Pacific Northwest
NATIONAL LABORATORY

PNNL-36023

Countering Weapons of Mass Destruction Office (CWMD) Data Categorization Study

Chemical, Biological, Radiological, and Nuclear (CBRN) Detection Device Data

May 2024

Starr Abdelhadi
Mark D Watson
Sumit Purohit
Quinn J Wright-Mockler
Beau R Morton
Penny L McKenzie

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-36023

Countering Weapons of Mass Destruction Office (CWMD) Data Categorization Study

Chemical, Biological, Radiological, and Nuclear (CBRN) Detection Device Data

May 2024

Starr Abdelhadi
Mark D Watson
Sumit Purohit
Quinn J Wright-Mockler
Beau R Morton
Penny L McKenzie

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830.

Pacific Northwest National Laboratory
Richland, Washington 99354

Summary

Pacific Northwest National Laboratory (PNNL) seeks to address critical questions related to chemical, biological, radiological, and nuclear (CBRN) detection devices. This research aims to enhance the security and understanding of these devices by investigating various aspects of their identification, communication, and functionality. The primary focus is on network security, malware detection, device identification, and intelligence gathering.

CBRN data can be categorized in various ways depending on the purpose of CBRN detection devices and the specific context of the applications for analysis. Criteria that can be used to assist in this effort include but are not limited to data type, data protocol, source/destination, application, time, security, and content. This study will inform additional paths for data classification, data profiling, data mapping, and data modeling. This will help the Countering Weapons of Mass Destruction Office (CWMD) better understand their data and make informed decisions based on the insights gained from this study and their application. The CBRN Data Categorization study includes the identification of three different CBRN detection devices with unique characteristics for assessing and analyzing the data that is being produced by and transmitted from these devices.

Data categorization is the process of grouping or classifying data based on certain characteristics, features, or attributes. The purpose of categorizing data is to make it easier to understand, manage, analyze, and make decisions based on the information that the data contains. Data categorization is a crucial step in data management and analysis. It is important to choose appropriate criteria for categorization to ensure that the data is useful and meaningful for the intended purposes.

CBRN detection devices and their operational data must be properly categorized, identified, and secured throughout their life cycle to reduce the threat surface to the mission of the CWMD program in critical environments. Additionally, the categorization of data obtained by CBRN detection devices must support structured data standards across disparate ownership platforms (e.g., state, local, tribal, territorial [SLTT], etc.). PNNL’s research seeks to provide valuable insights and solutions to enhance the security and understanding of CBRN detection devices. Addressing these issues is a crucial step towards mitigating the risks associated with these devices and safeguarding critical networks and infrastructures.

Acknowledgments

Pacific Northwest National Laboratory would like to acknowledge the T&E Cyber Team of the Countering Weapons of Mass Destruction Office for their dedication to securing the nation and their forward-leaning approach to cybersecurity throughout the data categorization process, of which device data categorization remains a foundational element of the secure system life cycle.

Acronyms and Abbreviations

AD	Advertising Data
ALF	Acquisition Lifecycle Framework
BLE	Bluetooth Low Energy
CBRN	chemical, biological, radiological, and nuclear
CBRNDR	Chemical, Biological, Radiological and Nuclear Defense Repository
CIA	Confidentiality, Integrity, Availability
CNSS	Committee on National Security Systems
CNSSI	Committee on National Security Systems Instruction
CWMD	Countering Weapons of Mass Destruction Office
DCM	data categorization model
DHS	U.S. Department of Homeland Security
DMAMC	Data Mining Analysis and Modeling Cell
DSC	Device System Categorization
DT	decision tree
FIPS 199	Federal Information Processing Standard Publication 199
HTTP	Hypertext Transfer Protocol
IL	Instrument Library
IoT	Internet of Things
IP	Internet protocol
IR	Interagency Report
MAC	Media Access Control
MS	Milestone
NCCoE	National Cybersecurity Center of Excellence
NIEM	National Information Exchange Model
NIST	National Institute of Standards and Technology
OSINT	open-source intelligence
PM	program manager
PNNL	Pacific Northwest National Laboratory
RMF	Risk Management Framework
RPM	radiation portal monitor
S&T	Science and Technology
SDP	Solution Development Process
SELC	Systems Engineering Life Cycle
SLTT	state, local, tribal, territorial
SP	Special Publication
SSP	System Security Plan

SUT	system under test
T&E	Test and Evaluation
UI	user interface

Contents

Figures

Tables

1.0 Introduction

When chemical, biological, radiological, and nuclear (CBRN) data lacks proper categorization, it can invite a host of cyber implications. Without clear classification, essential CBRN data becomes vulnerable to unauthorized access, manipulation, or even destruction by malicious actors. This jeopardizes the reliability and integrity of CBRN detection capabilities, potentially leading to manipulated data, data losses, and misconfigurations of sensing capabilities.

Misclassified CBRN data also poses a significant challenge to effective decision-making and resource allocation. Without accurate categorization, the Countering Weapons of Mass Destruction Office (CWMD) will struggle to identify and prioritize critical data sharing across organizations, hindering the ability to respond swiftly and effectively to emerging threats or misleading data points. Moreover, regulatory compliance becomes a daunting task, as data protection requirements hinge on precise classification and handling protocols.

The lack of proper categorization worsens the complexities of data governance, impeding efforts to enforce access controls, maintain data quality, and ensure accountability across CWMD and partner organizations. This undermines operational efficiency and erodes trust and confidence in CWMD's ability to protect sensitive information.

