System Requirements and  
Concept of Operations  
for  
  
**FROM LOCAL TO GLOBAL AWARENESS: A DISTRIBUTED INCIDENT MANAGEMENT SYSTEM (DIMS)**

**Table of Contents**

1.2 MNS Capability Gap

The principal high-level gaps in the missions described in the previous section that exist have to do with the availability and affordability of tools that support those missions. Each of them has limitations or impediments to their use:

* There are managed security services that could be engaged to handle all security incident response and forensics. The cost of these services is prohibitive for all but the most serious incidents with potential losses that rise to the level of existential threats to the viability of the enterprise. The availability of affordable open source tools to improve response and recovery is a gap that DIMS is intended to fill.
* There are agent-based systems and network-based that can provide the level of detail and pervasive collection of event data at the host, server, and network levels. These, too, are prohibitively expensive. They only work in environments where policy can dictate the deployment of agents on all end hosts and servers, and where network topology and administrative responsibility at the enterprise level is such that one group can deploy, manage, and interact on a daily basis with the security system. Most SLTT government sites cannot afford to have this level of in-house security monitoring and response capacity. At present, even if one site in a region can afford such capabilities, their use is limited to protection of that site alone and there is little benefit to other inter-related entities in the region (hence the need to share not only IOCs and Observables, but also Course of Action and analytic results.)
* Most SIEM systems focus on the problem of collecting and correlating millions of events per day, distilling them down to a reasonable (N<=100/day) level, and directing them to the entities with administrative control over the system identified in the alerts. Correlation across a confederated population is not typically done (most deployments are for one enterprise, perhaps with multiple business units under the same top level corporate structure). These systems are also primarily focused on detection and alerting on input of events, not on after-the-fact triage and respond/recover operations. When they do support forensic analysis of past events, these systems typically do not support confederated cross-organizational correlation and collaborative response (e.g., by sharing analysis between multiple enterprises, or distributing Course of Action information.)
* The existence of the Ops-Trust community proves that volunteers can self-assemble to respond and react to issues that impact everyone on the internet, but these groups frequently operate on email and chat communication channels that are unstructured, ad-hoc, and are very difficult to keep up with. Unless one reads every message in every email thread, extracts all attached files or processes all in-line data, and manually searches for IOCs and Observables that can be manually used to search data sources that that person controls, the benefit of information sharing is lost. And for any emergent situation of global significance, the threads are many and the messages in each thread can flow for days or weeks. It is impossible to keep up with this without moving to structured data and machine processing to identify messages of interest.
* There have been many formats for structured security data sharing developed over the years. Each one has seen a similar lifecycle, where there is interest and excitement at the start of the project, a slow deliberative process of developing the standard, going through the process of vetting and acceptance of the standard by an official body, and then a push to get the industry and researchers to adopt the standard. STIX may encounter this same fate. It is too early to tell. What some (like Wes Young, developer of the Collective Intelligence Framework) suggest as an alternative is to “blow up the standards process”[[1]](#footnote-6) and simply implement something quickly, get it used by as many people as possible, adapt and modify it to address limitations that are encountered, and keep moving forward. “We believe traditional standards processes not only have a high barrier to entry, but are often slow and use the design by committee approach. We believe the best way to create a protocol is from the ground up using CONOPs. Push design out to the edge and let operations influence design in real-time.”[[2]](#footnote-7)

**CURRENT SITUATION**

**SECTION 2: Operations and Support Description**

This section identifies and explains the mission objectives of the DIMS project in terms of users and other stakeholders, and specific operations and support missions.

As the DIMS system relies upon and integrates multiple existing and future open source software components, it will be developed using an Agile programming development methodology (as opposed to the classic “waterfall” development methodology with its sequential processes.) This document, therefore, is a “living document” that will be updated as the project proceeds and as cyclic input/feedback from users and testers is received. Sections to be addressed in future releases of this document are listed as **TBA**.

