# STIX Sightings

The intent of this document is to provide suggested best practices, use cases, and user feedback questions around STIX 2.x Sighting Objects. Considerations, potential gaps, and open questions for future versions of STIX are given throughout the text in *red italicized* font.

A sighting can be of any STIX Domain Object (SDO), but this document focuses on Indicators and other SDOs that are sighted most frequently, such as Malware, Campaigns, and Threat Actors. The STIX specification does not currently permit or define semantics for sighting STIX Relationship Objects (SRO); therefore, they are outside the document scope. In addition, this paper represents an initial attempt at covering this topic, and is not intended to be definitive, as the concept of Sightings in the context of STIX is still evolving.

The semantics surrounding Sightings are outlined in Section 1, and suggested best practices for producing and consuming Sightings are given in Section 2. Sighting use cases are outlined in Section 3, and a summary of Sighting property uses is given in Section 4. Open questions are discussed in Section 5.

## 1 Semantics of Sightings

Any STIX SDO can be sighted, however the semantics surrounding each type of sighting are not equally understood. Table 1 summarizes semantics that are best understood.

**Table 1**. Semantics of selected SDO Sightings

| STIX 2.x SDO | What it means to be sighted |
|---|---|
| Indicator | The pattern in the Indicator was seen in some data (e.g., host-based data, network/PCAP data, log files, etc.). |
| Malware | A binary file, network traffic, or other evidence associated with the malware or its execution was seen. |
| Threat Actor | A direct sighting of a Threat Actor OR an Attack Pattern, Malware, or Tool Object was seen that is known to be associated with the Threat Actor. |
| Campaign | An Attack Pattern, Malware, or Tool Object was seen that is known to be associated with the Campaign. |
| Intrusion Set | An Attack Pattern, Malware, or Tool Object was seen that is known |

| | |
|---|---|
| | to be associated with the Intrusion Set. |
| Attack Pattern | Evidence of the usage of a particular Attack Pattern was seen. For example, receipt of an email message containing a crafted link and subject line may result in a sighting of a “spear phishing” Attack Pattern SDO. |
| Tool | A binary file, network traffic, or other evidence associated with the tool or its execution was seen. |

- *The CTI TC may want to consider whether Sightings should be updated to capture Sightings of Relationships (Sighting Objects cannot currently capture SROs). For example, “I saw the use of this Malware by a particular Threat Actor."*

- *If the* **sighting_of_ref** *property refers to a Campaign or Threat Actor Object, the* **observed_data_refs** *property would relate to Cyber Observable data associated with a TTP without directly identifying the TTP. The CTI TC may want to discuss whether this is an issue, whether a separate sighting of the TTP is needed, and whether sightings of campaigns and threat actors should be made directly or indirectly (via sightings of their known TTPs).*

# 2 Suggested Best Practices

Best practices for producing and consuming Sighting Objects are given below. Because best practices depend on the SDO sighted and the particular sharing community and organizations involved, several open questions are included.

Note that automation aimed at optimizing the production and consumption of Sighting Objects may be considered in later versions of this document; for example, false positives could be effectively reduced through natural language processing and machine learning implementations.

## 2.1 Producing Sightings (Reporting)

The discussion below relates to sightings of the same SDO (**`sighting_of_ref`** property is constant).

### 2.1.1 Reporting Frequency

The production of Sighting Objects is largely dependent on the procedures and operational processes of the organization that is creating them and the sharing communities they belong to.

However, because sightings may enable others to act more quickly on cyber threat intelligence, we generally recommend that organizations submit and produce Sightings as frequently as possible. In Table 2, we outline several possible frequencies for reporting Sightings, along with their relative pros and cons.

