, **high-tech**, higher **education**, **telecommunications**, and **travel** services.

Unit42/ Travel themed phising

This article provides early warnings for the **travel** industry and global **travelers** […]

Securelist / Moonbounce the dark side of UEFI firmware

One particular target corresponds to an organization **in control of several enterprises** dealing with **transport** technology.

Unit42/ APT41 using new Speculoos backdoor to target organizations globally

We also used this data to identify multiple victims in industries such as **healthcare**, higher **education**, **manufacturing**, **government** and **technology** services in multiple regions around the world, such as North America, South America, and Europe.

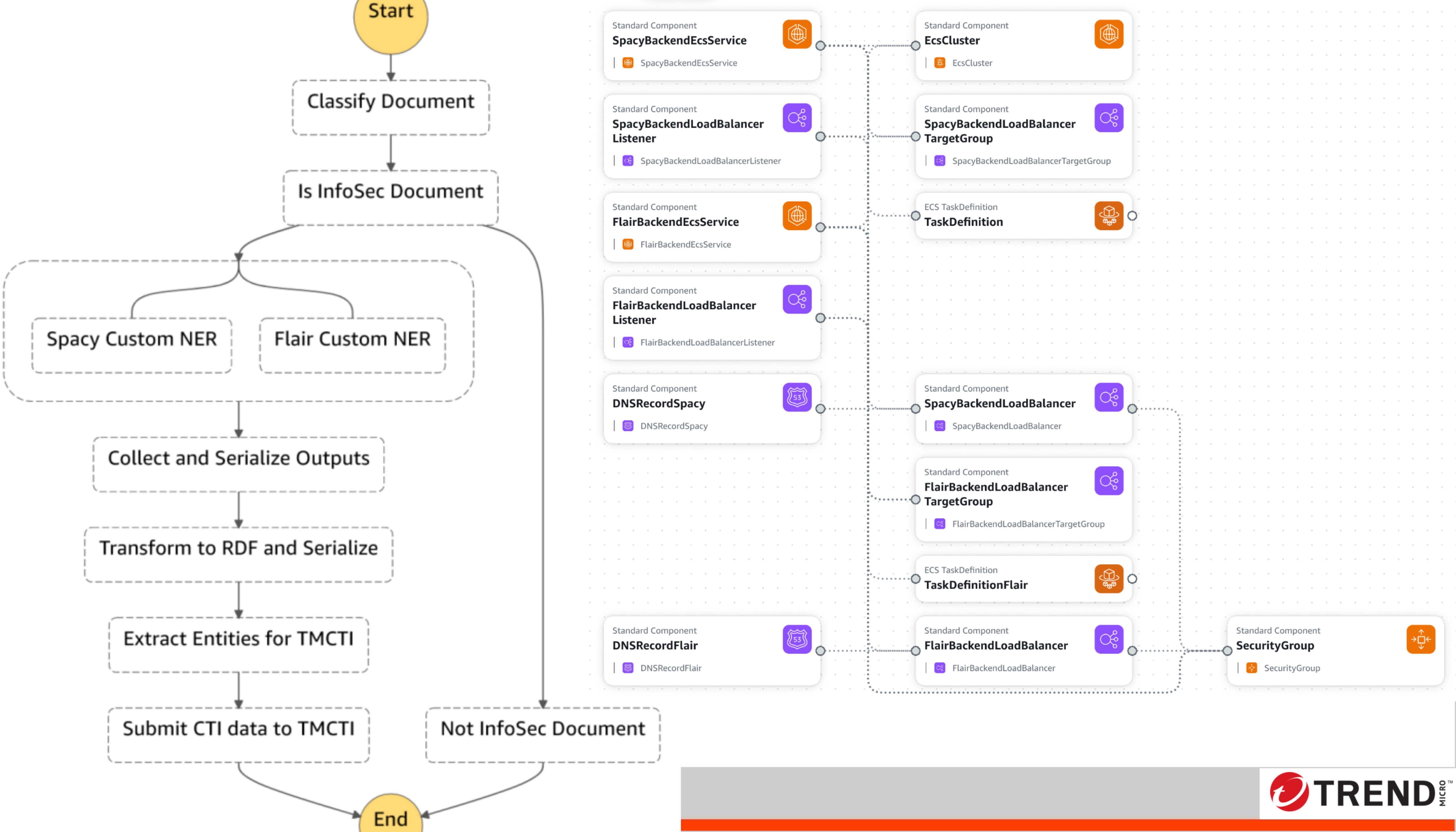
What happens over 3 years of adding data to OpenCTI

- Ingesting free-form text & structured data + manual edits
- 50k+ external references
- From 2k source domain names
- Of 46 well-known sources
- By 50+ threat researchers who enrich the data and create hundreds of threat reports

- System was based on ANSSI sectors, countries, regions, cities …

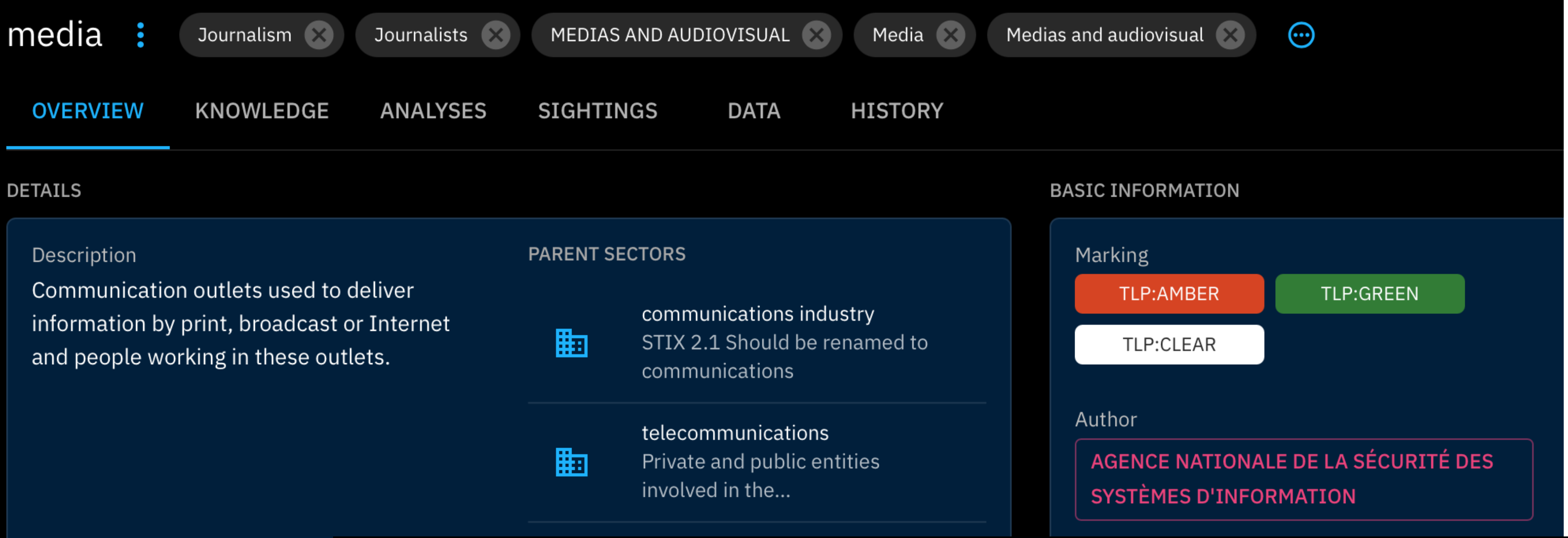
In addition to high quality reports from