The misclassification of operational data opens the door to a myriad of cyber risks, undermining CWMD's resilience, regulatory compliance, and organizational trust. It underscores the urgent need for robust data management practices that prioritize accurate classification, effective governance, and proactive security measures.

This study will inform additional paths for data classification, data profiling, data mapping, and data modeling. This will help CWMD better understand their data and make informed decisions based on the insights gained from this study and its application.

1.1 Purpose

The U.S. Department of Homeland Security (DHS) CWMD depends on an interconnected complex environment for their systems and networks. These environments can be inherently vulnerable. Device system assets, including data and connection types, need to be protected from identified security threats and vulnerabilities that could impact the CWMD mission. To protect device system assets and their information, the cybersecurity categorization of CWMD device system data will assist in the development and fielding of more secure and resilient systems. The data categorization performed by CWMD supports the life cycle of CWMD mission priorities to design, implement, and maintain a device system security program that aligns with DHS Cybersecurity and Cyber Resilience processes.

Data categorization is the process of grouping or classifying data based on certain characteristics, features, or attributes. The purpose of categorizing data is to make it easier to understand, manage, analyze, and make decisions based on the information that the data contains. Data categorization is a crucial step in data management and analysis. It is important to choose appropriate criteria for categorization to ensure that the data is useful and meaningful for the intended purposes.

The operational data of CBRN detection devices must be properly categorized, identified, and secured throughout these devices' life cycles to reduce the threat surface to the mission of the

CWMD program in critical environments. CBRN data that the device is capturing and/or disseminating, including data sensitivities and thresholds, must be properly categorized to support structured data standards across disparate ownership platforms (e.g., state, local, tribal, territorial [SLTT], etc.).

CBRN data can be categorized in various ways depending on the purpose of the device and the specific context of the applications for analysis. Criteria that can be used to assist in this effort include but are not limited to data type, data protocol, source/destination, application, time, security, and content.

1.2 Organization of the Report

This document contains seven primary sections, including this introductory section.

- Section 2.0 provides an up-front perspective on the data gathering challenges and use cases needed for development.
- Section 3.0 lays out the implementation of cybersecurity categorization for CBRN detection device systems across the acquisition life cycle.
- Section 4.0 provides an approach to the CWMD categorization processes and methods essential to support the application of policy and guidance to establish data categorization priorities based on a series of factors and dependencies related to the specific needs and data categorization objectives of CWMD and their organizational partners.
- Section 5.0 outlines the next steps identified during the categorization process.
- Section 6.0 provides a list of references and applicable sources.
- Appendix A provides additional guidance and policies that align with CWMD missions.

2.0 CBRN Data Challenges

To comprehensively address the challenges associated with capturing and categorizing CBRN data, several key factors need consideration. These include the sensitivity of the data, the methods of capture and storage, its intended use and audience, and the necessary controls to safeguard it effectively.

In the realm of countering weapons of mass destruction, accurately classifying unstructured CBRN data presents a substantial hurdle. The majority of CBRN data lacks predefined formats, complicating conventional classification approaches. Moreover, the diverse nature of dispersed datasets across various sectors (e.g., SLTT, organizational, governmental, etc.) introduces additional complexities, necessitating a need-to-know basis for swift and precise responses to cybersecurity incidents involving CBRN data.

Data collected from devices (such as network, operational, location, and user data) may harbor additional sensitivities, varying in importance across organizations. Reviewing and refining these datasets often involves consultation with stakeholders to ensure their relevance and representativeness within their respective domains. There is a lack of a centralized repository for holding relevant device data for review.

The absence of specific policies, compliance regulations, and recommendations tailored to CBRN data further exacerbates the challenge. Efforts to bolster data security must undergo peer review to ensure accuracy, with stakeholder engagement integral to the development of data categorization models (DCMs).

Effective information sharing is critical for precise cyber incident response endeavors. However, existing response procedures vary widely among organizations, leading to potential delays and inaccuracies, particularly in incidents involving CBRN data. The absence of guidance for CBRN-related incidents can result in disparate response protocols and data inconsistencies, undermining overall response efficacy. Therefore, there is no centralized cyber incident response sharing mechanism for CBRN data.

2.1 Data User Use Case

CBRN device users will collect raw data based on attributes and parameters that may be useful for classification later in the data life cycle. The next step they will take is to define the patterns and criteria for classifying the data assets. Collaboration efforts with the other SLTT stakeholders and organizations will better determine the categorization processes with interviews, additional document reviews, the use of scanning tools, and classification workflows. This has ensured that the data is handled in accordance with policy guidelines and regulatory requirements. Once the datasets have been properly categorized, they will be categorized with the appropriate security controls and additional protection mechanisms. This ongoing process will require regular reviews of policies and regulations as they are updated to ensure security compliance.

2.2 Data Security Use Case

A security audit may uncover inconsistencies in the categorization of CBRN data within a database. This miscategorized data could compromise national security and endanger lives. This scenario could include the disclosure of mislabeled data, with potential hazardous

materials categorized at a lower security level. The data would have to be isolated and restricted to authorized personnel only. The data would also have to be painstakingly reprocessed and reevaluated to ensure that it was accurately classified according to its level of sensitivity and potential risk. Numerous challenges to properly categorizing data would occur with poor documentation and the lack of sensitivity context.