**Table 2**. Pros and cons of sighting frequencies

| Frequency/Type | Pros | Cons |
|---|---|---|
| **Immediately/Daily** | ● May enable quicker response, especially to SDOs that may have shorter useful life spans such as Indicators | ● Potentially high volume, depending on SDO<br>● Some Sightings may be noise<br>● Can impact usefulness with respect to correlation and reconstruction of sighted activity |
| **Weekly** | ● Limits data size<br>● Provides regular snapshot of current threats | ● May not provide context for longer duration SDOs, such as Threat Actors and Campaigns |
| **Monthly** | ● Balances short-term and long-term reporting<br>● Provides historical context for longer duration SDOs, such as Campaigns and Threat Actors | ● Time interval may be too long to be useful for SDOs with a limited lifespan, such as Indicators |
| **Batch-based**<br>(e.g., 100 sightings at a time) | ● Reporting is more predictable<br>● Effort is not expended on once-off sightings | ● Waiting for a batch to fill may delay submission, negating the value of Sightings<br>● One batch size may not suit all use cases |
| **Fixed Frequency / Batch-based**<br>(e.g., 100 sightings at a time OR at the end of the day, whichever comes first) | ● Reporting is more predictable, less erratic<br>● Submissions are not delayed (compared to pure batches) | ● One batch size/frequency may not suit all use cases |

### 2.1.2 Observed Data

Whenever possible (e.g., in cases where it doesn't reveal sensitive or proprietary information) Sightings should include supporting Cyber Observable data via the **`observed_data_refs`** property. Such data can be valuable for various use cases, including help consumers understand the context around the sighting and whether it's something that they should care about or prioritize. For example, the sighting of a particular malware instance as a file on an endpoint may indicate that the malware is able to evade traditional anti-virus based defenses, whereas the sighting of the same malware instance as an email attachment would indicate a

more traditional (and potentially less serious) propagation mechanism. Care should be taken when including custom properties or Cyber Observables for this purpose, as they may limit the value of the Sighting to downstream consumers.

### 2.1.3 New versus Updated Sighting Objects

An existing Sighting Object should be versioned and have its **modified** timestamp updated to incorporate any non-material changes; a new Sighting Object should be created for material changes. Material versus non-material changes are defined next.

Changes to properties that are expected to change regularly, namely **`last_seen`** and **`count`**, constitute *non-material* changes to a Sighting Object. Changes to other non-list-type properties of semantic importance constitute *material* changes (e.g., **`sighting_of_ref`**).

Changes to list-type properties may or may not constitute material changes (e.g., **`observed_data_refs`**, or **`where_sighted_refs`** fields). If existing list values are unchanged and new items are added, the additions constitute a non-material change. However, if an *existing* list value is changed, the change is material.

- *The CTI TC may want to formally define material versus non-material changes.*

Note that the **`first_seen`** and **`last_seen`** properties will convey overall timespan, but they cannot convey timing details such as gaps between sightings. For example, two Sighting Objects could have the same timestamp property values, but one could have been seen daily for three months and the other could have been seen three months ago and today. This is discussed further in the next bulleted item.

Figure 1 illustrates example sighting information captured over three days, which requires two different Sighting Objects (shown in light and dark purple).

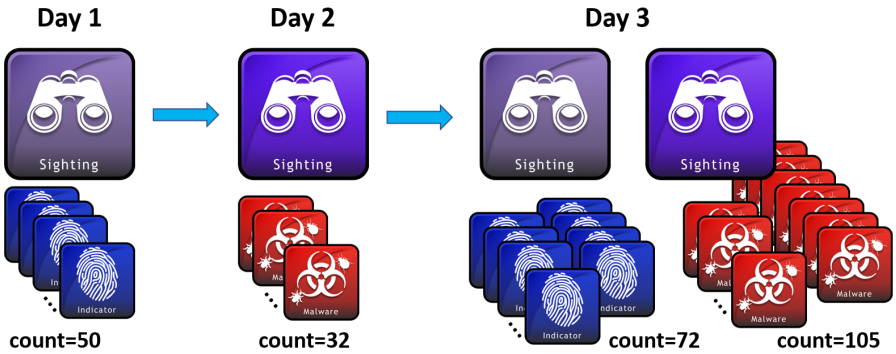


**Figure 1**. Sighting Objects captured over three days

Corresponding JSON examples are given below. For brevity, the associated objects referenced (e.g., observed_data_refs) are not included.