- Cisco Talos
- Mandiant / Fireeye
- Kaspersky Securelist
- PAN / Unit42
- ESET
- Proofpoint


Start
Classify Document
Is InfoSec Document
Spacy Custom NER
Flair Custom NER
Collect and Serialize Outputs
Transform to RDF and Serialize
Extract Entities for TMCTI
Submit CTI data to TMCTI
Not InfoSec Document
End
Standard Component
SpacyBackendEcsService
SpacyBackendEcsService
Standard Component
EcsCluster
EcsCluster
Standard Component
SpacyBackendLoadBalancer Listener
SpacyBackendLoadBalancerListener
Standard Component
SpacyBackendLoadBalancer TargetGroup
SpacyBackendLoadBalancerTargetGroup
Standard Component
FlairBackendEcsService
FlairBackendEcsService
ECS TaskDefinition
TaskDefinition
Standard Component
FlairBackendLoadBalancer Listener
FlairBackendLoadBalancerListener
Standard Component
DNSRecordSpacy
DNSRecordSpacy
Standard Component
SpacyBackendLoadBalancer
SpacyBackendLoadBalancer
Standard Component
FlairBackendLoadBalancer TargetGroup
FlairBackendLoadBalancerTargetGroup
ECS TaskDefinition
TaskDefinitionFlair
Standard Component
DNSRecordFlair
DNSRecordFlair
Standard Component
FlairBackendLoadBalancer
FlairBackendLoadBalancer
Standard Component
SecurityGroup
SecurityGroup

But why sectors?

1. Avoid naming customers
2. Abstraction → Aggregation
3. Attacker group has preferences
4. "Are we attacked?"
5. If *(fingers crossed)* everything is right, we have beautiful graphs.



media

Journalism ⊗ Journalists ⊗ MEDIAS AND AUDIOVISUAL ⊗ Media ⊗ Medias and audiovisual ⊗

OVERVIEW KNOWLEDGE ANALYSES SIGHTINGS DATA HISTORY

DETAILS

Description
Communication outlets used to deliver information by print, broadcast or Internet and people working in these outlets.

PARENT SECTORS

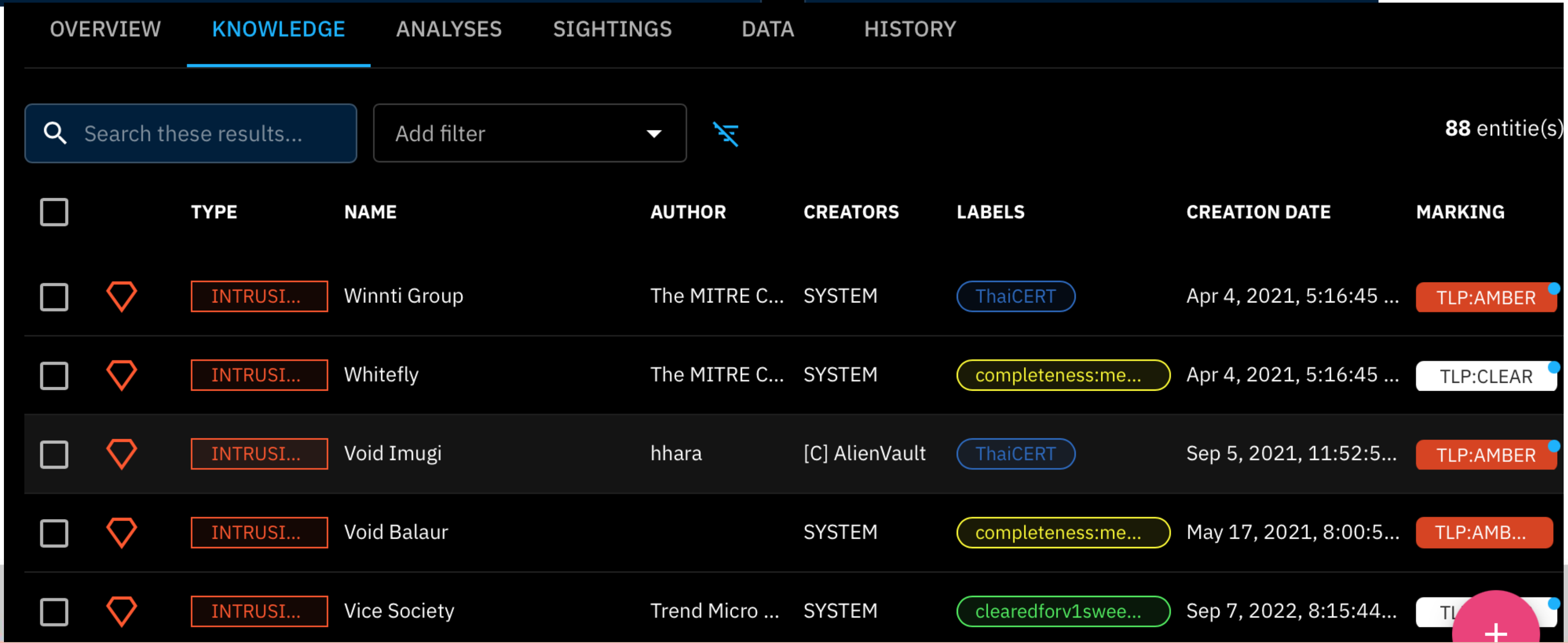
communications industry
STIX 2.1 Should be renamed to communications

telecommunications
Private and public entities involved in the...

BASIC INFORMATION

Marking
TLP:AMBER TLP:GREEN TLP:CLEAR

Author
AGENCE NATIONALE DE LA SÉCURITÉ DES SYSTÈMES D'INFORMATION



OVERVIEW KNOWLEDGE ANALYSES SIGHTINGS DATA HISTORY

Search these results...

