2013 5th International Conference on Cyber Conflict
K. Podins, J. Stinissen, M. Maybaum (Eds.)
2013 © NATO CCD COE Publications, Tallinn

Towards Improved Cyber Security Information Sharing

Requirements for a Cyber Security Data Exchange and Collaboration Infrastructure (CDXI)

Luc Dandurand
Cyber Defence and Assured
Information Sharing
NATO Communications and
Information Agency
The Hague, Netherlands

Oscar Serrano Serrano
Cyber Defence and Assured
Information Sharing
NATO Communications and
Information Agency
The Hague, Netherlands

Abstract: There is a requirement for improved information sharing and automation in the cyber security domain. Current practices and supporting technologies limit the ability of organizations to take full advantage of their staff's expertise and the trust relationships they have established with each other in their efforts to secure their communication and information systems. Limitations include the lack of interoperable standards, the absence of mechanisms to govern and control the use of sensitive information, and problems validating data quality. While centralized repositories, distribution lists and web services have been adopted in an attempt to address the requirement, the underlying needs are only partly met by these approaches, which do not deliver the required efficiency and effectiveness.

Analysis of the specific constraints applicable in the cyber security domain led to definition of the Cyber Security Data Exchange and Collaboration Infrastructure (CDXI) capability. CDXI provides a knowledge management tool for the cyber security domain whose objectives are to facilitate information sharing, enable automation, and facilitate the generation, refinement and vetting of data through burden-sharing collaboration or outsourcing. The capability is defined through a set of high-level requirements that are both necessary and sufficient. This paper describes the high-level requirements and provides a brief description of the work performed to develop the CDXI concept to date as well as planned future work.

Keywords: *Cyber security, knowledge management, data sharing, collaboration, automation*

This work was sponsored by NATO's Allied Command Transformation under the 2012 Cyber Defence Programme of Work. This document is a working paper that may not be cited as representing formally approved NCIA, ACT or NATO opinions, conclusions or recommendations, and represents the views of only the authors.

1. INTRODUCTION

Knowledge management is commonly used as an umbrella term that covers the generation, representation, storage, transfer, transformation, application, embedding, and protecting of an organization's information ([1], [2], [3]). Knowledge management has become increasingly important to various communities as the amount of information being produced has been growing exponentially in the last decades, and timely information exchange has become essential if not critical in a broad range of domains.

In the cyber security community, there is currently a strong need for the exchange of data to support the management of vulnerabilities, threats and incidents, as well as other cyber security activities. The exchanges are necessary to achieve common goals in federated environments and to exploit collaboration opportunities. Furthermore, given the speed at which cyber-attacks unfold, there is also a need to support timely decision-making and automate responses to the greatest extent possible. These two goals can be achieved only if structured and quality-assured data is available for automated processing.

Having recognized these issues in the cyber security domain, NATO's Allied Command Transformation (ACT) sponsored the NATO Communications and Information Agency to develop the concept for a Cyber Security Data Exchange and Collaboration Infrastructure (CDXI), whose objectives are to:

- Facilitate information sharing
- Enable automation
- Facilitate the generation, refinement and vetting of data through burden-sharing collaboration or outsourcing.

As part of the development of the CDXI concept, high-level requirements that must be met to achieve the above objectives in the cyber security domain have been identified. The high-level requirements, which define the capability needed by the Alliance to manage cyber security information, are described and justified in this paper.

The remainder of the paper is structured as follows. Section 2 introduces the problem associated with information sharing and automation in cyber security, and the current state of affairs. Section 3 lists and describes the high-level requirements identified. Section 4 introduces an illustrative high-level architecture, and Section 5 presents conclusions and outlines future work that is planned or recommended.

2. BACKGROUND

The INFOSEC "Hard Problems List", under the heading "Information Provenance", identifies assuring the quality of shared data by tracking its evolution as one of the most fundamental problems in information security [4]. It can be argued that the difficulty stems from the loss of metadata that occurs when information is exchanged over systems that favour general availability and re-use over integrity, quality assurance and traceability. The problem is not exclusive to the cyber security community; areas as diverse as medicine [5], genetics [6] and law enforcement [7] are also affected by this issue.