3.0 Data Gathering Processes and Methods

Gathering CBRN data supports various purposes and critical functions for decision-makers, including assessments (e.g., risks, threats, vulnerabilities), planning (e.g., incident and emergency response activities), and policy development. Data gathering processes and methods build awareness to help decision-makers better understand, respond to, and mitigate potential risks. CBRN data gathering processes and methods stem from a variety of resources and applications including open-source intelligence (OSINT), specialized device equipment and systems, sensor networks, remote sensing, environmental monitoring, surveillance systems, biological sampling, simulation and modeling, historical data analysis, and stakeholder engagement and collaboration. Using a combination of specialized instruments and techniques to gather CBRN detection device data from handheld devices (e.g., portable units), detector networks (i.e., deployed detectors for continuous monitoring), sampling equipment, wireless sensor networks, testing and evaluation, maintenance and calibration, and other disparate sources requires complex training and tradecraft to create data fusion, analysis, and integration for a more complete picture of the threat landscape.

CWMD mission owners and partners benefit from shared situational awareness and a common operating picture of contested CBRN operating environments to effectively share mission-relevant data in real time. To address the underlying capabilities of accelerated mission awareness of CBRN system performance and rapid threat detection and the deterrence of nation-state actors, tactics, techniques, and procedures, it is important to acknowledge the interconnected nature of CBRN mission partners and stakeholders, system-of-systems components and services, device dependencies and limitations, and related operational data. Pacific Northwest National Laboratory (PNNL) has developed the CBRN Ecosystem to express various data components and dependencies related to the operational environment from end users to disparate edge CBRN devices and the resident data that may traverse across these data components. In addition, instrument and data repositories for storing and accessing data artifacts captured by data collection systems are further discussed below.

3.1 CBRN Ecosystem

The CBRN Ecosystem in Figure 1 represents the complex nature of the interconnected systems and data components that coexist in dynamic interoperable environments. An ecosystem is unique from other system counterparts in that all interoperable elements, including the data produced, transmitted, or stored, share a common point of interconnection at the highest level. Systems, however, are constrained to subsystem dependences and device-specific components at lower levels and do not necessarily share higher points of interconnection in their operating environment, including shared information network resources within the CBRN Ecosystem. For example, CBRN detection devices do not provide inherent operational data in real time to the higher points of interconnection in the CBRN Ecosystem and are not exclusively connected with each other within the same environment. However, these devices can provide operational data and immediate value within the operating environment they support. Another key element of the CBRN Ecosystem is establishing a network of CBRN detection devices that provides the capability for data gathering, delivery, and accessibility to end users that can process and analyze multilayered data with appropriate tools and platforms.

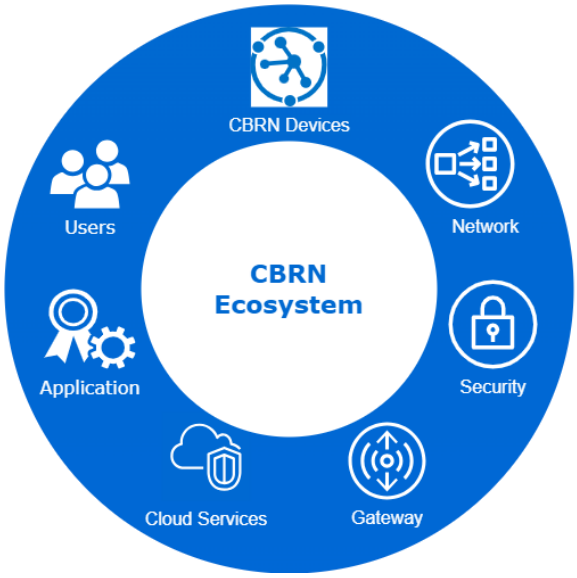

Figure 1. Data components of the CBRN Ecosystem.

The CBRN Ecosystem in Figure 1 consists of seven components: CBRN Devices, Network, Security, Gateway, Cloud Services, Application, and Users. At a high level, some or all of these components are interconnected, whereby CBRN detection devices gather data at their end point and securely transfer it through the network to a gateway connected through a demilitarized zone, where the data is further compressed and sent to the cloud for further analyses. The cloud provides data repositories and services that support data enrichment, enumeration, and visualization using preferred applications that provide end users with data insights. Not all CBRN ecosystems may be alike, nor do they require all data components; however, these ecosystems are constrained and likewise dependent on the function of their end points, namely CBRN detection devices. An overview of the CBRN Ecosystem components is summarized in **Error! Reference source not found.**, with a brief description for each component.

Table 1. Summary of Components in the CBRN Ecosystem.

Data Component	Description
CBRN Detection Devices	• CBRN detection devices are physical devices that interact with the environment. The two types are sensors and actuators: ○ Sensors – devices that are supposed to gather information about the environment and measure its physical parameters like temperature, motion, people flow, etc. In other words, sensors convert physical phenomena into a digital form. ○ Actuators – devices that perform a physical action on things after they get such a command.
Network	• The network is responsible for the communication within a CBRN ecosystem between smart device(s), a gateway, and the cloud.
Security	• The security component is responsible for access control to the CBRN network, the security of data transfers, data leakage prevention, and scanning for malicious software. The security component is presented by firmware and software from security providers.