**Day 1**

On Day 1, a new sighting of SDO-A (indicator) is made.


```
{
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--11111111-1111-1111-1111-111111111111",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-10T20:08:31.000Z",
  "modified": "2018-09-10T20:08:31.000Z",
  "first_seen": "2018-09-10T19:00:00Z",
  "last_seen": "2018-09-10T19:00:00Z",
  "count": 50,
  "sighting_of_ref": "indicator--aaaaaaaa-aaaa-aaaa-aaaa-aaaaaaaaaaaa",
  "observed_data_refs": ["observed-data--57485968-034c-cbd6-00ac-33cfee38adcf"],
  "where_sighted_refs": ["identity--fbcd4777-0098-498a-3ffc-0ff342111cfa"]
}
```


**Day 2**

On Day 2, a new sighting of SDO-B (malware) is made.


```
{
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--22222222-2222-2222-2222-222222222222",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-11T10:18:43.000Z",
  "modified": "2018-09-11T10:18:43.000Z",
  "first_seen": "2018-09-11T08:15:00Z",
  "last_seen": "2018-09-11T08:15:00Z",
  "count": 32,
  "sighting_of_ref": "malware--bbbbbbbb-bbbb-bbbb-bbbb-bbbbbbbbbbbb",
  "observed_data_refs": ["observed-data--b67d30ff-02ac-498a-92f9-32f845f448cf"],
  "where_sighted_refs": ["identity--b67d30ff-02ac-498a-92f9-32f845f448ff"]
}
```


**Day 3**

On Day 3, additional sightings are made of SDO-A and SDO-B. A second Observed Data Object, along with the Observed Data Object previously associated with SDO-A, is seen 22 times; seventy-two more sightings are made of SDO-B. None of the changes are material, so the existing Sighting Objects are updated.


```
[
 {
  "type": "sighting",
  "spec_version": "2.1",
```

```
  "id": "sighting--11111111-1111-1111-1111-111111111111",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-10T20:08:31.000Z",
  "modified": "2018-09-12T13:05:00.000Z",
  "first_seen": "2018-09-10T19:00:00Z",
  "last_seen": "2018-09-12T13:05:00Z",
  "count": 72,
  "sighting_of_ref": "indicator--aaaaaaaa-aaaa-aaaa-aaaa-aaaaaaaaaaaa",
  "observed_data_refs": ["observed-data--57485968-034c-cbd6-00ac-33cfee38adcf",
                         "observed-data--45fd19ef-099f-ab22-ff8a-fe4530068bcd"],
  "where_sighted_refs": ["identity--fbcd4777-0098-498a-3ffc-0ff342111cfa"]
 }
 {
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--22222222-2222-2222-2222-222222222222",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-11T10:18:43.000Z",
  "modified": "2018-09-12T05:23:41.000Z",
  "first_seen": "2018-09-11T08:15:00Z",
  "last_seen": "2018-09-12T05:23:41Z",
  "count": 105,
  "sighting_of_ref": "indicator--bbbbbbbb-bbbb-bbbb-bbbb-bbbbbbbbbbbb",
  "observed_data_refs": ["observed-data--b67d30ff-02ac-498a-92f9-32f845f448cf"],
  "where_sighted_refs": ["identity--b67d30ff-02ac-498a-92f9-32f845f448ff"]
 }
]
```


Alternatively, if the new Observed Data Object associated with SDO-A is sighted, but the original Observed Data Object is not, the change is considered material, and a new Sighting Object would be created for SDO-A as follows:


```
 {
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--11111111-1111-1111-1111-111111111111",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-12T13:05:00.000Z",
  "modified": "2018-09-12T13:05:00.000Z",
  "first_seen": "2018-09-12T13:05:00Z",
  "last_seen": "2018-09-12T13:05:00Z",
  "count": 22,
  "sighting_of_ref": "indicator--aaaaaaaa-aaaa-aaaa-aaaa-aaaaaaaaaaaa",
  "observed_data_refs": ["observed-data--45fd19ef-099f-ab22-ff8a-fe4530068bcd"],
  "where_sighted_refs": ["identity--fbcd4777-0098-498a-3ffc-0ff342111cfa"]
 }
```


▼ *The CTI TC may want to consider whether Sightings should convey gaps and trends.*

### 2.1.4 Reporting Multiple Sightings

How should multiple Sightings over some period of time be reported? For example, if there are ten sightings of an SDO today and three tomorrow, how should the sightings be reported?