The use of ontologies for knowledge-sharing activities has long been an important research topic ([8], [9], [10]). The importance of mapping overlapping ontologies has also been highlighted [11], and research has been conducted in the area of distributed knowledge management ([12], [13]). However, cyber security organizations have traditionally addressed information-sharing using *ad hoc* solutions such as email exchange, web-based collaboration tools such as portals and wikis, shared databases, and automated feeds of data.

In the last few years, a number of standards and initiatives that facilitate cyber security information exchange have been developed and they are gaining acceptance. ENISA (European Network and Information Security Agency) is trying to support its member states by deploying the European Information Sharing and Alert System (EISAS) [14], while the MITRE Corporation has developed a number of standardized enumeration structures and languages: Common Vulnerabilities and Exposures (CVE), Common Platform Enumeration (CPE), Common Configuration Enumeration (CCE), Common Attack Pattern Enumeration and Classification (CAPEC), and the Open Vulnerability and Assessment Language (OVAL), amongst others [15]. Industry adoption of these standards as well as other relevant standards appears to be progressing well.

The 2011 X.1500 CYBEX (Cyber Security Information Exchange Framework) Recommendation of the ITU's Study Group 17 "*describes* **techniques** *for exchanging cyber security information*" [16]. The ITU-T's X.15xx series of standards includes many of the standards and techniques developed by the U.S. National Institute of Standards and Technology under the Security Content Automation Protocol (SCAP) initiative, the MITRE enumeration structures and standards previously mentioned, and standards and techniques for the actual exchange of data, for establishing trust and policy agreement between parties, and for assuring the integrity of exchanges. Finally, a number of standards produced within the Internet Engineering Task Force (IETF) are aimed at facilitating cyber security information exchange, e.g. Real-time Inter-network Defense (RID) under RFC 6545 [17] and RFC 6546 [18] and the

Incident Object Description Exchange Format (IODEF), RFC 5070 [19].

Commercial products are just beginning to incorporate the previously mentioned standardization efforts and their supporting technologies. In a 2008 review of existing security ontologies, it was argued that *“existing ontologies are not prepared for being reused and extended and the security community still needs a complete security ontology that solves these lacks and provides reusability, communication and knowledge sharing”* [20]. While a single, complete security ontology may be an unreachable goal, it is possible to make a set of ontologies interoperable, covering all aspects of security, and this would address the requirements. More recently, subject-matter experts from the RSA organization stated that,

“Data standards for describing and transmitting threat information have advanced significantly, but much progress is needed to extend existing standards and drive wider adoption in vendor solutions. [...] Threat information-sharing and collaboration programs help organizations augment their expertise and capabilities in detecting and remediating advanced threats, but most sharing programs are hindered by a heavy reliance on manually intensive, non-scalable processes and workflows.” [21].

While the development of interoperable ontologies is progressing well, a number of major challenges remain with respect to achieving effective and efficient exchange of data and automation in the cyber security domain:

- There are no mechanisms available to automate large-scale information sharing.
- Many different sources of data containing inconsistent and in some cases erroneous data exist.
- It is difficult, in some cases, to access the desired information from the large volumes of data stored on the Internet or embedded in specific products (e.g. vulnerability repositories, signatures for anti-virus products, etc.).
- Many protocols and access mechanisms are proprietary or not interoperable.
- Incompatible semantics using the same or similar words are used in different data sources covering the same topics.
- The quality of data varies and information and assurance regarding the level of quality provided is lacking.
- There is very limited support for efficient collaboration, despite the availability of subject-matter experts in a large number of organizations willing to collaborate.

- Concerns regarding the confidentiality of exchanged data in the absence of means by which redistribution can be satisfactorily controlled must be addressed.