Add filter

88 entitie(s)

□		TYPE	NAME	AUTHOR	CREATORS	LABELS	CREATION DATE	MARKING
□		INTRUSI...	Winnti Group	The MITRE C...	SYSTEM	ThaiCERT	Apr 4, 2021, 5:16:45 ...	TLP:AMBER
□		INTRUSI...	Whitefly	The MITRE C...	SYSTEM	completeness:me...	Apr 4, 2021, 5:16:45 ...	TLP:CLEAR
□		INTRUSI...	Void Imugi	hhara	[C] AlienVault	ThaiCERT	Sep 5, 2021, 11:52:5...	TLP:AMBER
□		INTRUSI...	Void Balaur		SYSTEM	completeness:me...	May 17, 2021, 8:00:5...	TLP:AMB...
□		INTRUSI...	Vice Society	Trend Micro ...	SYSTEM	clearedforv1swee...	Sep 7, 2022, 8:15:44...	TL

But the world is not ideal. ☹

Label misuse

Type	Name	# of STIX objects
loc	americas	3
loc	antarctica	0
org	Ministry of Foreign Affairs	0
platform	Android	11
platform	Win	64
???	Critical Industries	3

Duplicated entries

Type	Name	# of STIX objects
sector	**Education**	304
sector	education industry	242
sector	**Energy**	407
sector	energy industry	100
sector	TELECOMMUNICATIONS	616
sector	telecommunications industry	145
Sector	Telco	5
???	UNKNOWN	66

Misspelling, and …

Type	Name	# of STIX objects
sector	Gas And	5
sector	GAS AND OIL	22
sector	GOVERNEMNT	577
sector	Oil And	17
sector	Oil	0
sector	government-**public-services** industry	153

We ended up with

195 sectors on production
214 sectors on staging

Whereas

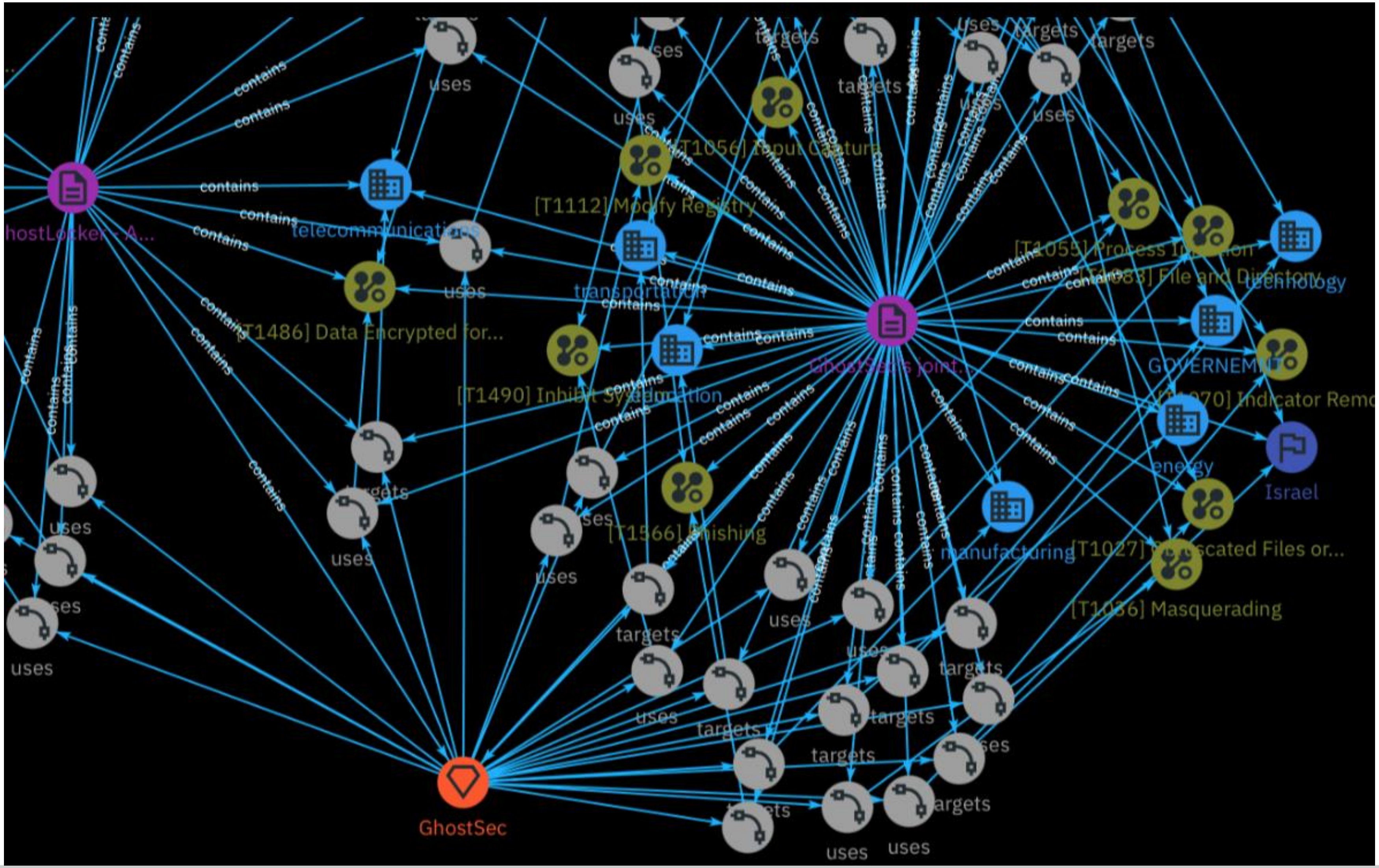
34 sectors in STIX 2.1
16 sectors in US CISA

Must maintain compatibility with other systems >_<"

Sectors used in different sources

Vendor	Not-so-STIX Sectors
Trend Micro	**aerospace**, **agriculture**, **automotive**, aviation, banking, business, **chemical**, communication, computer, **construction**, consumer electronics, corporate, cosmetics, **defense**, device fabrication, diving, **education**, **energy**, energy and water, enterprises, esports, financial, food and beverage, food and logistics, gaming, gems, government agencies, government ministries, governments, **healthcare**, heavy industries, high-technology, hospitality, industrial equipment, information technology, **insurance**, internet-related services, law enforcement, lights and fixtures, logistics, management consulting, **manufacturing**, marine, media, **mining**, movies and music, non-business industries, oil, oil and gas, online gaming, petroleum, **pharmaceutics**, public sector, pyrotechnics, real estate, **retail**, sailing, satellite, science and technology, services, shipping, space, tax and employment services, tech, **technology**, telecom, **telecommunications**, tiles, tourism, **transportation**, vehicle manufacturing, videogame, water treatment plants, web hosting, wireless
Palo Alto Unit42	**aerospace**, **agriculture**, banking, call centers, **construction**, **defense**, **energy**, finance, gaming, **government**, government institutions, **healthcare**, higher education, high-tech, high technology, hosting providers, **manufacturing**, marketing, medical, private sector, public sector, restaurant, semiconductor, services, software, **technology**, technology service, telecom, **telecommunications**
Kaspersky SecureList	activists, **automotive**, banks, business, **chemical**, cosmetics, cryptocurrency, **defense**, educational institutions, embassies, **energy**, financial, fintech, government institutions, high technology, investment capital, large electronics and manufacturing, law enforcement, maritime and ship-building, mass media, military, military contractors, NGO, oil and gas, online gaming, online video game, outsourcing, **pharmaceuticals**, political parties, private equity, private sector, research institutes, satellite operators, technological startup, telecommunication, telecom operators
Cisco Talos	**energy**, engineering, financial, **government**, logistics, **manufacturing**, public / private sector, telecommunication