If differences between the sightings are non-material (see previous question for definitions), it is recommended that they be captured in one Sighting Object that is updated and versioned each time another sighting is made (for immediate reporting) or at the end of each day (for daily reporting). Updating existing Sighting Objects is recommended because it results in fewer objects in the STIX/TAXII ecosystem.

For daily reporting, an update to an existing Sighting Object should be sent each day; otherwise, it will not be clear when sightings were made. For example, if 10 sightings are made one day and three more sightings *two* days later, an intermediate update to the Sighting Object should be sent on the second day that shows the 10 sightings (with an updated **`modified`** property). Otherwise, it will not be clear whether all three additional sightings were made on the third day or or whether one or more were made on the second day.

The sightings of SDO-A in the previous example can be expanded with an intermediate update on Day 2 as follows:

**Day 1**

On Day 1, a new sighting of SDO-A (indicator) is made.

```
{
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--11111111-1111-1111-1111-111111111111",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-10T20:08:31.000Z",
  "modified": "2018-09-10T20:08:31.000Z",
  "first_seen": "2018-09-10T19:00:00Z",
  "last_seen": "2018-09-10T19:00:00Z",
  "count": 50,
  "sighting_of_ref": "indicator--aaaaaaaa-aaaa-aaaa-aaaa-aaaaaaaaaaaa",
  "observed_data_refs": ["observed-data--57485968-034c-cbd6-00ac-33cfee38adcf"],
  "where_sighted_refs": ["identity--fbcd4777-0098-498a-3ffc-0ff342111cfa"]
}
```

**Day 2**

On Day 2, no new sightings of SDO-A are made, but the modified time is updated. The **modified** timestamp could correspond to any interval the producer wants to use. Below, the intermediate update (i.e., no new sightings) is made on the last second of the day.


```
{
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--11111111-1111-1111-1111-111111111111",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-10T20:08:31.000Z",
  "modified": "2018-09-11T23:59:59.000Z",
  "first_seen": "2018-09-10T19:00:00Z",
  "last_seen": "2018-09-10T19:00:00Z",
  "count": 50,
  "sighting_of_ref": "indicator--aaaaaaaa-aaaa-aaaa-aaaa-aaaaaaaaaaaa",
  "observed_data_refs": ["observed-data--57485968-034c-cbd6-00ac-33cfee38adcf"],
  "where_sighted_refs": ["identity--fbcd4777-0098-498a-3ffc-0ff342111cfa"]
}
```


**Day 3**

On Day 3, additional sightings of SDO-A are made, which increases the count. Because the update of the Sighting Object was given on Day 2, it is clear the additional 22 sightings were made on Day 3.


```
{
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--11111111-1111-1111-1111-111111111111",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-10T20:08:31.000Z",
  "modified": "2018-09-12T13:05:00.000Z",
  "first_seen": "2018-09-10T19:00:00Z",
  "last_seen": "2018-09-12T13:05:00Z",
  "count": 72,
  "sighting_of_ref": "indicator--aaaaaaaa-aaaa-aaaa-aaaa-aaaaaaaaaaaa",
  "observed_data_refs": ["observed-data--57485968-034c-cbd6-00ac-33cfee38adcf",
                         "observed-data--45fd19ef-099f-ab22-ff8a-fe4530068bcd"],
  "where_sighted_refs": ["identity--fbcd4777-0098-498a-3ffc-0ff342111cfa"]
}
```


Intermediate updates are especially important over longer periods of time because without historical data, a consumer will only know that “X” sightings were reported between a particular set of timestamps. However, it is up to the consumer of the sightings to understand their distribution over time, either by keeping track of them as they are received, or querying the TAXII server for all versions of a particular sighting.