CDXI is designed to address these challenges by providing an enterprise-level capability that facilitates information sharing, enables automation, and facilitates the generation, refinement and vetting of data through burden-sharing collaboration or outsourcing.

3. CDXI HIGH-LEVEL REQUIREMENTS

To define the capability needed to meet the objectives stated in Section 1, the problems associated with information sharing and automation in the cyber security domain were examined. As a result the challenges listed in Section 2 as well as a number of key considerations applicable to that domain were identified, which in turn led to the identification of eleven high-level requirements that the CDXI capability must meet in order to achieve its objectives. These high-level requirements are considered to be both necessary and sufficient, and are described below.

A. PROVIDE AN ADAPTABLE, SCALABLE, SECURE AND DECENTRALIZED INFRASTRUCTURE BASED ON A FREELY AVAILABLE CORE

Collecting data from a heterogeneous set of data sources, sharing some of it with partners, and supporting automated cyber security operations while exploiting collaboration and outsourcing opportunities is a daunting challenge. While many organizations have established trust relationships with each other, few are able to agree on a single system that fits every organization's specific requirements. Adaptability is therefore required so that organizations of different sizes, different types, facing different constraints and seeking different objectives can deploy CDXI in a way that meets their specific situation. The organizations that CDXI must support range from a very small, single-site company to a large multinational federated organization. In many cases, the need to exchange information will be the only common point, and mandating a fixed configuration will lead to an ineffective and inefficient solution, if not outright failure.

CDXI must be scalable, not so much for reasons of data quantity, which remains quite modest in cyber security, but rather because an "agile data model" and correlation capabilities are necessary (see requirement B), as is the need to support dissension (see requirement I). These two requirements are expected to increase the need for storage capacity. As well, CDXI components must be scalable to meet a wide range of hosting constraints and performance requirements in different deployment scenarios.

Because the increased need to share does not diminish the confidentiality, availability, and integrity requirements of the exchanged data, CDXI must also be secure. Therefore CDXI must provide flexible access controls to allow protection of the data as well as the possibility for custom workflows that will enable multi-step approval for actions affecting sensitive data. In order to allow greater exploitation of shared data while maintaining privacy requirements, CDXI must allow organizations to identify data elements that must be consistently replaced by privacy-protecting labels before being shared, as well as provide privacy-preserving query functionality. CDXI must allow organizations to contribute data anonymously. The CDXI architecture must also allow an organization to replace individual components in order to achieve a higher degree of assurance where it thinks it is necessary. Finally, organizations relying on CDXI must be able to review data exchanges in order to allow detection of security issues.

Organizations that need to exchange information with each other do not always recognize a single common centralized authority for establishing trusted channels for the exchange. Organizations must therefore be able to deploy and interconnect their own CDXI "instance" as they see fit. CDXI must provide for "knowledge exchanges" that allow organizations to offer their data to others as well as discover others' data offerings. As establishment of such knowledge exchanges is open to any organization, they will provide a way to mimic the current practice whereby organizations meet with each other in different, independent communities of interest (COI) that they control. In the service offerings published through the knowledge exchanges, data providers must be able to set the terms and conditions under which others can gain access to the offered data. A decentralized model allows COIs to emerge and subside without a central authority being aware of or needing to approve this.

By making the CDXI software freely available NATO will have access to data of improved quality that is contributed by the global security community. If there is convergence towards CDXI then a "critical mass" will be reached, at which point the monetary value of the data will far exceed the cost of implementing CDXI, which will be to NATO's benefit.