Data Component	Description
Gateway	• A gateway is a physical device that allows for data streams from sensors to the cloud and in the opposite direction. • A gateway performs data preprocessing before the information will be transferred to the cloud. • A gateway is not a necessary element since CBRN devices can set connections to the Internet by themselves without a gateway as an intermediary.
Cloud Services	• The cloud is a resource that is responsible for data storage, deep analysis, and management. It is a group of computers people access through the Internet to use their compute capacity for some purpose. • The cloud is enhanced by powerful analytic and visualization tools, big data algorithms, and machine learning technology.
Application	• An application is the graphical user interface that provides remote control and management of devices connected to the CBRN ecosystem.
Users	• Users are all the people who affect the CBRN ecosystem and use it for their purposes. • Users comprise people with personal CBRN detectors, researchers who use analytics from the CBRN cloud, decision-makers who use CBRN data in their operational processes, and stakeholders.

3.2 Data Repositories

Example devices and datasets will be used when designing processes to categorize operational data from CBRN detection devices. Tools such as the Data Mining Analysis and Modeling Cell (DMAMC) Instrument Library (IL) and the Chemical, Biological, Radiological and Nuclear Defense Repository (CBRNDR) will be used to develop a current device inventory list and provide data artifacts, respectively.

3.2.1 Instrument Library

The DMAMC IL web application serves as the instrument repository for CWMD Test and Evaluation (T&E). Device metadata such as physical dimensions, weight, accessories, data storage, communication types, and manuals can be found on the site. Device information is manually uploaded to the site and can be updated as needed. Currently, the site holds information for over 250 different types of instruments: radiological (233), biological (8), chemical (3), other (6).

The IL will be leveraged when searching for device information.

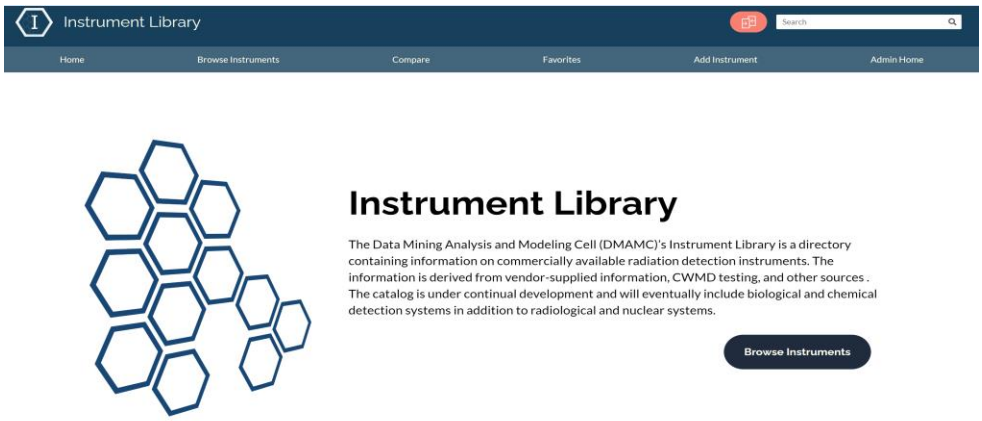

Figure 2. Data Mining Analysis and Modeling Cell Instrument Library.

3.2.2 Chemical, Biological, Radiological and Nuclear Defense Repository

The DMAMC CBRNDR web application is the data repository for CWMD T&E. The repository stores data artifacts captured by data collection systems used when executing a test. After Milestone (MS)-5, the Test Manager for an event will notify the assigned Data Manager to archive the data that was captured. These datasets include information such as detector output files, manuals, and vendor training information.

Currently, there is data for 324 systems under test (SUTs) and over 5 million files associated with different SUTs. Example datasets from radiological, biological, and chemical detectors will be used to back-test and design the DCM that will be applied to CBRN devices.