### 2.1.5 Maximum Age

For reporting Sightings that haven't been reported in the past (for various reasons), what's the maximum threshold in terms of age?

The maximum age threshold (e.g., three months) will vary depending on circumstances. For example, an organization new to a sharing community may want to report past sightings to provide historical context about the cyber threat activity that they've seen.

### 2.1.6 Zero-Count Sightings

When should a Sighting with **`count`**=0 be reported (indicating a window of time when an SDO was not seen)?

The answer depends on the sharing community and the rules that govern it in terms of sightings. However, there are some specific cases where reporting the fact that a sighting was not seen can be useful. For example, if a provider has previously reported a sighting of an indicator but has not seen it in a day or week then reporting that it was not seen this helps downstream systems to potentially clean up deployed rules that were updated based on active instances. This is especially important where downstream systems have limited space for active rules and they need to actively manage rules out.

### 2.1.7 Aggregated Sightings

Should Sightings of the same SDO be aggregated by threat feed providers? For example, should five sightings of the same SDO received from different organizations be aggregated?

Yes, ideally they would be aggregated into a single Sighting Object, as this provides value to consumers of the threat feed; in such Sighting Objects, the **`summary`** property would be set to `true`. However, this assumes any differences between the sightings are non-material, as defined above. For example, a change to an existing list value of the **`observed_data_refs`** property is material—sightings with materially different **`observed_data_refs`** properties should not be aggregated.

Threat feed providers tend to focus on whether an SDO has been seen frequently or infrequently, and less importance is placed on *where* the SDOs were seen. The resulting loss of fidelity with respect to individual sightings may impact aggregation-related metrics generated by organizations that are consumers of the threat feed.

## 2.2 Consuming Sightings (Triage)

The use of Sightings will vary between different consumers because each may have different use cases and will prioritize individual properties differently. Suggestions and considerations for which properties to use to determine priority are given below.

- The **`where_sighted_refs`** property may predict the geographic location of future attacks or the type of agency most susceptible to an attack, enabling an organization to prioritize according to their own location and/or mission. However, it's important to note that this should not be used as a reliable source of attribution, since attackers will often deliberately obfuscate or hide behind multiple proxies to avoid attribution.

  The **`where_sighted_refs`** property may also identify industry sectors most relevant (and of higher priority) to an organization. For example, a user may be part of the financial industry sector and therefore may care only about attacks that targeting the financial industry. Note that the **`created_by_ref`** property corresponds to the *provider* of the intelligence, not necessarily the organization that made the sighting; therefore it's value in prioritization is marginal.

- The **`observed_data_refs`** property may drive prioritization. For example, Sightings involving File Objects may be assigned higher priority than those involving URL Objects.

- Similarly, the **`sighting_of_ref`** property may also drive prioritization.

- The **`confidence`** property of the SDO sighted may support prioritization. For example, sightings of SDOs for which there is higher confidence in the correctness of their data may be given correspondingly higher priority. This information must be considered in conjunction with who is asserting the confidence.

- The **`first_seen`** and **`last_seen`** properties may also support prioritization. For example, an organization may prioritize Sightings seen recently or those which have been seen over a long span of time.

# 3 Use Cases

A variety of use cases for Sightings Objects are given below. A table summarizing the roles of Sighting Object properties is provided in Section 4.

## 3.1 Situational Awareness

Sightings can be used by a SOC/Detect Team to determine new threats and identify trends. For example, sightings of TTP SDOs such as Malware, Attack Patterns, and Tools can help an organization prioritize their detection of these active threats.