B. PROVIDE FOR THE CONTROLLED EVOLUTION OF THE SYNTAX AND SEMANTICS OF MULTIPLE INDEPENDENT DATA MODELS AND THEIR CORRELATION

In early work related to cyber security information exchange, one of the key difficulties encountered was obtaining agreement within a community to a standard data model. Over time, the situation has improved and there are now a number of standards that define data models and protocols that support information sharing and automated cyber security. However, there is no consistent use of these many standards, models and protocols, which makes information sharing, collaboration and automation difficult, particularly in the absence of mappings between existing data models. Furthermore, organizations are often compelled to use the data models (standardized or not) implemented in the commercial products they have acquired. These are sometimes not interoperable, which means additional effort is required to correlate the data across products. In some cases they are also inadequate, which means an organization must complement them in order to meet its specific needs. Thus despite the existence of standardized data models, organizations must still perform a substantial amount of effort to manage data models.

Therefore, to achieve the stated objectives in the cyber security domain, CDXI must allow organizations to implement standardized data models of their choosing via an "agile data model" that allows easy definition of new or existing data models without requiring a software development cycle. The proposed CDXI approach is to use "independent topic ontologies" (ITO) that capture each data model independently; this approach allows correlation of data elements across ITOs.

In this context, the term ontology is used as defined in [22]: *"a formal explicit specification of a shared conceptualization"*, and does not necessarily imply the use of ontological languages. From a software development point of view, an ITO can be seen as a logical container for a set of classes and relationships with associated attributes. An ITO is therefore a data model covering a defined domain of interest, and CDXI does not limit the size, scope, or depth of ITOs in any way. Each instance of a class or relationship must have a globally unique identifier that can be used to correlate data across available ITOs, subject to access controls.

The use of an agile data model implies that CDXI can support any data model and does not try to force a particular one on an organization or community of interest. The latter condition is necessary because defining a single, standardized ontology that covers the entire cyber security domain is not practical. Moreover, the agile data model allows CDXI users to easily implement new data models for which no current standards exist, as is the case for enterprise security models [23][1] and network security policies [24]. Sharing ITOs while they are in the process of being defined, and collaborative refinement of them, may also facilitate standardization efforts [25]. The agile data model allows existing data sources to be brought into CDXI relatively easily, thus taking advantage of prior investments. CDXI's support for correlation across ITOs will facilitate interoperability by allowing organizations to compose data queries that exploit ITOs that are covering the same topics at the same granularity. In a large organization, this work would be done by ontologists for the benefit of end-users.

Finally, controlled evolution of ITOs must be possible. The CDXI objective of enabling automation will be achieved when organizations use data obtained through CDXI in cyber security applications. However, the agile data model allows users to modify existing ITOs as domain knowledge evolves by adding, modifying or deleting classes, relationships or attributes and by modifying the ITO syntax or semantics. Allowing ITOs to be freely changed would give rise to problems because organizations would have to revise their cyber security applications after every ITO change to accommodate the new syntax and semantics. By enforcing comprehensive version control of ITO definitions, CDXI will allow data providers to modify their data models and data consumers to adjust their automated applications independently and at their own pace.

C. SECURELY STORE BOTH SHARED AND PRIVATE DATA

CDXI must allow an organization to store cyber security data that can be either kept private or shared with other organizations. When user data is identified as being private, CDXI must ensure that the data is never made available outside of the organization. This will allow organizations to exploit the agile data model and correlation capabilities in CDXI to store organization-specific data that is never intended to be shared, link it to data obtained from external data sources, and use the correlated information to support automated applications.

D. PROVIDE FOR CUSTOMIZABLE, CONTROLLED MULTILATERAL SHARING

Since most cyber security organizations need to interact with a range of partners for different information exchanges, CDXI must provide mechanisms that allow customizable, controlled multilateral sharing. Organizations must be able to create and manage information-sharing relationships with their partners using the

[1] Although Anderson provides an enterprise security model, it is not a standardized model.

security protocols most appropriate for each individual case. All exchange of data must be through “Information Exchange Policies” (IEP) set up by the organizations themselves. It must be possible to define any number of IEPs in order to meet the various exchange requirements.

CDXI must allow for the definition of any number of “communication channels” that implement encryption, authentication and authorization mechanisms. CDXI must allow organizations to freely associate IEPs with communication channels in order to select the most appropriate means over which a particular exchange can take place. The decision to share can be applied to entire ITOs or sub-elements of ITOs, and to all of the data or to individual data records. It must be possible to define a custom workflow for activating an IEP, as well as for authorizing the sharing of individual records in an IEP when needed.