2.1 Missions (Primary/Secondary)

The primary mission objectives for the DIMS system are operational in nature, focused on facilitating the exchange of operational intelligence and applying this intelligence to more efficiently respond and recover from cyber compromise. The secondary mission objectives are to create a framework in which tools to support the primary mission objectives can more quickly and easily be integrated and brought to bear on advancing techniques on the attacker side or the equation. These missions will be described in this section in detail by way of examples of how users and stakeholders will use the DIMS system.

* 1. Users and Other Stakeholders

DIMS is being designed to facilitate trusted information sharing among multiple stakeholder groups, as well as enhancing the ability of a federate group like the PRISEM membership to manage security incidents using a model of event and alert information sharing. A diagram showing a representative subset of these stakeholders and the types of data that will be shared is seen in Figure 3.

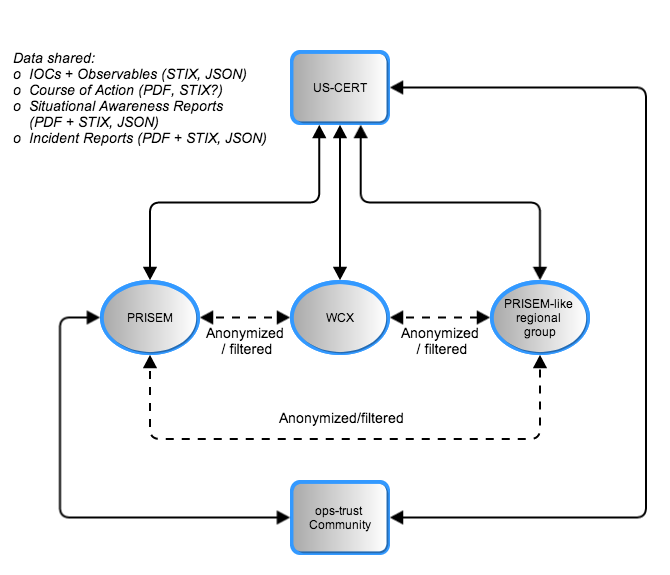
* + 1. 

Figure - Data Flows Between Stakeholders

1. Mission Support Scenarios

The following general Mission Support Scenarios focuses on improving the efficiency of daily communications workflow processes.

* + 1. **Tracking Status of Remediation Efforts**

A regular occurrence within the Ops-Trust community is someone reporting a large number of hosts or network autonomous system (AS) numbers that have vulnerable, exploited, or infected computers. The Subject line usually reflects something about the data (e.g., “1.2M NTP amplifiers identified”) Members of the list will read these email messages, extract the list from the body of the message or attached files, process the list (often with a custom script), and do what they can to mitigate the threat within their own network. Some will respond to the email with something like “ACK for AS123, AS456, and AS678”. While these acknowledgement messages are nice, nobody is responsible for tracking them, updating a list with status, etc. It is impossible for one to know, without themselves tracking the entire thread and accumulating the results from all responses, what percentage of the original list of 1.2M items has been mitigated, which ones are left, etc. Such lists are sometimes sent in the body of the message in what is known as a “Cymrufied list” (columns of IP addresses, AS numbers, etc, separated by vertical bar “|” characters, made popular by Team Cymru.[[3]](#footnote-12) See Figure 12). Sometimes they are Excel spreadsheets attached to the message, or Comma Separated Value (CSV) files. Sometimes people just put a CIDR block in the Subject line of a message. The method is ad-hoc, random, and often requires writing custom scripts to process and extract just the data relevant to one’s own network. It is not uncommon to receive a “Cymrufied list” that is placed in a GZIP compressed Unix/Linux tar archive file, which is then attached to an email message (necessitating extraction, unpacking the archive, processing the included file with a script, then deleting the .tar.gz file, all *manually*.)

The DIMS system will automate this process by supporting the automatic recognition and processing of structured data files either uploaded into the system, attached to email messages, or sent over TAXII or an AMQP message bus. These structured files can then be processed and the context used to track activity (i.e., is this the initial report, an acknowledgement that certain items have been mitigated, etc.) This also allows tracking of the status of mitigation, statistics over time, etc.