What we want

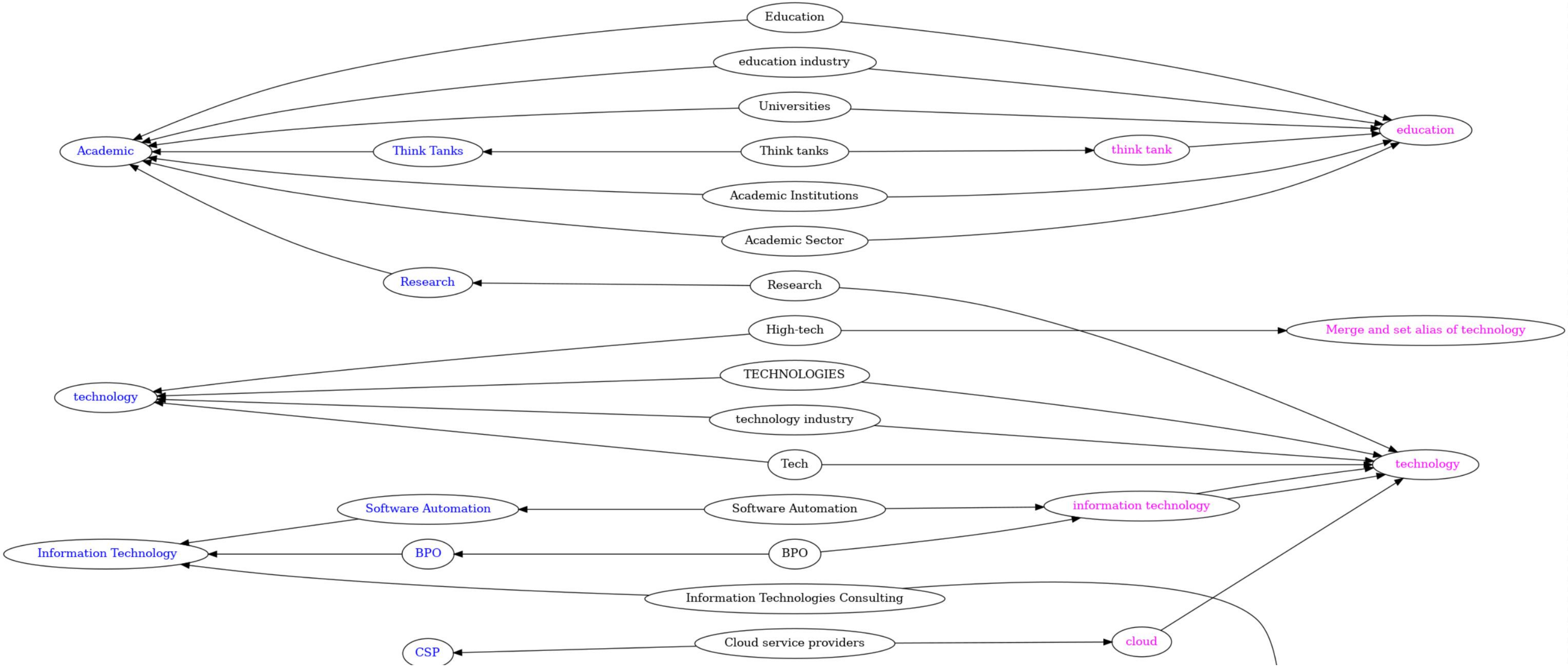

GhostLocker - A...
telecommunications
transportation
[T1056] Input Capture
[T1112] Modify Registry
[T1486] Data Encrypted for...
[T1490] Inhibit System Recovery
[T1566] Phishing
GhostSec's joint...
[T1055] Process Injection
[T1083] File and Directory...
technology
GOVERNEMENT
[T1070] Indicator Remo
energy
Israel
manufacturing
[T1027] Obfuscated Files or...
[T1036] Masquerading
GhostSec
contains
uses
targets

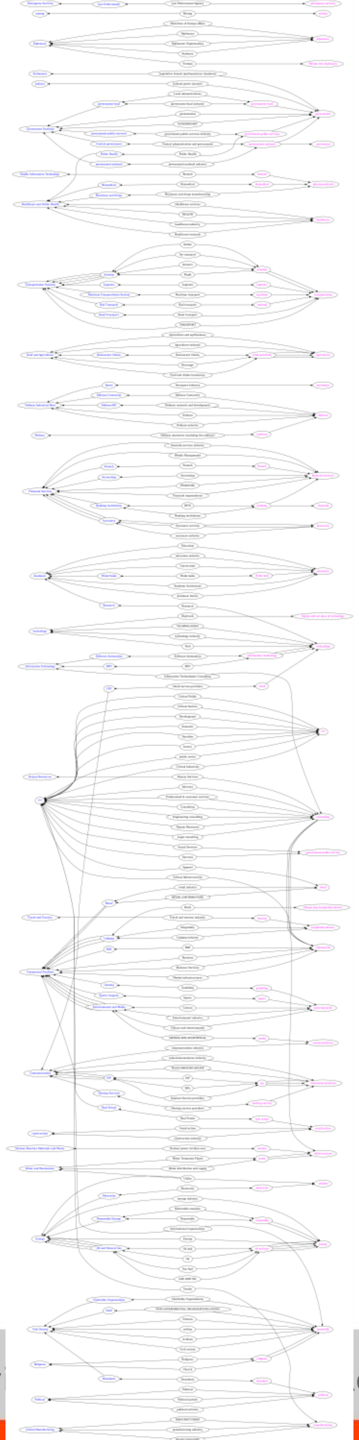
With canonicalized sectors

- Answer Customers' concern: Am I at risk?
- Spot potential correlation between actor groups
- Data is more coherent
- APT researchers are happy
- We can fix the typos

CISA

STIX 2.1


Education
education industry
Universities
Academic
Think Tanks
Think tanks
think tank
education
Academic Institutions
Academic Sector
Research
Research
High-tech
Merge and set alias of technology
TECHNOLOGIES
technology
technology industry
Tech
technology
Software Automation
Software Automation
information technology
Information Technology
BPO
BPO
Information Technologies Consulting
CSP
Cloud service providers
cloud