Therefore when two or more organizations agree to exchange information with each other, they must select the applicable ITOs (thus choosing a particular ontology that describes the syntax and semantics of the data to be exchanged), identify the parties to the exchange, capture the terms and conditions under which the exchange will take place, and select the communication channels that CDXI will use to execute the exchange. This approach decouples the technical details of how to create a secure tunnel for the information over possibly insecure networks from the details related to fine-grained access controls and the terms and conditions of the exchange, such as the intellectual property rights, rights to further distribute the data and uses that can be made of it. IEPs must also allow organizations to choose a suitable accounting mechanism to support commercial activities (see requirement K). Finally IEPs must also indicate whether or not recipients can edit the exchanged data; such authorization would be given to support collaboration or outsourcing.

All exchanges must be logged and made available for audit review. Furthermore, exchanged data must always remain associated with the IEP under which it was received. CDXI must enforce the terms and conditions set forth in IEPs, and specifically the condition for redistribution of the data.

E. ENABLE THE EXCHANGE OF DATA ACROSS NON-CONNECTED DOMAINS

CDXI is expected to be deployed in various CISs that may not be directly interconnected (e.g. highly secure networks). CDXI must provide mechanisms to facilitate exchange across these “air gaps”. Such mechanisms must provide for the auditing of the transfers in a manner that would allow for the detection of sensitive information leakage or the introduction of malicious code. The exchange of data across non-connected domains must facilitate the efficient reconciliation of conflicting changes concurrently made in all CDXI deployments participating in an exchange of data.

F. PROVIDE HUMAN AND MACHINE INTERFACES

A key requirement of CDXI is that it provide both human-specific and machine-specific interfaces. CDXI must provide a set of graphical user interfaces (GUI) that facilitate human interaction with the data, and a set of application programming interfaces (API) that facilitate machine interaction with the data. These interfaces must be well adapted to the needs of these very different types of user.

G. PROVIDE COLLABORATION TOOLS THAT ENABLE BURDEN SHARING FOR THE GENERATION, REFINEMENT, AND VETTING OF DATA

One of the objectives of CDXI is to facilitate burden-sharing collaboration and/or outsourcing for the generation, refinement and vetting of cyber security data. While a number of organizations have established a sufficient degree of trust between each other to allow for collaboration, current information systems do not provide sufficient support to make collaboration an effective and efficient approach to generating, refining, and vetting of data, and in many cases the associated level of effort for collaboration is simply too high. Where collaboration does take place, it is often inefficient due to the absence of a facilitating system. CDXI must therefore provide tools that will address this issue.

As a minimum, CDXI must provide a timely threaded discussion mechanism that can be used to annotate different data elements. As well, it must provide a chat facility that is subject to access controls and IEPs and that provides a capability to quickly establish a shared context to support discussing a particular data element.

H. PROVIDE CUSTOMIZABLE QUALITY-CONTROL PROCESSES

CDXI will be used to aggregate and transform information from many sources to feed decision-making and automated processes. Inaccurate information could cause a business process to fail, resulting in undesired effects that can vary greatly in significance. To successfully enable automation in cyber security, CDXI must provide the means to assure the quality of the data it provides.

Quality assurance (QA) within CDXI refers to the planned and systematic activities that ensure that the data in the CDXI system meets the quality requirements specific to its intended use. QA is achieved through the application of custom quality-control processes (QCP) that are defined by users and partly managed within CDXI. Because CDXI data can be re-used for many different purposes, ITOs, QCPs and quality requirements are associated to the use that will be made of the data, based on the concept of "curation". The curation identifies the ITOs that are needed to support an automated application as well as the QCPs that will be used to filter the data to provide only that data that meets the required quality. This allows QCPs to be re-used for different ITOs where applicable, and for ITOs to be re-used for different purposes (i.e. for different curations) even if those purposes have different quality requirements. QCPs can also be included in IEPs to ensure that data exchanged with external parties meets the desired quality requirement. In addition, CDXI must allow organizations to exchange QCPs and associated information so that QCPs can be re-used, outsourced or performed in a collaborative fashion.