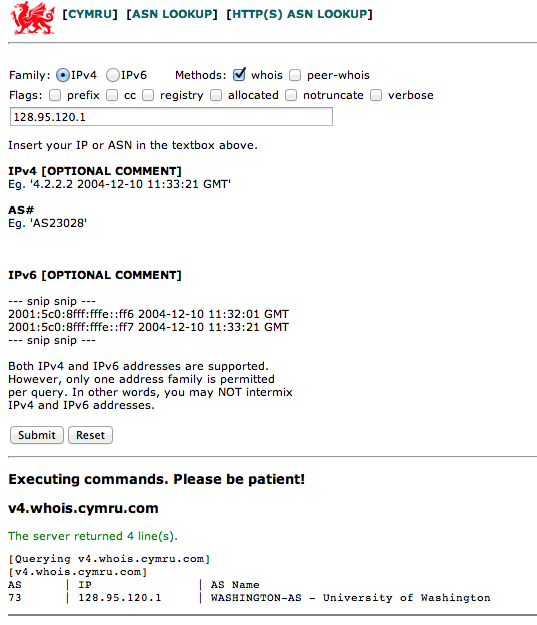


Figure - "Cymrufied list" Example

* + 1. **Situational Awareness Through “Identifying Friend or Foe”**

When trying to analyze events and alerts in a haystack of data, one method of extracting meaning from the data is to organize it according to facts that are known about the entities that are identified in the haystack of data. A first order of meaning can be derived from taking the end points of connections and categorizing them according to which sets they belong to: known to be a PRISEM participant (a.k.a., “friend”), or known to *not* be a PRISEM participant.

Figure 13 illustrates how organizational top-level domains and/or CIDR blocks for a subset of PRISEM participants are mapped to their Site ID strings and chosen anonymization strings (i.e., the label that participant would like to use to mask their internal IP addresses and host names in reports that are shared outside the trust group.) When events are logged, and those logs are ingested into the PRISEM system, they are processed so as to associate them with the site from which they came. Once in the historic log archives, an analyst may search for a specific observable (e.g., “show me all connections to/from a specific suspect IP address.”)