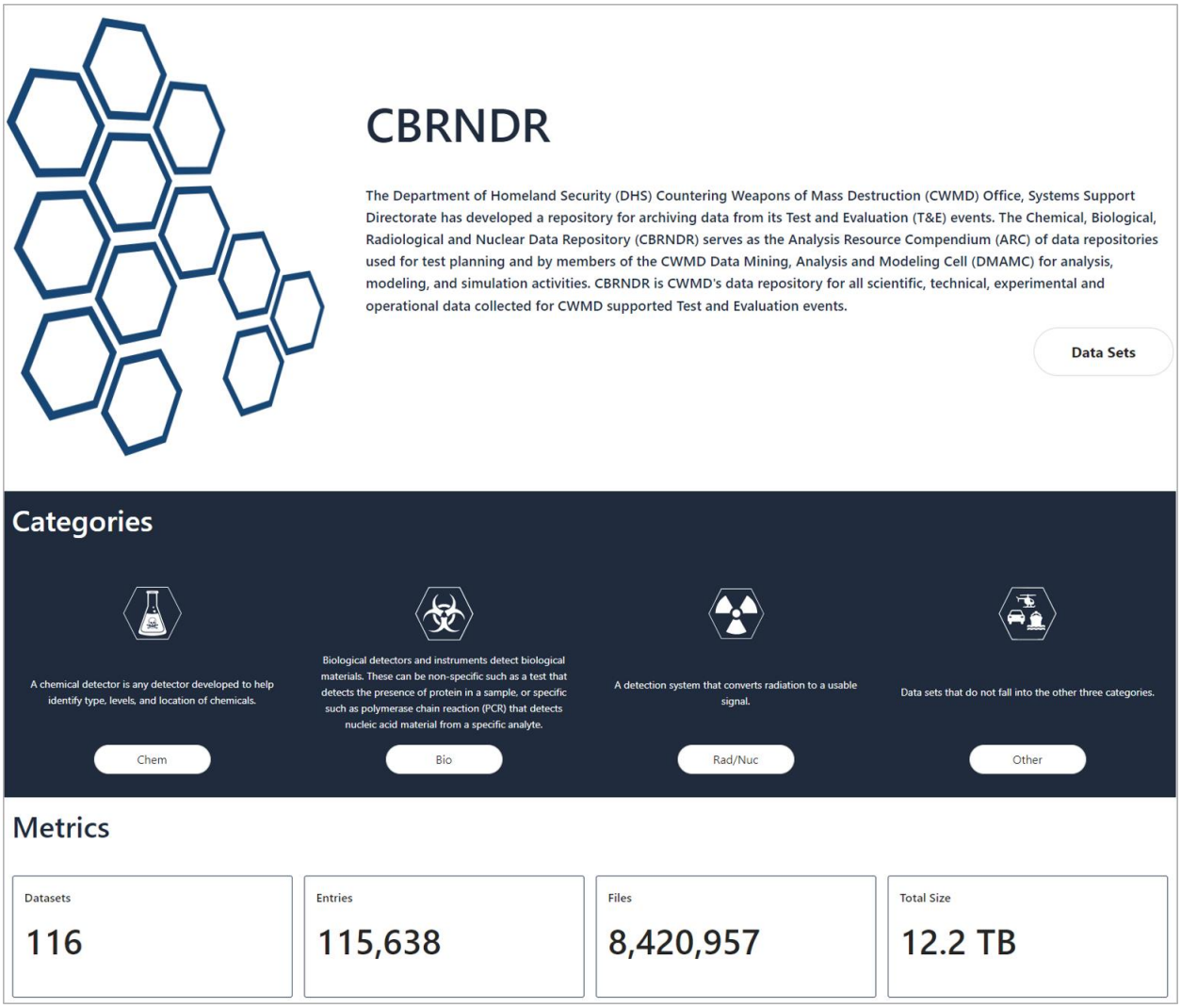

Figure 3. Chemical, Biological, Radiological and Nuclear Defense Repository.

3.2.3 NIEM Open Framework

The National Information Exchange Model (NIEM) is a framework for exchanging information between public and private sector organizations. The framework includes a reference data model for objects, properties, and relationships; and a set of technical specifications for using and extending the data model in information exchanges.

NIEM promotes scalability and reusability of messages between information systems which allows organizations to share data and information more efficiently. For over 20 years, NIEM has facilitated information exchanges across a variety of mission spaces and subject areas. What began as a solution for the law enforcement and homeland security communities has since evolved into a wide range of subject matters and areas.

- Since 2005, NIEM has been updated and maintained in a collaboration between the U.S. federal government, state and local government agencies, private sector, and non-profit and international organizations with new versions released roughly annually.

- In September 2022, the NIEM Executive Steering Committee announced that NIEM would transition to an OASIS Open Project known as the NIEM Open Project.[1] As an OASIS Open Project, the governance of NIEM is further expanded to include opportunities for public sector, private sector, non-profit and international organizations.
- In December 2022, NIEM released the latest version 5.2[2] which defines a set of common objects, "NIEM Core," and 17 sets of objects that are specific to certain government or industry verticals, "NIEM Domains."

Figure 4. NIEM Open Reference Model.

The NIEM model integrates terms, definitions, relationships, and formats for data exchange of up to 17 unique NIEM Domains, including a CBRN Domain for the CBRN Community of Interest (COI).[3] The NIEM core consists of data elements that are commonly understood and defined across subcommittees, such as person, activity, document, location, and item. It's governed jointly by all NIEM subcommittees. The CBRN COI governs the CBRN subcommittee under the stewardship of CWMD.

In general, the NIEM model consists of two related vocabularies: core elements that are commonly agreed to by all of the communities who use NIEM, and the community-specific elements that align to individual NIEM domains such as the CBRN Domain. To support the need for information-sharing in the radiological/nuclear prevention and detection mission area, CWMD, with the support and cooperation of the CBRN COI, developed a standard messaging protocol for interoperability and information sharing. During the development of the N.25 Protocol, the CBRN subcommittee was established to publish a set of data elements and common radiological and nuclear vocabulary that did not already exist in the NIEM model. The CBRN Domain is currently the primary source for data types and data properties for the N.25 information exchange package documentation (IEPD).

[1] https://lists.oasis-open-projects.org/g/niemopen
[2] https://niem.github.io/niem-releases/5.2/
[3] https://www.niem.gov/communities/cbrn

The CBRN Domain is an NIEM-conformant schema intended to be the baseline definition of business objects required to support data interchange needs of the chemical, biological, radiological and nuclear detection and interdiction mission area. Its initial definition has been developed by analysis of existing data specifications specific to radiation detection devices and messaging.

The CBRN Domain follows the NIEM data model object-oriented concepts of "types" (of things) that have "properties" and that participate in "relationships" with other "types" (of things). For example, the CBRN Domain contains 153 object types that represent various data types. In addition, the current release of the NIEM Model for the CBRN Domain[4] represents 592 data properties for the CBRN Domain, as shown in Table 2 below.

- An *object* is an instance of a type and is an abstraction of a specific physical thing or a conceptual thing. Also, in an object, the properties have values.
- An *object type* is a description of a set of things that share the same properties, relationships, and semantics.
- A *property* is a named characteristic of an object type. Furthermore, the property is of a specific type itself.
- A *relationship* may be modeled as either a type or a property.

Table 2. NIEM CBRN Domain Schema.

