Importance of a shared Data Model for Intelligence Collaboration

Andrew Foster
Senior Cyber Threat Intelligence Analyst
Introduction to Foster:

Background:

- 7 years working for the U.S. Department of Defense in various roles, primarily Threat Intelligence & Security Operations
- 2 years at North American Power Company doing Security Operations, Threat Hunting, & Tactical Forensics

Hobbies:

Introduction to EclecticIQ

EclecticIQ: First European supplier of:
- Threat Intelligence Technology
- Threat Information Aggregation
- Fusion Center Operation

Mission: Enable Government & Private Organizations to establish Threat Intelligence Practices

Customers: Central Government / National Cyber Security Centers/ Defense Regulated Industries / Critical Infrastructure (Finance, Telecom, Oil, Etc..)
Overview

• Introduction to and the business value of CTI
• Need for Intelligence Collaboration
• Problems space & existing solutions for Intelligence Collaboration
• Importance of a shared Data Model
• EclecticIQ’s Data Model
What is Cyber Threat Intelligence (CTI)?

The process of analyzing Cyber Threat Information to produce intelligence.

Indicators of Compromise, system logs, antivirus detections, actor behavior patterns, and other forms of evidence can be collected, processed, analyzed, and reported to answer a question; creating Cyber Threat Intelligence.
The need for CTI

- Informs decision makers of the capabilities and intentions of the specific adversaries relevant to their organizations.
- Enables SOC/IR leadership to align security decisions with business priorities.
- Guides network defenders to focus their efforts, improving reaction & response time.
What CTI do outside the IT umbrella?

CTI can inform decisions outside traditional IT, but only if business units engage the CTI team and levy intelligence requirements.

In return, the CTI practice gains a larger pool of stakeholders, re-framing the CTI practice as a core business unit, instead of a cost-center for IT Security.
Strength through Intelligence Collaboration

• Bidirectional exchange of Threat Intelligence strengthens the defenses of all participants.
• An attack on one is an attack on all, changing the adversaries risk vs reward calculation.
• Participation in an Intelligence Collective can reduce the need for in-house analytical talent.
Intelligence Collaboration is your problem,

Analysts match incoming information to intelligence requirements, process data into analyzable format, and make associations between known threats and new threats. That is a different skill set than establishing, building & maintaining relationships with intelligence sharing organizations.

In fields such as Intelligence and other knowledge-based disciplines, the task of ensuring that the appropriate people have access to information should not fall on the shoulders of those attempting to collaborate.

- McIntyre, Palmer, Franks (2013)
...but it’s a manageable problem.

Some of the problems associated with intelligence collaboration remain unsolved.

Most of the technical and legal problems have been resolved and functional real-world intelligence collaboration exists.

CTI Teams that are not actively participating in an intelligence collective (of some kind) are operating in the past.
Unsolved “problems”:

• Approx. 30 companies produce nearly all of the intelligence
• High barrier to entry
• Requires some form of trust
• No quality guarantees
• ‘All or Nothing’ licensing
Why Collaboration was hard:

• Trust
  “Doesn’t sharing intelligence with my business competition give them an advantage?”

• Transport
  “Ok, we have agreed to share data, now how do we do it?”
High Level Solutions - Trust

Information Sharing and Analysis Centers seem to have partially solved the problem of trust by enforcing submission anonymity & vetted access.

This solution is far from perfect, but the model has been proven and is working for many industry groups.
STIX & TAXII specifications have solved the problem of transport. Any vendor can choose to support these technologies, and today analysts are using STIX to consume, produce and share intelligence.

These aren’t the only solutions, but they are the only free & community-driven specifications specifically to solve this problem.
How is Cyber Threat Intelligence made?

Analysts receive an Intelligence Requirement (question). To answer this question, analysts must collect, process and analyze information to produce a product. Once that product is delivered to the appropriate person, enabling them to make a decision, the product becomes intelligence.
Why is Data Model important?

When intelligence begins flowing between organizations, it will need to be reprocessed, so that it can be analyzed.

For analysts, processing is the most boring and time consuming step of the Intelligence Cycle, relying more on engineering skill sets and less on analytical ones.

Establishing a shared Data Model minimizes or eliminates the need for reprocessing.
Doesn’t STIX solve this problem?

If STIX is the language of CTI, your Data Model is its **dialect**.

In STIX, analysts working on the same topic can easily create multiple different interpretations of the same data.

Having a shared language makes exchange of information and intelligence possible, having a shared dialect makes the exchange of information and intelligence **fast and effortless**.
Data Model – Whose responsibility is this?

While more vendors are claiming support for STIX, most do so in a minimally viable way, as vendors are not incentivized to sell intelligence that is accessible outside of their proprietary format.

Intelligence Sharing groups (ISACs and similar) are best positioned to encourage or enforce Data Models, but these efforts are still at the early stages, and enforced Data Models are likely a few years away.
Proposed Solution – EclecticIQ Data Model

The EclecticIQ Fusion Center designed a Data Model tailored to address two specific use-cases:

• Enabling fusion of 3rd Party Data
• Compatibility with STIX 2.x

The Data Model has been used in production since December 2017, and has proved to be a viable Data Model for CTI Analysts.
Bite Size Intelligence

STIX supports multiple different object types, enabling intelligence to be sub-divided.

Expressing intelligence in many small objects, opposed to a single massive pane of glass, gives produces & consumers of intelligence significant flexibility.

Bite size intelligence can be independently updated, analyzed, classified and shared.
Macro vs Micro || General vs Specific

When defending against a Threat Actor, it is important to know what TTPs they employ generally.

When conducting attribution or campaign analysis, nuance becomes more important.

Intelligence expressed in Macro & Micro relationships enables both types of analysis & simplified integration with MITRE ATT&CK.
Designed for Scale & Vendor Neutrality

The Data Model was designed to enable analysts to fuse intelligence at scale, regardless of source vendors, with minimal impact on engineering resources. When engineering is necessary, Non-CTI Engineers can work with Non-Engineer Analysts to design & build integrations with minimal data loss.
Real World Examples:
Data Model – Frequently Asked Questions

• Doesn’t the Data Model prevent analysts from <verb + noun>?

• Don’t you lose some of the data when “fusing” it?

• Does analysis take longer?

• How do you enforce the Data Model?
Thank you for your time

If you would like to discuss Data Models further, please exit the hall, making left turns until you find our booth.

For more reading on:
- our data model - https://tinyurl.com/NIASRED1
- the process of fusing - https://tinyurl.com/NIASRED2
- our fusion center - https://tinyurl.com/NIASRED3