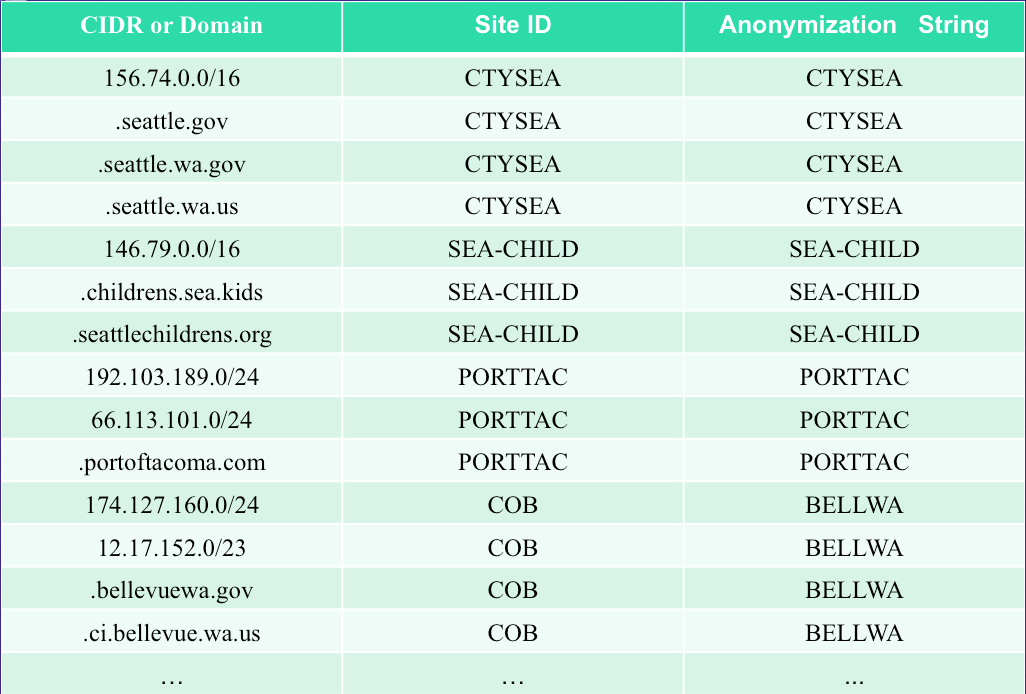


Figure – Partial Participant ID Mapping

Using this mapping of domains and CIDR blocks to participants, it is possible to identify all records in search results that are associated with any of the PRISEM participants, count how many discrete hosts within each participant site were found, and produce cross-organizational correlation statistics that describe the percentage breakdown of all identified records in the search results. An example of what this process produces can be seen in Figure 14. In this example, hosts from seven different PRISEM sites were found, with the three most frequent results being in Seattle Childrens Hospital (70.65%), Kitsap County (26.61%), and Port of Olympia (1.38%).

Making only one pass over a set of data only allows us to extract IP address and domain names known to be in the map, or not in the map, deriving two non-intersecting sets of entities that are either “matching” and “not matching”. This is depicted graphically with the Venn diagram in Figure 15. Without any other information or context about the “not matching” entities that were identified, there is not much that can be deduced about those entities, other than they were involved in connections associated with whatever the analyst was searching for. We can define the results of this pass as identifying “friend” (because we are using a mapping of what constitutes “friend” sites). This is, in fact, how the output of the Cross Correlation service is tagged in Figure 14.

Now that we have the list of entities that are not our “friends”, we can make a second pass and add context that will be useful in helping make decisions. Rather than just “known” and “not known,” we can determine, based on information provided by selected authorities to have a certain level of probability of being involved in malicious behavior, that an end point of communication is believed to be hostile (a.k.a., “foe”). The Collective Intelligence Framework accumulates reputation data from sources that the security community deems to be trustworthy in determining which are malicious. If an IP address or domain name occurs in a CIF feed of 65% confidence, then we can assume with 65% confidence that any connections from a PRISEM participant are highly suspicious indicators of malicious activity. If that IP address is not known to *any sources that feed CIF*, it may or may not be malicious. It could be associated with an “advanced persistent threat” actor who performs targeted attacks and evades the security industry’s sandboxes. Or it could be a totally innocent new social network site related to an animal rescue organization. The context and search criteria used by the analyst to get the data being processed holds some clues as to whether the connections are innocent or malicious, and adding context regarding reputation from the security industry and researchers assists even more in making a determination of “innocent” or “malicious” activity.



Figure - Cross-organizational Correlation of Query Results (Redacted)

**4.0 Functional Capabilities**

The functional capabilities of the DIMS system are described in this section as a set of high-level *user stories* as used in Agile programming to provide sufficient detail to developers to allow them to code features that deliver the desired capabilities. These stories will be expanded upon and/or altered based on feedback from stakeholders and prospective users, and will form some of the test and evaluation criteria used to validate the final product.

Users of the system fit into one or more of several categories. They may be analysts, investigators (either private, or law enforcement), system or network administrators, security operations staff, network operations staff, managers, or CISOs. Similar categories are described in research into management of forensic or incident response processes. (Beebe, 2005; Ciardhua ́in, 2004; Hutchins, 2011; Khurana, 2009; Ieong, 2006) In the following user stories, multiple roles may be listed in braces or the general term “user” may be applied. (The stories presented here will need sketched out prototypes of the type of output being described to guide programmers in producing the desired user interface and presentation mechanism.)