## 3.2 Transport Efficiency

A Sighting Object enables capture of relationships between an Indicator and multiple objects in a single payload. For example, the Sighting Object shown below captures a sighting of two URLs (Observed Data Object) in the banking sector (Identity Object) in Germany (Location Object)—all with a single JSON object. An associated illustration is given in Figure 2.

```
{
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--8495dfea-ff89-3722-fe9a-9887de2f3cee",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-09-10T15:28:12.000Z",
  "modified": "2018-09-10T15:28:12.000Z",
  "first_seen": "2018-09-10T15:00:00Z",
  "last_seen": "2018-09-10T15:00:00Z",
  "count": 47,
  "sighting_of_ref": "indicator—0002de09-3434-9a2e-221d-907786dafbb1",
  "observed_data_refs": ["observed-data--57485968-034c-cbd6-00ac-33cfee38adcf",
        "observed-data--94732dffe-899a-ed23-0042-cff26378ade1"],
  "where_sighted_refs": ["identity--fbcd4777-0098-498a-3ffc-0ff342111cfa",
        "location--ab872900-0392-39af-3ccd-cb890009ff76"]
}
```


**Figure 2**. Sighting Object embedded relationships

## 3.3 Indicator Confidence

Sightings can be used as a measure of confidence for Indicator Objects (when the **`sighting_of_ref`** property references an Indicator Object), independent of the actual **`confidence`** property that exists on Indicators and other SDOs. For example, up to some threshold, more sightings of an indicator (in an aggregate sense) might give higher confidence that an Indicator is effective; conversely, too many sightings might suggest an Indicator is too noisy to be useful.

It is important to consider other factors as well: confidence in an Indicator shouldn't be based solely on sightings. One factor of note is the Opinion Object, which captures other organizations' assessment of the correctness of information, such as an Indicator. Another is whether an indicator is seen by a single source or multiple sources.

The potential significance of Sighting count will vary based on Indicator pattern type (e.g., DNS, SMTP, WCF, etc.), and context will require Indicator type-specific thresholds. For example, currently, NCCIC E3A Operational Instructions say that DNS Indicators with more than 5,000 Sightings per hour and SMTP Indicators with more than 25 alerts an hour shall not be deployed.

Accordingly, it might be interesting to consider whether it is possible to define useful thresholds based on the number of Sightings for an indicator. For example:

- **`count`** = 0/week → Indicator is inactive, ineffective, or not implemented properly. Response to indicators with zero counts will vary, depending on the Indicator, and some indicators will be deployed long term, regardless of count.
- 0 < **`count`** < 10/day → Indicator is useful.
- 10 < **`count`** < 100 → Indicator is effective.
- 100 < **`count`** → Indicator may be noisy.
- 10000 < **`count`** → Indicator is so noisy as to be meaningless.

## 3.4 Indicator Data Feeds

This section provides use cases of how Sightings of Indicators can be used operationally.

### 3.4.1 Presence/Absence of Sightings

- Presence of Sightings might suggest the Indicator is (still) valid.

- Absence of Sightings might suggest the Indicator is ineffective or the TTPs have changed. Or it could indicate a problem of instrumentation (e.g., one IDS can find the Indicator, but others cannot). Studying this data could allow refinement of valid time windows or Indicator patterns.
- If the time between the **`first_seen`** and **`last_seen`** properties is large and there are few false positives, the attacker/threat might be considered persistent.
- If the **`last_seen`** property is more than X days ago, the Indicator might be stale. This would require reviewing historical Sightings for an Indicator because there are no new Sightings.

### 3.4.2 Sighting Counts

- Sighting counts above some threshold might indicate the Indicator is valid, increasing confidence in the Indicator, or it could mean the Indicator pattern is too inclusive and needs refinement. It may be important to distinguish between overall sightings counts and counts per organization. Confidence might change dynamically as the number of recent sightings changes.
- A sudden spike of Sightings of an old Indicator might suggest the associated Threat Actors or TTPs (e.g., malware) are again active.
- A low number of Sightings might suggest Indicator is of low quality or is stale.
- Sightings from one organization that a second organization should have also seen (but did not) might point to Indicator and/or detection problems at the second organization.
- Sightings on a per-organizational basis could be used to track narrow- versus wide-targeting by specific Threat Actors.
- Sightings could be tracked by sector (or another factor). Doing so would support the need for more context on Indicators.