CBRN Domain	CBRN Domain Count
NIEM Domain Schema	CBRN - Chemical, Biological, Radiological, and Nuclear
NIEM Code List Schema	NIEM CBRN domain Radiological and Nuclear Code List
Facet	2,436
Local Term	40
Metadata	2
Property	592
Type	269
TypeContainsProperty	619

[4] http://release.niem.gov/niem/domains/cbrn/5.2/

4.0 Cybersecurity Data Categorization Approach

The device-level data and system-level attributes are foundational elements in the cybersecurity categorization approach introduced in this document. The approach is based on policy and governance guidance identified through online research, documentation provided by CWMD, and other documentation about security and data privacy and guidance set by DHS and the National Institute of Standards and Technology (NIST).

4.1 Policy and Guideline Review

The goals of data categorization are to (a) support the development of a data-centric device-system-dependent profile that aligns with device network capabilities and maps to DHS policy and governance and (b) create a device-agnostic, security categorization profile that includes the recommended security impact levels and security objectives.

4.1.1 Open-Source Review

OSINT information relating to CBRN data categorization, including data models and schemas related to similar Internet of Things (IoT) data and frameworks, was identified and reviewed. A short overview of open-source documentation is as follows:

- The Kaa IoT Platform[5] is an out-of-the-box, end-to-end enterprise platform providing data collection, device management, data processing, and IoT dashboards and analytics.
- NIST Special Publication (SP) 800-213A, IoT Device Cybersecurity Guidance for the Federal Government: IoT Device Cybersecurity Requirement Catalog[6] focuses on device cybersecurity capabilities (i.e., the features and functions needed from a device to support security controls) and nontechnical supporting capabilities (i.e., actions and support needed from device manufacturers and other supporting entities to support security controls).
- NIST SP 1800-36, Trusted Internet of Things (IoT) Device Network-Layer Onboarding and Lifecycle Management[7] is a draft document focusing on establishing trust between a network and devices with the credentials and policy needed to join a network known as network-layer onboarding. The National Cybersecurity Center of Excellence (NCCoE) demonstrates how to achieve trusted network-layer onboarding by applying standards, recommended practices, and available technology to securely provide network credentials to IoT devices and maintain security throughout their life cycle.
- NIST Interagency Report (IR) 8259, Foundational Cybersecurity Activities for IoT Device Manufacturers[8] focuses on foundational cybersecurity activities and recommendations that manufacturers should consider performing to improve the built-in security functions of new IoT devices produced.
- IoT Security Foundation, Secure Design Best Practice Guides[9] focus on 14 topic areas, including the classification of data, that serve as holistic guidance for security practitioners to implement in the design of IoT devices. Data classification recommendations include

[5] https://www.kaaiot.com/products/overview
[6] https://csrc.nist.gov/pubs/sp/800/213/a/final
[7] https://csrc.nist.gov/pubs/sp/1800/36/2prd
[8] https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8259.pdf
[9] https://iotsecurityfoundation.org/best-practice-guidelines/

defining a data schema into classes or levels of sensitivity for data protection to support data processing and compliance with policy and regulatory drivers.

The use of IoT frameworks offers significant advantages in categorizing CBRN data, including comprehensive data collection, real-time monitoring, enhanced situational awareness, potential integration with analytical tools, and improved efficiency in the prioritization of data. Leveraging these frameworks will enhance the approach to categorizing CBRN data while gaining a better understanding of the CBRN landscape.

4.1.2 Government Policy

CWMD leverages the application of the Risk Management Framework (RMF) for authorized systems across operating environments. The DHS Sensitive Systems Policy Directive 4300A, Version 13.1[10,11] specifically states that components, such as CWMD, should assure that the RMF is applied to designed critical infrastructure for improving and ensuring adequate critical infrastructure cybersecurity, as recommended in the NIST Framework for Improving Critical Infrastructure Cybersecurity (i.e., The Cybersecurity Framework).[12] To ensure that appropriate mission impact levels and security objectives are aligned, the RMF specifies, in Step 1 of the process, that all systems are properly categorized, whereby the categorization of asset information types and system/subsystem types is determined, in accordance with NIST 800-60 Volumes I[13] and II[14] for sensitive but unclassified information and the Committee on National Security Systems Instruction (CNSSI) 1253[15] for national security systems, where applicable.

The impact levels for CWMD systems and subordinate devices should be categorized using the Federal Information Processing Standard Publication 199 (FIPS 199), “Standards for Security Categorization of Federal Information and Information Systems,” as Low (L), Moderate (M), or High (H) to align with respective minimum security control baselines.[16]

[10] Note: This policy implements DHS Management Directive 140-01, “Information Technology Security Program.”
[11] U.S. Department of Homeland Security (DHS) Sensitive Systems Policy Directive 4300A, Version 13.1, July 27, 2017.
[12] U.S. Department of Homeland Security (DHS). Sensitive Systems Policy Directive 4300A, Version 13.1, October 2, 2019.
[13] National Institute of Standards and Technology (NIST) Special Publication (SP) 800-60, Volume I: Guide for Mapping Types of Information and Information Systems to Security Categories, Revision 1, August 2008. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v1r1.pdf.
[14] National Institute of Standards and Technology (NIST) Special Publication (SP) 800-60, Volume II: Appendices to Guide for Mapping Types of Information and Information Systems to Security Categories, Revision 1, August 2008. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v2r1.pdf.
[15] Committee on National Security Systems (CNSS) Instruction 1253, Security Categorization and Control Selection for National Security Systems, March 27, 2014. https://www.cnss.gov/CNSS/openDoc.cfm?n4wWvsdOQfNKeOqvvE1XSQ==.
[16] FIPS 199 is a federal standard that guides the level of security needed for a system based on the potential impact of a security breach, or in this case compromised data.