1. ~~“As {a security operator, investigator, analyst, CISO} I want to be able to define multiple sets of attributes that the system can then use to inform me about when new data is seen that matches those attributes. Attributes can include anything that might be seen in indicators of compromise, observables, or alerts. (The most basic being IP addresses and/or CIDR blocks, domain names, MD5 or other cryptographic hash values, file names, Registry key settings, etc.)”~~
2. ~~“As {an investigator, analyst} I want to be able to keep track of cases and campaigns (i.e., groups of related incidents). I want the system to inform me, if I so chose, of any time new data that is determined to be associated with the sets I am tracking comes into the system. For example, if I log in and open a case, I can easily tell which data has been entered into the case since the last time viewed the case. This allows me to stay on top of new evidence or activity that I am investigating.”~~
3. ~~“As {a security operator, investigator} I want to be told when an email thread or received set of indicators includes systems that I am responsible for securing, ideally pointing out to me those hosts that are involved without requiring that I read the entire thread, extract attachments, write scripts to parse and search data, etc. I want to be given a list of those records that are important, in a format that I can submit directly to query interfaces without having to write scripts to parse and process.”~~
4. ~~“As {an investigator, analyst} I want to be able to preserve the results of searches, and in some cases the data that was identified while searching, in order to have copies that are subject to expiration and purging from the system. Some investigations may take many months, which could bump up against the data retention period (approximately 12 months, at present).”~~
5. ~~“As an investigator, I would like to be able to timestamp files I create (i.e., calculate multiple different cryptographic hashes of the contents of files to validate their integrity, associate a timestamp from a trusted time source, then cryptographically sign the result with a private key). This allows validation of the existence of a file at a point in time, who produced the file, and maintenance of a form of “chain of custody” of the contents of the file. To ensure privacy as well as integrity and provenance, the file would first be encrypted (or both cleartext and encrypted files included in the timestamping operation).”~~
6. ~~“As an {analyst, investigator, security operator}, I would like to be able to get context about “external” hosts that includes what kind of malicious activity has been observed, by whom, starting and ending when, have they been involved in precious incidents I have dealt with, etc. This view could combine a timeline aspect (first seen to last seen time ranges along the X axis), for one or more sources of threat intelligence (discrete items along a non-linear Y axis) with some method of mapping to these external hosts (grouping into AS, etc.). The objective is to quickly associate context about threats within observed flows or logged events.”~~
7. ~~“As an {analyst, investigator, security operator}, I would like to be able to step through large volumes of output records in a manner that reduces the set of remaining items as quickly as possible. I would like to see related entries visually identified as being part of a common set, and have the ability to select one representative entry, tag it, categorize it as being benign or malicious, then filtering all of the related records out so as to focus on categorizing the remaining records. If the system can remember the tags and automatically apply them when similar records are seen in the future, it will be easier to identify new unknown records that require analytic scrutiny.”~~
8. ~~“As a system administrator, I would like to have a picture of the operational state of all of the system components that make up DIMS (and related underlying SIEM, etc.) This will allow me to quickly diagnose outages in dependent sub-systems that cause the system as a whole to not function as expected. The less time that it takes me to diagnose the trouble and remediate, the better.”~~
9. ~~“As a system administrator, I would like to be able to update or reconfigure DIMS subsystem components from a central location (rather than having to log in to each system and copy/edit files by hand). I would like to be assured that those changes are applied uniformly across all subsystem components, and that I have a mechanism to back out to a previous running state if need be to maintain uptime.”~~