### 3.4.3 Sighting Details

- The **`observed_data_refs`** property may help refine the Indicator. For example, if an Indicator specifies a thousand IP addresses, but only a few are seen, the Indicator IP list could be truncated. As another example, it may be possible to extract more indicators, if the **`observed_data_refs`** property of an Indicator for an email address includes a full email message from that sender.

- Geographical location of Sightings (**`where_sighted_refs`** property) may predict location of future attacks. Other identity details may offer similar value (sector, nationality, language, etc.).

### 3.4.4 Sighting Accuracy and Indicator Confidence

This use case requires distinguishing a true positive from a false positive. It might require the sighting producer  to have a mechanism for false positive rate reporting.

- Proportion of true positive versus false positive sightings may change confidence in the Indicator. This could be done dynamically as the accuracy of recent sightings changes.
- False positives would suggest decreased confidence in Indicator.

### 3.4.5 Network Role/Location of Device

Note that it may not be possible for the sighting producer to get this information.

- The type of equipment reporting the Sighting of the Indicator (e.g., honeypot vs backbone router) affects the meaning. Such information would be captured via the **`where_sighted_refs`** property of a Sighting Object by setting the **`identity_class`** property of the associated Identity SDO to “system.” Details would be specified in the Identity’s **`description`** and/or **`roles`** properties (e.g., “honeypot”).
- A false positive on one type of device and not on another may indicate different things about the Indicator.
- Sighting information shared among members of a collaborative threat sharing group enables organizations to leverage collective defenses/sensors. In other words, Sightings leverage visibility provided by distributed sensors.

### 3.4.6 Prioritization Based on Sighting Interest Level

Targeted/APT Sightings are of higher interest than generic Sightings. Therefore, CTI from Organizations that (sometimes) make more interesting Sightings might be prioritized. See Section 2.2 for further discussion of prioritization.

### 3.4.7 Sighting Extrapolation

It may be possible to extrapolate Sighting information:

- A Sighting may indicate other potential targets (e.g., **observed_data_refs** might include email addresses of other organizations (targets)).
- Variation between similar sightings can lead to broader and stronger indicators. For example, multiple sightings of a phishing-oriented Indicator might lead to discovery of a domain name pattern. Details (raw data) would likely be necessary.
- Targets from similar Sightings can be studied to determine a larger target list, leading to warnings to other organizations.
    - For example, if both Boeing and Lockheed report a Sighting, maybe attendees of an aerospace conference are being targeted, leading to questions such as, “Was the attendee list shared publicly?” or “Was the conference organizer compromised?”
    - Or, if Sightings are reported by organizations with names, email addresses, etc. early in the alphabet (e.g., A-K), then organizations with names at the end of the alphabet might be targeted next; but with warning, those organizations can prepare.

### 3.4.8 Sharing with Restrictions

A Sighting Object can be used by organizations with information sharing restrictions to relay minimal, useful information. For example, the Sighting Object below says that the indicator is valid and was seen, but other details, such as associated Observed Data objects, are not specified.

```
{
  "type": "sighting",
  "spec_version": "2.1",
  "id": "sighting--ee20065d-2555-424f-ad9e-0f8428623c75",
  "created_by_ref": "identity--f431f809-377b-45e0-aa1c-6a4751cae5ff",
  "created": "2018-07-10T20:08:31.000Z",
  "modified": "2018-07-10T20:08:31.000Z",
  "sighting_of_ref": "indicator--8e2e2d2b-17d4-4cbf-938f-98ee46b3cd3f"
}
```

## 3.5 Open FAIR Risk Analysis

Sighting details (**first_seen**, **last_seen**, **count**, **where_sighted_refs**, **sighting_of_ref**, **observed_data_refs** properties) could inform an Open FAIR risk analysis. Analysis of the **observed_data_refs** dataset could be used to calculate annual Threat Event Frequency and inform the estimate for Threat Capability, particularly when the analysis of multiple threat actors

are compared. See The Open Group library of documents related to FAIR at https://publications.opengroup.org/editors-picks/open-fair?limit=15.