Emergency Services
Law Enforcement
Law Enforcement Agency
emergency services
mining
Mining
mining
Ministries of foreign affairs
Diplomacy
Diplomacy
Diplomatic Organizations
Embassy
diplomacy
Foreign
Merged into diplomacy
Parliament
legislative branch (parliamentary chambers)
Judicial
Judicial power (justice)
Local administrations
government local
government local industry
government local
governmental
government
GOVERNMENT
Government Facilities
government public services
government public services industry
government public services
Central government
Central administration and government
government national
government
Public Health
Public Health
government national
government national industry
Health Information Technology
Biotech
biotech
pharmaceuticals
Biomedical
Biomedical
biomedical

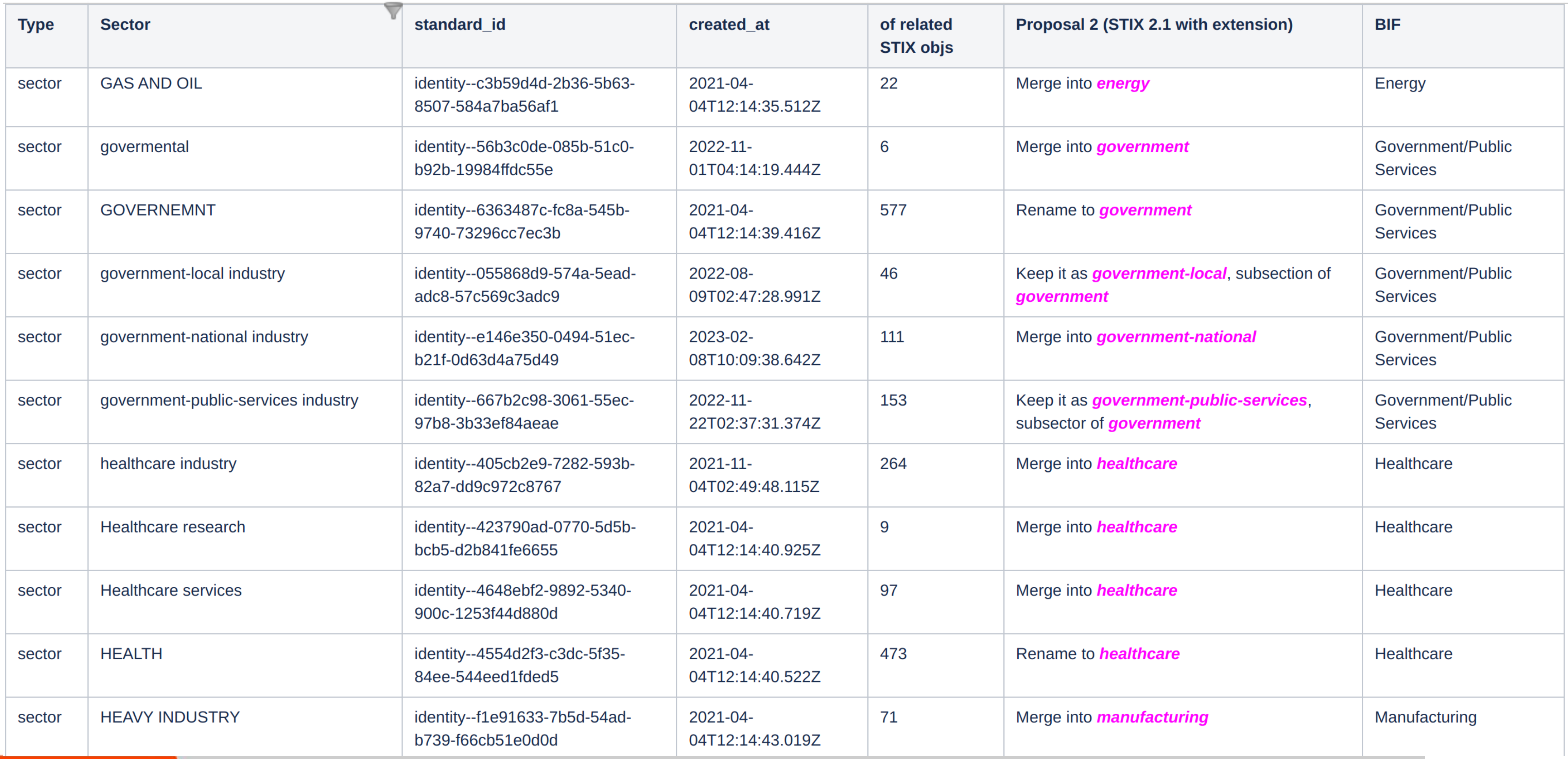
APT Researchers may want these

Proposed subsector	Parent sector (STIX 2.1)	Notes
diplomatic	government	Was: embassy, because diplomatic entities are not necessary embassies.
dissident	non-profit	
media	communications	
political	government	Political parties, DNC-like organizations
information-technology	technology	
~~fintech~~	~~financial-services~~	
~~gambling~~	~~entertainment~~	

Granularity

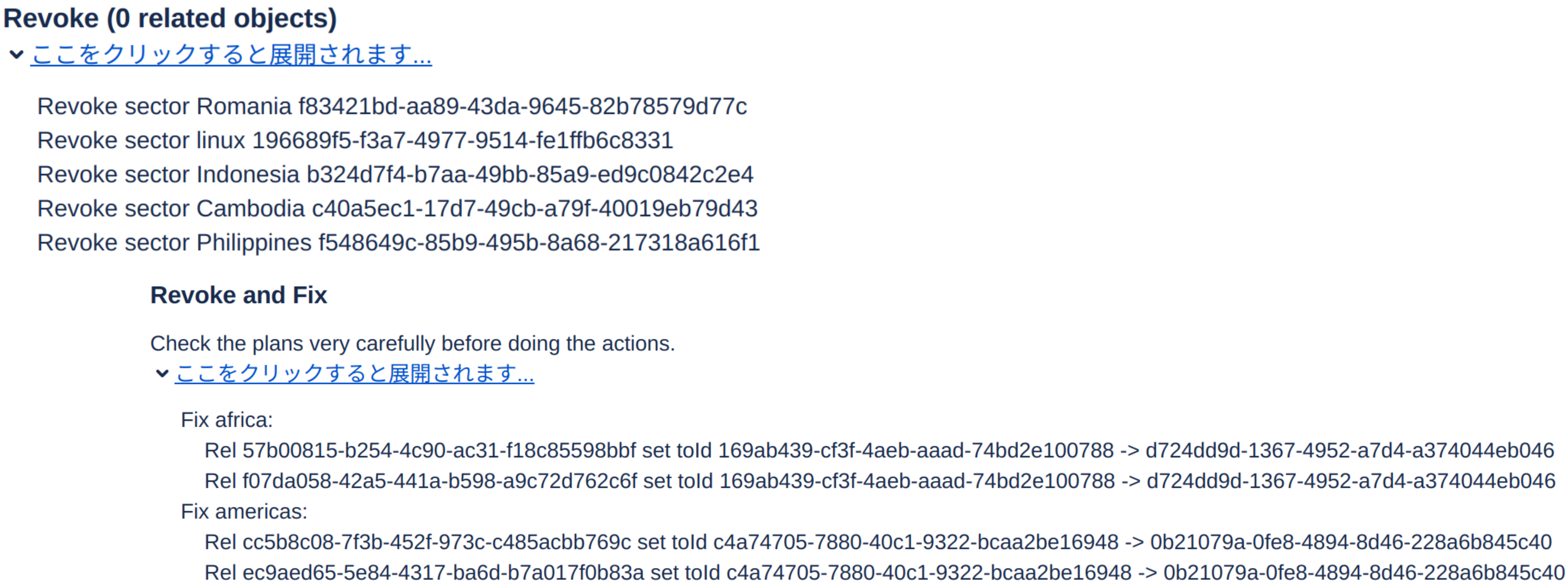
- Tendency to use more nuanced terms
- OpenCTI supports aliases, so we must have consensus
- Ontology might help / externalized