10. ~~“As a {system administrator, security operator}, I would like to know that the DIMS system components are being monitored for attempted access by any of the same malicious actors who are seen to be threatening my constituent users. It is only natural to assume that an attack on any participant site could lead to discovery of the security monitoring system and for that system to be attacked as well, so the system should be monitoring itself using the same cross-organizational correlation features as are used internally.”~~
11. ~~“As a system administrator, I would like to be able to deal with a breach of the security system in a tactical way. If a user is found to have had a compromise of their account, all access to that user should be disabled uniformly across all system components via the single-signon authentication subsystem. All cryptograph keys should also be revoked. Once the user has been informed and the computer systems they use cleaned, all cryptographic keys, certificates, and password should be updated and re-issued”~~
12. ~~“As a user of the system, I would like to see the status of any asynchronous queries or report generation requests I have made. It is reasonable for a search through the entire history of billions of events to take some time to complete, but I would like to be able to tell approximately how long I will have to wait. Ideally, the system would keep track of previous requests, the time span and complexity of filtering applied, and to provide a time estimate when a new query is being formulated so as to guide me in deciding what I really need to ask for to get an answer in the time frame I am faced with at the moment.”~~
13. ~~“As an {analyst, security operator}, I would like to have links to detailed analyses and reports that are available in public sources when a query I have made results in identifying known malware or malicious actors. This way I can more quickly come up to speed on what is (or is not) known about the threat behind the indicators or observables I am dealing with.”~~
14. ~~“As a {system administrator, security operator, network operator}, I would like to have links to Course of Action steps related to the threats that I identify using the DIMS system. This allows me to not only inform owners or compromised assets that have been identified by the system, but to also give them information about what they need to do, in what order they should take steps, and when/how to preserve evidence in the event that there is criminal investigation ongoing.”~~
15. ~~“As a user of the DIMS system, I would like the ability to (at any point in time during analysis of an incident or while viewing the situation associated with threats across the user population) produce an anonymized version of the output I am looking at so as to be able to share it with outside entities. The system should anonymize and filter the data according to the policies set by the entities that provided the underlying data, and I should be able to determine the policy for sharing of information (by clearly seeing its tagged TLP sensitivity level). Reports should similarly be tagged appropriately with TLP for the sensitivity level of the aggregate document.”~~
16. ~~“As a {system administrator, security operator}, I would like to be able to link indicators and observables that come in at the network level (e.g., IP addresses, domain names, URLs) to observables at the host level (e.g., Registry Keys and values, file names, cryptographic hashes of files) and search for those observables to confirm or refute assertions that computers under my authority have been compromised. If I get confirmation, I would then like to preserve evidence and maintain chain of custody for that evidence as easily and quickly as possible.”~~
17. ~~“As an {analyst, security operator} I would like to be able to start an analysis and annotate data files as I go through the analysis process, trying to derive meaning from what I am seeing in the data, and being able to (at any time seems appropriate) create a reference to the current data set(s) and my view of them so I can pass this reference identifier to another analyst, a CISO, or an investigator, to allow them to take a look at what I am seeing and provide their input. For example, if someone reports a DoS attack directed at SLTT government, and my analysis confirms that such an act can be seen in the PRISEM population, I would like to provide my observations to someone to help investigate targeting, etc., in order to develop a better picture of what is happening. If the result is a determination that a SITREP should be developed and information passed along to federal law enforcement, the updated annotated body of data can then be assembled into a SITREP (using a “break” or “step” reporting format, including both cleartext and anonymized versions for sharing with outside groups) and passed along with little added effort.”~~