# 4 Summary of Sighting Property Use

Highlights of property use are given in Table 3. Details are not included, but references to Sections 2 and 3 are given. All Sighting-specific properties are listed; common properties (blue shaded cells) are listed only if they are discussed in the document as enabling analysis. Descriptions are taken from the STIX 2.1 specification.

**Table 3**. Use of Sighting properties

| Property | Description | Sighting Use |
|---|---|---|
| **confidence** | Level of confidence that the creator has in the correctness of their data. | Sightings of SDOs for which there is higher confidence in the correctness of their data may be given higher priority. See Section 2.2. |
| **type** | The value of this property **MUST** be `sighting`. | n/a |
| **first_seen** | The beginning of the time window during which the SDO referenced by the **sighting_of_ref** property was sighted. | The **first_seen** and **last_seen** properties may drive prioritization. For example, an organization may prioritize Sightings seen recently or those which have been seen over a long span of time: see Section 2.2. The values can also relay information about an associated Indicator or Threat Actor: see Section 3.4.1. |
| **last_seen** | The end of the time window during which the SDO referenced by the **sighting_of_ref** property was sighted. | |
| **count** | The number of times the SDO referenced by the **sighting_of_ref** property was sighted. | The significance of sighting **count** will vary based on Indicator pattern type (e.g., DNS, SMTP, WCF), and context will require Indicator type-specific thresholds. See Section 3.3. |
| **sighting_of_ref** | An ID reference to the SDO that was sighted (e.g., Indicator or Malware). This property **MUST** reference only an SDO or a Custom Object. | The value of this property will impact the significance and meaning of the other properties. An organization may prioritize some types of SDOs over others. See Section 2.2. |

| | | |
|---|---|---|
| **observed_data_refs** | A list of ID references to the Observed Data objects that contain the raw cyber data for this Sighting. | The SDO type may drive prioritization. For example, sightings involving File Objects may be assigned higher priority than those involving URL Objects. See Section 2.2. |
| **where_sighted_refs** | A list of ID references to the Identity (victim) objects of the entities that saw the sighting. | Indications and predictions of the geographic location of future attacks or the sector most susceptible to an attack may drive prioritization. See Section 2.2. |
| **summary** | Indicates whether the Sighting should be considered summary data. Summary data is an aggregation of previous Sightings reports and should not be considered primary source data. Default value is `false`. | Aggregating Sightings with non-material differences via the **summary** property provides value to consumers of threat feeds. See Section 2.1.6. |

# 5 Open Questions/User Feedback

Questions related to Sightings are posed below with user feedback included below each item. Additional input from other parties (e.g., SOC analysts and larger organizations) would be useful.

- How willing are people/organizations to share Sighting Objects?
    - Sharing information is common.
    - Collaboration and transparency are key principles in cybersecurity. However, larger organizations are reluctant to share information given the competitive and globalized world. Government intervention might help enforce cyber threat information (CTI) sharing.
- What capability gaps are in the current implementation of Sightings?
    - Sightings of STIX Relationships Objects (SROs) should be considered.
- The scope of this document is limited to Sightings of SDOs that are best understood. Do users have experience that can be shared to expand the “Semantics of Sightings” section to types of SDOs that aren't covered currently?

- Should Sightings be transitive? For example, if a sighting is made of an attack pattern used by a threat actor, should the sighting imply the threat actor?
- Some SDOs can be sighted in two ways: (1) indirectly, through a Sighting of an Indicator that “indicates” the SDO, or (2) directly via a Sighting of the SDO itself. For example, a Sighting can be made of an Indicator SDO where its **`pattern`** field indicates a Malware SDO (an indirect sighting of malware); or a Sighting can be made directly of the Malware SDO. Should sightings of SDOs be captured directly or indirectly? Do indirect and direct Sightings of an SDO have the same meaning?