I. EXPOSE DISSENSION TO REACH CONSENSUS

The fact that most databases are designed to hold a single value for each attribute of a data element, in other words only "one truth", means that users cannot express disagreement about a value except by changing the value in the database (assuming they have the necessary privileges to do so), which would then change the value for all users. Since most common data repositories have no means to expose dissension about attribute values, errors and inaccuracies recognized by users remain hidden, which limits an organization's ability to improve the data upon which it relies for operations.

CDXI must therefore expose dissension by allowing multiple possible values to be shown for each field ("multiple truths") in order to allow users to see that there is disagreement and eventually either reach consensus on which value is correct or agree to disagree. Data managers in the organizations participating in an exchange of data would have the ability to see all proposed values for an attribute and to select the one they consider to be correct for their organization, or choose to have CDXI always use the most recently entered value if they do not have the expertise to decide themselves for a particular type of data. Finally, CDXI must also allow users to easily correct detected errors and inaccuracies by allowing "divergent values" to be used locally within an organization so that automated processing can proceed with the corrected data. This functionality can also help detect and address mischievous activities directed at data sources by malicious users.

J. SUPPORT CONTINUOUS AVAILABILITY OF DATA

CDXI must meet availability requirements, even in the presence of cyber-attacks. It cannot be assumed that an organization will always have external connectivity to obtain cyber security data. CDXI must therefore allow an organization to choose to hold a local copy of selected data previously exchanged so that it can continue to use that data after disconnecting all external communication links (subject to the terms and conditions set forth in IEPs).

K. ENABLE COMMERCIAL ACTIVITIES

The private sector will be more motivated to use CDXI if it provides accounting models and functionality for selling data or data-related services. This in turn will lead to better-quality data for CDXI and thus for NATO.

CDXI must therefore provide various accounting models for the usage of data, and the mechanisms must allow vendors of data and data services to control the dissemination of data exchanged under the terms of a commercial contract. Organizations must be able to sell any data element, such as content (ITO data), and the application of quality control processes, as well as professional services related to the management and refinement of CDXI data, such as assistance in defining ITOs, correlation and translation.

If commercial activities are supported, organizations that use CDXI will be able to make use of industry's extensive resources and expertise to obtain the data they require at a cost determined by market forces, and as a result NATO will have access to the best available data.

4. HIGH-LEVEL ARCHITECTURE

To illustrate an implementation approach that could address the adaptability requirement, a high-level architecture was developed. It consists of two major building blocks: the CDXI Administrative Domain (CAD) and the CDXI Security Domain (CSD). The CAD encompasses the set of CDXI components deployed by a single organization and managed through a coherent set of administrative and high-level security policies. The CSD groups the set of CDXI components deployed in a particular network that share a common set of security services and settings and that can be directly connected to each other. Any number of CSDs can be defined within a CAD, but a CSD can belong to only one CAD. In general, a CAD will correspond to an organization, but in some cases, a larger organization may wish to deploy more than one CAD to adapt the implementation of CDXI to its organizational structure and business practices.

The CAD is used to provide coherence in the management of CSDs and to define the IEPs used by CSDs for the exchange of data. Some aspects of the management of CSDs can be centralized at the CAD (e.g. management of user accounts) or performed using management interfaces in each CSD.