For this data categorization effort, we used the three levels of impact described in FIPS 199:

1. **Low Impact:** A breach resulting in a limited adverse effect on organizational operations, organizational assets, or individuals.
2. **Moderate Impact:** A breach resulting in a serious adverse effect.
3. **High Impact:** A breach resulting in a severe or catastrophic adverse effect.

The potential impact for each device system can be Low, Moderate, and High for device systems that connect and transmit different types of information and may have different impact levels. FIPS 199 states, "the potential impact values assigned to the respective security objectives shall be the highest value from among those security categories that have been determined for each type of information resident on the information system." To estimate the impact levels for the device systems, the system's security-related attributes are translated into the security categories and security objectives of the CWMD T&E mission: Confidentiality (C), Integrity (I), and Availability (A), known as the CIA triad. The CIA security categories and objectives stem from CNSSI 1253[17] for national security systems, where applicable. Each device system and the information it contains must be categorized based on the highest level of impact for any of the security aspects (confidentiality, integrity, or availability).

Once CWMD has determined the impact levels on the devices of systems based on network connections and data transfers, the impact levels should be aligned with the attributes associated with network connections and communications. For example, if a device is connected to a wireless network and transmits different data types based on the confidentiality of the data, the highest impact value for each security objective should represent the device system's CIA impact level. The nexus between the data categorization of devices and system-level mission requirements is crucial for guiding the selection of appropriate security controls and ensuring that devices supporting federal information systems are properly categorized and protected at an appropriate level.

Therefore, a process defined for categorizing data sensitivities for operational data, network data, and other shared information is an essential building block to determine data sensitivities.

4.2 Data Categorization Priorities

Data categorization priorities can vary depending on the specific needs and objectives of CWMD and their organizational partners. Categorization priorities are dependent on multiple factors that can include the following:

- Confidentiality Levels: Data may be categorized based on its level of confidentiality, ranging from public information to highly confidential data that only specific individuals or teams should have access to.
- Data Ownership: Categorizing data based on ownership can help clarify who is responsible for managing and maintaining different datasets within an organization.
- Data Quality: Prioritizing data based on its quality can help ensure that accurate and reliable information is readily available for decision-making and analysis. This may involve

[17] Committee on National Security Systems (CNSS) Instruction 1253, Security Categorization and Control Selection for National Security Systems, March 27, 2014. https://www.cnss.gov/CNSS/openDoc.cfm?n4wWvsdOQfNKeOqvvE1XSQ==.

categorizing data as high-, medium-, or low-quality based on factors such as accuracy, completeness, and consistency.

- Data Life Cycle Stages: Data can be categorized based on its stage in the data life cycle, including creation, storage, usage, and archival or deletion. Understanding where data falls in this life cycle can inform decisions about storage, access controls, and retention policies.
- Regulatory Compliance: Data categorization may be driven by regulatory requirements specific to CWMD or SLTT stakeholders.
- Data Sensitivity: Information may be sensitive because of its strategic value or potential impact on CWMD if compromised. Categorizing data based on sensitivity levels can help determine appropriate security measures and access controls.
- Data Usage Patterns: Understanding how data is used within CWMD and other government organizations can inform categorization priorities. For example, frequently accessed data may be prioritized for optimization and performance tuning, while rarely accessed data may be a candidate for archival or deletion.
- Data Interdependencies: Data may be categorized based on its relationships or dependencies with other datasets. Identifying these interdependencies can help ensure that data is managed and accessed in a way that maintains integrity and consistency across the organization.

Leveraging these data categorization priorities, a decision tree (DT)-based machine learning model to classify future CBRN measurements into categories can be developed.

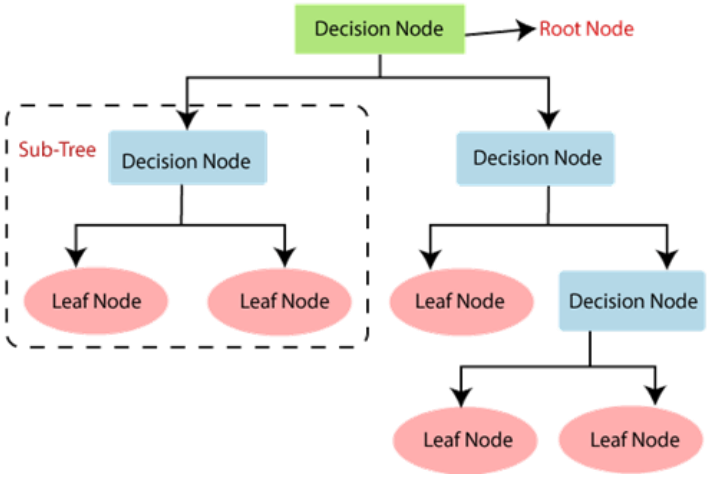

Figure 5. Baseline example of a decision tree.

DTs are nonparametric supervised models that can expand their structure in terms of width and depth of the model purely based on the training measurements. DTs are used for multiclass problems for each of the priority categories.