|  |  |  |
| --- | --- | --- |
| **Version #** | **Date** | **Description** |
| 0.1 | 30 March 2014 | Initial release. |

**6.4 References**

~~J. Aldridge. Targeted Intrusion Remediation: Lessons from the Front Lines, August 2012. Black Hat USA 2012 Presentation. <https://www.mandiant.com/blog/black-hat-usa-2012-presentation-targeted-intrusion-remediation-lessons-front-lines/>~~

~~N. Beebe and J. G. Clark. A hierarchical, objectives-based framework for the digital investigations process. Digital Investigation, 2(2):147–167, 2005. <http://faculty.business.utsa.edu/nbeebe/pubs/DIP%20Framework%20Journal%20Submission%20v4%20-%20FINAL%20JDI%20author%20copy.pdf>~~

~~Bluehat1. New MAPP Initiatives, July 2013. BlueHat Blog. <http://blogs.technet.com/b/bluehat/archive/2013/07/29/new-mapp-initiatives.aspx>~~

~~J. R. Boyd (Col.). Boyd’s OODA “Loop” From “The Essence of Winning and Losing”, 2008. Available at <http://www.d-n-i.net/fcs/ppt/boyds_ooda_loop.ppt>~~

~~S. O ́. Ciardhua ́in. An Extended Model of Cybercrime Investigations. International Journal of Digital Evidence, 3(1), Summer 2004. <http://www.utica.edu/academic/institutes/ecii/publications/articles/A0B70121-FD6C-3DBA-0EA5C3E93CC575FA.pdf>~~

~~D. Dittrich. PRISEM Analyst’s Handbook, December 2013.~~

~~D. Dittrich. PRISEM System Administration Handbook, December 2013.~~

~~D. Dittrich. Advanced Incident Response Capabilities Supporting Collaborative and Cooperative Responses. Unpublished manuscript, April 2007.~~

~~D. Dittrich. On Developing Tomorrow’s “Cyber Warriors”. In Proceedings of the 12th Colloquium for Information Systems Security Education, June 2008. <http://staff.washington.edu/dittrich/misc/cisse2008-dittrich.pdf>~~

~~D. Dittrich. On the Development of Computer Network Attack Capabilities. Unpublished manuscript, February 2008. This work was performed for the National Research Council under agreement D-235- DEPS-2007-001.~~

~~D. Dittrich. The Conflicts Facing Those Responding to Cyberconflict. In USENIX ;login: vol. 34, no. 6, December 2009. <http://www.usenix.org/publications/login/2009-12/openpdfs/dittrich.pdf>~~

~~Executive Office of the President. Homeland Security Presidential Directive 7: Critical Infrastructure Identification, Prioritization, and Protection. <http://www.dhs.gov/xabout/laws/gc_1214597989952.shtm>, December 2003.~~

~~Federal Emergency Management Agency. National Response Framework, January 2008. <http://www.fema.gov/pdf/emergency/nrf/nrf-core.pdf>~~

~~W. Gragido. Understanding Indicators of Compromise (IOC) Part I, October 2012. <http://blogs.rsa.com/will-gragido/understanding-indicators-of-compromise-ioc-part-i/>~~

~~M. Hamilton and D. Dittrich. An overview of the Public Regional Information Security Event Management Project, December 2013.~~

~~E. Hutchins, M. Cloppert, and R. Amin. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. In 6th Annual International Conference on Information Warfare and Security. Lockheed Martin Corporation, December 2011. <http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf>~~

~~H. Khurana, J. Basney, M. Bakht, M. Freemon, V. Welch, and R. Butler. Palantir: A Framework for Collaborative Incident Response and Investigation. In IDtrust ’09: Proceedings of the 8th Symposium on Identity and Trust on the Internet, pages 38–51, New York, NY, USA, April 2009. ACM. <http://middleware.internet2.edu/idtrust/2009/papers/05-khurana-palantir.pdf>~~

~~R. S. C. Ieong. FORZA - Digital forensics investigation framework that incorporate legal issues. Digital Investigation, 3(Supplement-1):29–36, 2006. <http://www.dfrws.org/2006/proceedings/4-Ieong.pdf>~~

~~Mandiant. Using Indicators of Compromise to Find Evil and Fight Crime, August 2011. <http://www.us-cert.gov/GFIRST/presentations/2011/Using_Indicators_of_Compromise.pdf>~~

~~Mandiant. APT1: Exposing One of China’s Cyber Espionage Units, February 2013. <http://intelreport.mandiant.com/Mandiant_APT1_Report.pdf>~~

~~Microsoft Developer Network. Chapter 3: Workflow and Process. <http://msdn.microsoft.com/en-us/library/bb833024.aspx>~~

~~The Mitre Corporation. Standarizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX), 2012. <http://makingsecuritymeasurable.mitre.org/docs/STIX-Whitepaper.pdf>~~

~~C. Richards. Briefings - Colonel John R. Boyd, USAF, November 2009. <http://www.ausairpower.net/APA-Boyd-Papers.html>~~

1. Language used by Wes Young in his presentation on data sharing frameworks at M3AAWG 30th General Meeting in San Francisco, February 20. [↑](#footnote-ref-6)
2. <http://csirtgadgets.org/rfc/getting-started/> [↑](#footnote-ref-7)
3. [↑](#footnote-ref-12)