In addition, the high-level architecture defines the CDXI Administrative and Security Boundary Managers (CABM and CSBM respectively). These components are used to control communications between domains. The role of the CSBM is to ensure that no data is exchanged between CSDs without a valid IEP and to take care of pulling and pushing data according to the terms of the applicable IEPs using the specified communication channel. The role of the CABM is to ensure that no data is exchanged between CADs without a valid IEP, to take care of pulling and pushing data according to the terms of the applicable IEPs using the specified communication channel, and to manage the interactions that occur with the knowledge exchanges. Both types of boundary manager provide buffering of data and a reliable exchange mechanism. Multiple instances could be deployed to provide scalability and high availability via load balancing.

5. CONCLUSIONS AND FUTURE WORK

The cyber security community requires tools to facilitate information sharing and automation, and the tools must allow for burden-sharing collaboration and outsourcing in the management of cyber security data. To address these needs, a knowledge management capability called the Cyber Security Data Exchange and Collaboration Infrastructure (CDXI) was defined. In the light of characteristics specific to the cyber security domain, the high-level requirements that must be met for the capability to achieve its objectives were identified.

As well, limited-depth investigative prototyping activities were conducted to determine which technologies are most suitable for implementing the agile data model. Possible options identified to date for implementing the agile data model include:

- Special constructs using relational database management systems (RDBMS):
    - Allowing the CDXI application to use the SQL Data Description Language (DDL) (e.g. CREATE, ALTER, DROP statements)
    - Use of an Entity, Attribute, Value (EAV) schema, which allows definition of the data model using only SQL Data Manipulation Language (DML).
    - Anchor modeling, which describes the data at high normalization levels using a graph notation based on anchors, attributes, ties, and knots (an approach similar to EAV).

- Triplestores for RDF (Resource Description Framework) or OWL (Web Ontology Language) as used for the development of semantic webs.
- Non-SQL solutions or schema-less databases such as MongoDB.

The prototyping activities conducted to date for the agile data model were based on the first two special constructs above and the use of a conventional RDBMS. The first prototype activity was based on the use of DDL, while the second was similar to the one introduced in [26] for the definition of genome ontologies. While the findings of these limited-depth trials suggest that the dynamic creation of ITOs using DDL would be a better approach than the use of an EAV schema, further work is required to confirm this and to assess the other approaches as well. At the moment the expectation is that the final implementation of an agile data model will likely not rest on a single solution but on a combination of the technologies mentioned above.

An initial proof-of-concept design was also developed. This work helped identify lower-level requirements and technical approaches for the implementation of CDXI, and is documented in NATO technical reports.

ACT has sponsored validation of the CDXI capability defined in this paper through an engagement with NATO stakeholders and subject-matter experts in NATO nations, industry, and academia, as well as a review of existing prototypes and capabilities that provide similar functionality. If the initial feedback indicates that it is necessary, the validation activity may be extended to include the development of a proof-of-concept. Once the CDXI capability is validated, options available for the procurement of an operational, production-grade CDXI will be considered. In parallel, further work will likely be conducted to refine specifications, identify minimum performance requirements, and investigate the suitability of existing technologies and standards in order to support the procurement process.

REFERENCES

[1] Hedlund, G. (1994). A Model of Knowledge Management and the N-Form Corporation. *Strategic Management Journal, 15*, 73-90.

[2] Alavi, M., & Leidner, D. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly, 25:1*, 107-136.

[3] Tyugu, E. (1993). Large Engineering Knowledge Bases. *Artificial Intelligence in Engineering*, 8(4), 265-270.

[4] INFOSEC Research Council. (2006). Hard Problems List. Washington DC, Cyber Security and Information Assurance Interagency Working Group (CSIA IWG).

[5] Nardon, F., & Moura, L. (2004). Knowledge sharing and information integration in healthcare using ontologies and deductive databases. *Medinfo*, 62-66.

[6] Camon, E., Magrane, M., Barrell, D., Lee, V., Dimmer, E., Maslen, J., Binns, D., Harte, N., Lopez, R., & Apweiler, R. (2004). The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Research, 32*(suppl 1), D262-D266.

[7] Ferrara, L., Mårtenson, C., Svenson, P., Svensson, P., Hidalgo, J., Molano, A., & Madsen, A. (2008). Integrating data sources and network analysis tools to support the fight against organized crime. *Intelligence and Security Informatics*, 171-182.