Step 1 – Make a plan



Type	Sector	standard_id	created_at	of related STIX objs	Proposal 2 (STIX 2.1 with extension)	BIF
sector	GAS AND OIL	identity--c3b59d4d-2b36-5b63-8507-584a7ba56af1	2021-04-04T12:14:35.512Z	22	Merge into ***energy***	Energy
sector	govermental	identity--56b3c0de-085b-51c0-b92b-19984ffdc55e	2022-11-01T04:14:19.444Z	6	Merge into ***government***	Government/Public Services
sector	GOVERNEMNT	identity--6363487c-fc8a-545b-9740-73296cc7ec3b	2021-04-04T12:14:39.416Z	577	Rename to ***government***	Government/Public Services
sector	government-local industry	identity--055868d9-574a-5ead-adc8-57c569c3adc9	2022-08-09T02:47:28.991Z	46	Keep it as ***government-local***, subsection of ***government***	Government/Public Services
sector	government-national industry	identity--e146e350-0494-51ec-b21f-0d63d4a75d49	2023-02-08T10:09:38.642Z	111	Merge into ***government-national***	Government/Public Services
sector	government-public-services industry	identity--667b2c98-3061-55ec-97b8-3b33ef84aeae	2022-11-22T02:37:31.374Z	153	Keep it as ***government-public-services***, subsector of ***government***	Government/Public Services
sector	healthcare industry	identity--405cb2e9-7282-593b-82a7-dd9c972c8767	2021-11-04T02:49:48.115Z	264	Merge into ***healthcare***	Healthcare
sector	Healthcare research	identity--423790ad-0770-5d5b-bcb5-d2b841fe6655	2021-04-04T12:14:40.925Z	9	Merge into ***healthcare***	Healthcare
sector	Healthcare services	identity--4648ebf2-9892-5340-900c-1253f44d880d	2021-04-04T12:14:40.719Z	97	Merge into ***healthcare***	Healthcare
sector	HEALTH	identity--4554d2f3-c3dc-5f35-84ee-544eed1fded5	2021-04-04T12:14:40.522Z	473	Rename to ***healthcare***	Healthcare
sector	HEAVY INDUSTRY	identity--f1e91633-7b5d-54ad-b739-f66cb51e0d0d	2021-04-04T12:14:43.019Z	71	Merge into ***manufacturing***	Manufacturing

Step 2 – Draw a detailed plan



Revoke (0 related objects)

⌄ ここをクリックすると展開されます...

```
Revoke sector Romania f83421bd-aa89-43da-9645-82b78579d77c
Revoke sector linux 196689f5-f3a7-4977-9514-fe1ffb6c8331
Revoke sector Indonesia b324d7f4-b7aa-49bb-85a9-ed9c0842c2e4
Revoke sector Cambodia c40a5ec1-17d7-49cb-a79f-40019eb79d43
Revoke sector Philippines f548649c-85b9-495b-8a68-217318a616f1
```

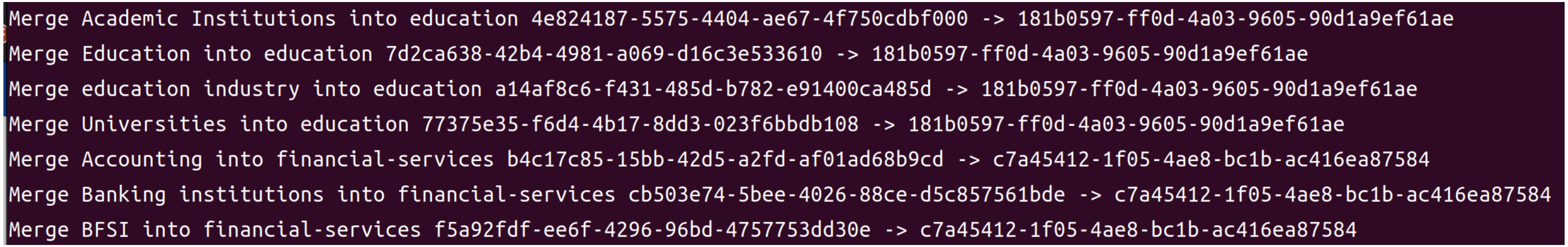
Revoke and Fix

Check the plans very carefully before doing the actions.