[8] Fulton, J. (1992). Technical report on the semantic unification meta-model – Standards working document ISO TC184/SC4/WG3 N103. Seattle, IGES/PDES Organization, Dictionary/Methodology Committee.

[9] Allen, J., & Lehrer, N. (1992). DARPA/Rome Laboratory Planning and Scheduling Initiative Knowledge Representation Specification Language (KRSL), Version 2.0.1 Reference Manual. ISX Corporation.

[10] Gruber, T. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human-Computer Studies, 44*(5-6), 907-928.

[11] Kalfoglou, Y., & Schorlemmer, M. (2003). Ontology mapping: the state of the art. *The Knowledge Engineering Review, 18*(1), 1-31.

[12] Bonifacio, P., Bouquet, P., Mameli, G., & Nori, M. (2004). Peer-Mediated Distributed Knowledge Management. *Lecture Notes in Computer Science, 2926*, 31-47.

[13] Ehrig, M., Tempich, C., Broekstra, J., van Harmelen, F., Sabou, M., Siebes, R., Staab, S., & Stuckenschmidt, H. (2003). SWAP: Ontology-based Knowledge Management with Peer-to-Peer. *Workshop ontologiebasiertes Wissensmanagement* (pp. 17-20). Lucern: Gesellschaft für Informatik, Lecture Notes in Informatics (LNI), P-28, Bonn.

[14] ENISA. (2011). EISAS (enhanced) report on implementation.

[15] Martin, R. (2008). Making Security Measurable and Manageable. *Proceedings of the IEEE Military Communications Conference*, 19. San Diego.

[16] ITU-T. (2011). Overview of cybersecurity information exchange. Geneva, ITU-T.

[17] Internet Engineering Task Force Request for Comments 6545, “Real-time Inter-network Defense (RID)”, K. Moriarty, IETF, April 2012.

[18] Internet Engineering Task Force Request for Comments 6546, “Transport of Real-time Inter-network Defense (RID) Messages over HTTP/TLS”, B. Trammell, IETF, April 2012.

[19] Internet Engineering Task Force Request for Comments 5070, “The Incident Object Description Exchange Format”, R. Danyliw, J. Meijer, & Y. Demchenko, IETF, December 2007.

[20] Blanco, C., Lasheras, J., Valencia-García, R., Fernández-Medina, E., Toval, A., & Piattini, M. (2008). A Systematic Review and Comparison of Security Ontologies. *The Third International Conference on Availability, Reliability and Security*, 813-820.

[21] Hartman, B. M. (2012). RSA Security Brief February 2012 – Breaking Down Barriers to Collaboration in the Fight Against Advanced Threats.

[22] Gruber, T. (1993). A translation approach to portable ontologies. *Knowledge Acquisition, 5*(2), 199-220.

[23] Anderson, E., Choobineh, J., & Grimaila, M. (2005). An Enterprise Level Security Requirements Specification Model. *38th Hawaii International Conference on System Sciences.* IEEE Computer Society.

[24] Cuppens, F., Cuppens-Boulahia, N., Sans, T., & Miège, A. (2005). A Formal Approach to Specify and Deploy a Network Security Policy. *International Federation for Information Processing*, *173*, 203-218.

[25] Sofia Pinto, H., Staab, S., & Tempich, C. (2004). DILIGENT: Towards a fine-grained methodology for DIstributed, Loosely controlled and evolvInG Engineering of oNTologies. *Proceedings of the 16th European Conference on Artificial Intelligence (ECAI 2004)*, 393397. Valencia, Springer.

[26] Nadkarni, P., Marenco, L., Chen, R., Skoufos, E., Shepherd, G., & Miller, P. (1999). Organization of Heterogeneous Scientific Data Using the EAV/CR Representation. *Journal of the American Medical Informatics Association*, 6(6), 478–493.