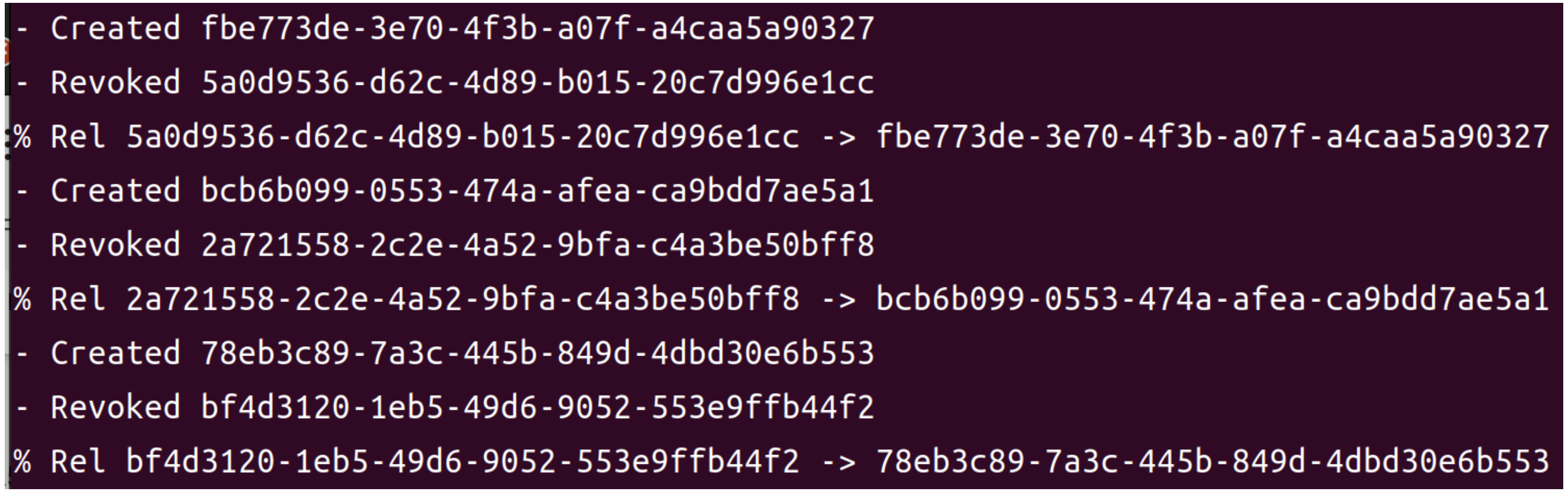
⌄ ここをクリックすると展開されます...

```
Fix africa:
   Rel 57b00815-b254-4c90-ac31-f18c85598bbf set toId 169ab439-cf3f-4aeb-aaad-74bd2e100788 -> d724dd9d-1367-4952-a7d4-a374044eb046
   Rel f07da058-42a5-441a-b598-a9c72d762c6f set toId 169ab439-cf3f-4aeb-aaad-74bd2e100788 -> d724dd9d-1367-4952-a7d4-a374044eb046
Fix americas:
   Rel cc5b8c08-7f3b-452f-973c-c485acbb769c set toId c4a74705-7880-40c1-9322-bcaa2be16948 -> 0b21079a-0fe8-4894-8d46-228a6b845c40
   Rel ec9aed65-5e84-4317-ba6d-b7a017f0b83a set toId c4a74705-7880-40c1-9322-bcaa2be16948 -> 0b21079a-0fe8-4894-8d46-228a6b845c40
```

Step 3 – Log every single action

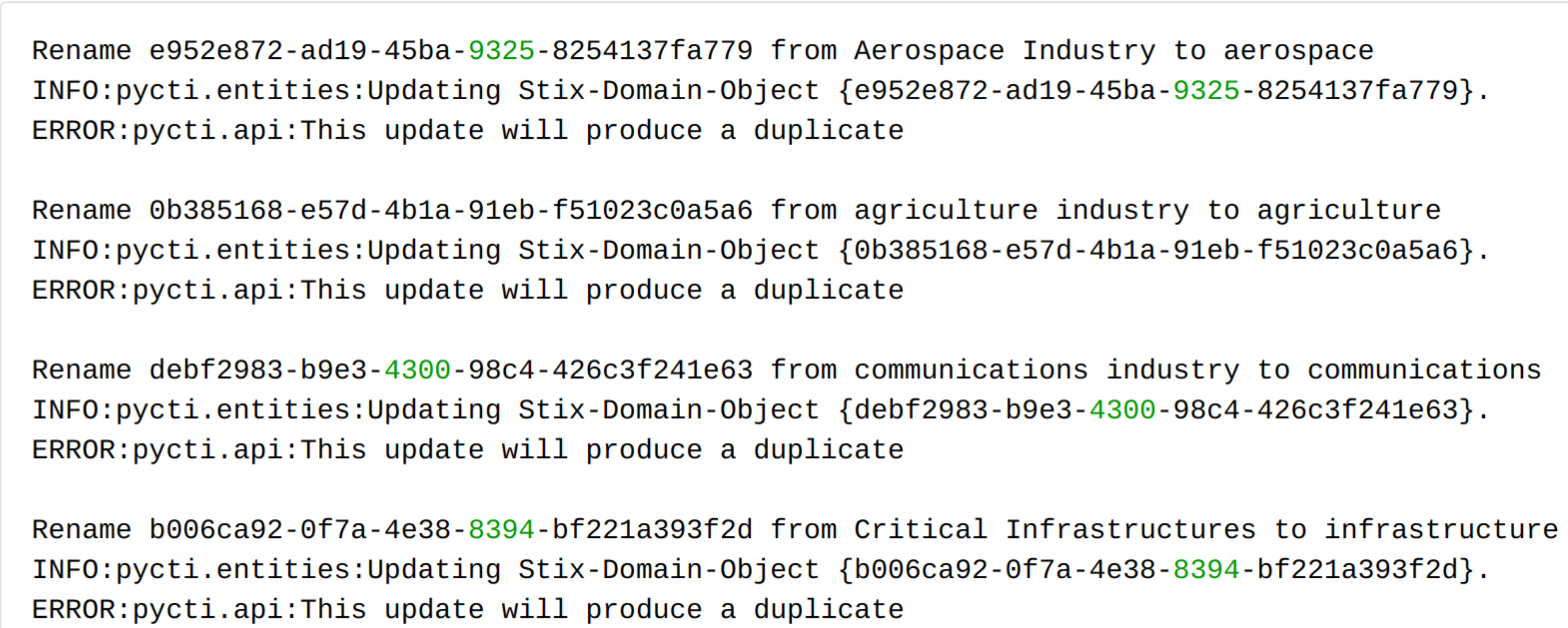


```
Merge Academic Institutions into education 4e824187-5575-4404-ae67-4f750cdbf000 -> 181b0597-ff0d-4a03-9605-90d1a9ef61ae
Merge Education into education 7d2ca638-42b4-4981-a069-d16c3e533610 -> 181b0597-ff0d-4a03-9605-90d1a9ef61ae
Merge education industry into education a14af8c6-f431-485d-b782-e91400ca485d -> 181b0597-ff0d-4a03-9605-90d1a9ef61ae
Merge Universities into education 77375e35-f6d4-4b17-8dd3-023f6bbdb108 -> 181b0597-ff0d-4a03-9605-90d1a9ef61ae
Merge Accounting into financial-services b4c17c85-15bb-42d5-a2fd-af01ad68b9cd -> c7a45412-1f05-4ae8-bc1b-ac416ea87584
Merge Banking institutions into financial-services cb503e74-5bee-4026-88ce-d5c857561bde -> c7a45412-1f05-4ae8-bc1b-ac416ea87584
Merge BFSI into financial-services f5a92fdf-ee6f-4296-96bd-4757753dd30e -> c7a45412-1f05-4ae8-bc1b-ac416ea87584
```



```
- Created fbe773de-3e70-4f3b-a07f-a4caa5a90327
- Revoked 5a0d9536-d62c-4d89-b015-20c7d996e1cc
% Rel 5a0d9536-d62c-4d89-b015-20c7d996e1cc -> fbe773de-3e70-4f3b-a07f-a4caa5a90327
- Created bcb6b099-0553-474a-afea-ca9bdd7ae5a1
- Revoked 2a721558-2c2e-4a52-9bfa-c4a3be50bff8
% Rel 2a721558-2c2e-4a52-9bfa-c4a3be50bff8 -> bcb6b099-0553-474a-afea-ca9bdd7ae5a1
- Created 78eb3c89-7a3c-445b-849d-4dbd30e6b553
- Revoked bf4d3120-1eb5-49d6-9052-553e9ffb44f2
% Rel bf4d3120-1eb5-49d6-9052-553e9ffb44f2 -> 78eb3c89-7a3c-445b-849d-4dbd30e6b553
```

and Traps …



```
Rename e952e872-ad19-45ba-9325-8254137fa779 from Aerospace Industry to aerospace
INFO:pycti.entities:Updating Stix-Domain-Object {e952e872-ad19-45ba-9325-8254137fa779}.
ERROR:pycti.api:This update will produce a duplicate

Rename 0b385168-e57d-4b1a-91eb-f51023c0a5a6 from agriculture industry to agriculture
INFO:pycti.entities:Updating Stix-Domain-Object {0b385168-e57d-4b1a-91eb-f51023c0a5a6}.
ERROR:pycti.api:This update will produce a duplicate

Rename debf2983-b9e3-4300-98c4-426c3f241e63 from communications industry to communications
INFO:pycti.entities:Updating Stix-Domain-Object {debf2983-b9e3-4300-98c4-426c3f241e63}.
ERROR:pycti.api:This update will produce a duplicate

Rename b006ca92-0f7a-4e38-8394-bf221a393f2d from Critical Infrastructures to infrastructure
INFO:pycti.entities:Updating Stix-Domain-Object {b006ca92-0f7a-4e38-8394-bf221a393f2d}.
ERROR:pycti.api:This update will produce a duplicate
```

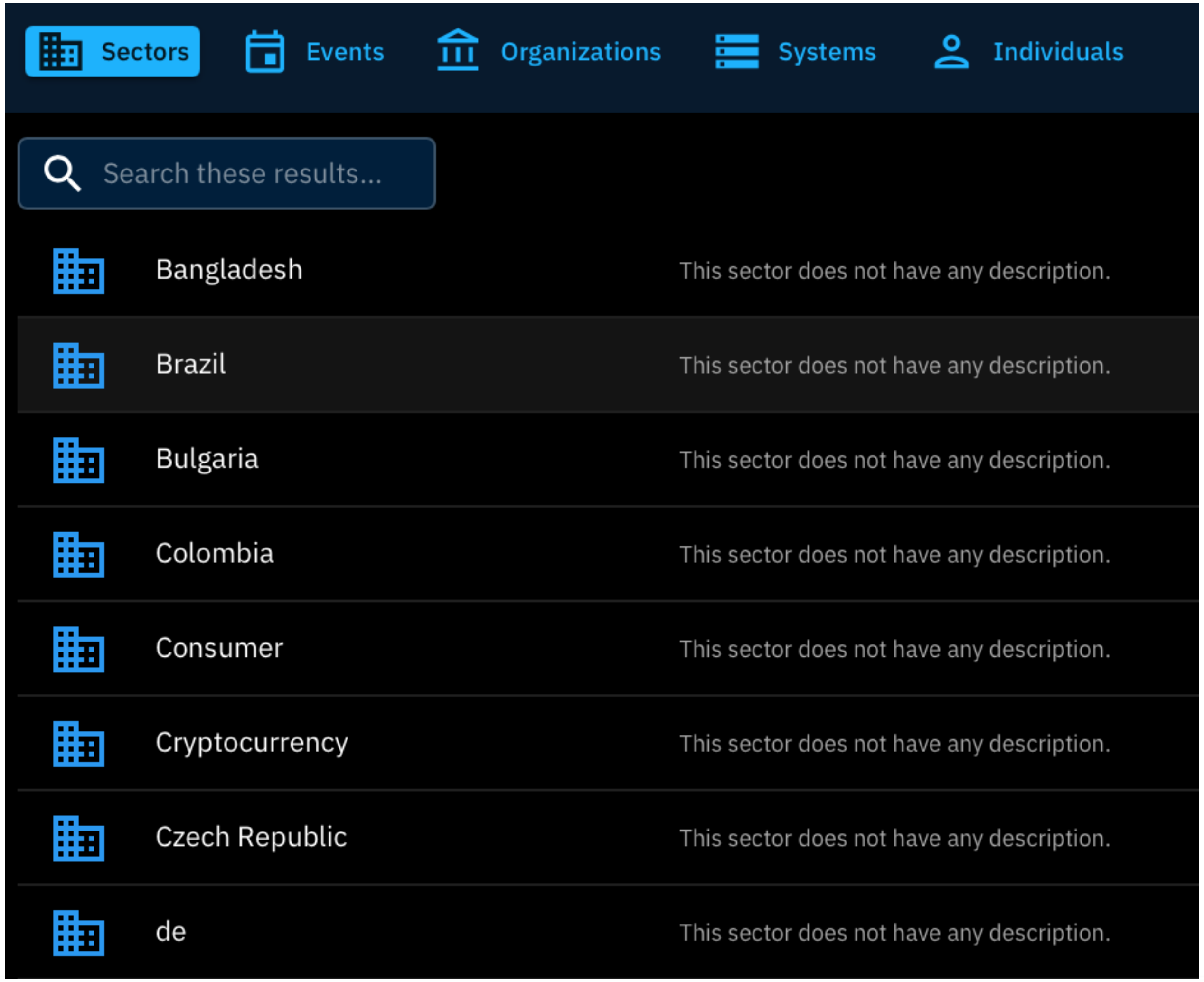
The story doesn't end here.

1. Keep monitoring, as automation creates new sectors.
2. In a big company, you can't always find who did it.
3. As people use it, more chaos will be introduced.

How can we do better?

- Define a clear data pipeline and track your sources
- Do data cleansing before importing them
  ☞ And do not tolerate dirty data.
- Keep original reports.
- (Optimal) Use ontology and try to keep the original sectors as shown in the reports – and do the translation later

Face-palm moment



Bangladesh	This sector does not have any description.
Brazil	This sector does not have any description.
Bulgaria	This sector does not have any description.
Colombia	This sector does not have any description.
Consumer	This sector does not have any description.
Cryptocurrency	This sector does not have any description.
Czech Republic	This sector does not have any description.
de	This sector does not have any description.

2024 **FIRST Cyber Threat Intelligence Conference**

Berlin, Germany
April 15-17, 2024

Josiah Hagen was one of the main contributors of the research.

@miaoski

LinkedIn: mswimmer