4.2.1 Data Buckets for Decision Trees

Hierarchical data buckets for DTs based on low, moderate, and high sensitivities involve organizing data into nested categories or levels according to the sensitivity of the information. This hierarchical structure allows for a more nuanced approach to managing and protecting data, with different levels of sensitivity receiving varying degrees of attention, security

measures, and access controls. Security impact levels, introduced above and denoted from FIPS 199 and NIST SP 800-60 Volumes I & II, represent data sensitivities that are further characterized as follows:

- **Low:** This data is considered nonsensitive or of minimal risk if accessed or disclosed. This might include publicly available information, and access to this data may be relatively unrestricted, with minimal security measures in place to protect it.
- **Moderate:** This data is moderately sensitive and may pose a moderate risk if accessed or disclosed inappropriately. Access to data in this category may be restricted to authorized CWMD personnel, with additional security measures such access controls in place.
- **High:** This data includes highly sensitive information that could pose significant risks to CWMD if accessed or disclosed without proper authorization. Access to this data is tightly controlled and restricted to only those individuals or teams with a legitimate need to know. Security measures such as encryption, multifactor authentication, and strict access controls are typically implemented to safeguard this data.

Within each sensitivity level, further subcategories or nested levels can be defined based on specific criteria or attributes of the data. A DT can be used to guide data management and protection strategies based on the sensitivity levels allocated to the device- and operational-level data. Furthermore, DTs can outline processes for data classification, access control policies, security requirements, data retention and disposal practices, incident response procedures, and other considerations tailored to each sensitivity level.

The NIEM data exchange model can be utilized to formulate the connections and sensitivities of data gathered from CBRN detection devices. The repository currently has data elements that have been shared with the NIEM community based on CBRN data from state and local authorities. The data category sensitivity in Figure 5, provides context for how each data element can change the data sensitivity based on how the data is aggregated and connected.

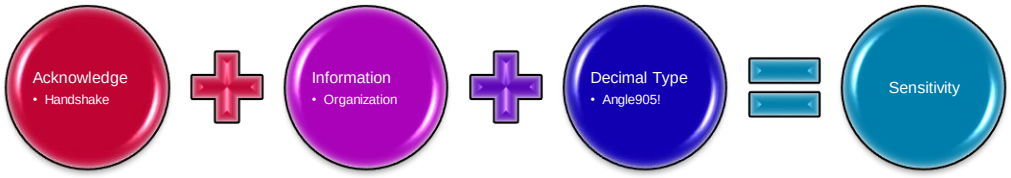

Figure 6. Data Category Sensitivity.

The NIEM also includes meta-data elements that can be used in a decision tree process to further determine how to categorize CBRN device data (Figure 6).

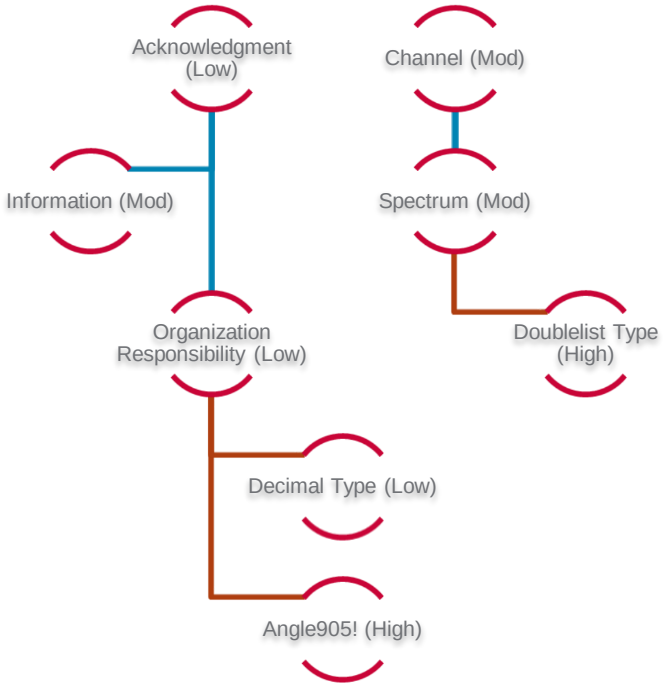

Figure 7. Example Metadata.

4.2.1.1 Data Sensitivities

To measure data sensitivities, certain factors need to be taken into consideration that include the source of the data, the intended use of the data, the potential for harm if the data is disclosed, and the applicable legal and regulatory requirements. This study has taken into consideration the nature of the data (e.g., operational, network), how the data is stored, and how the data is transmitted and answered questions related to who has access to the data and what the potential fallout would be if the data has been compromised.

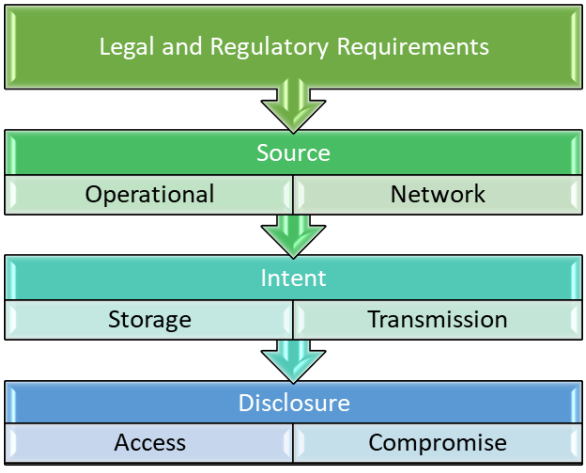

Figure 8. Data Factors.

Categories or buckets that the data will fall into include the regulatory requirements for organizational compliance; the accessibility of the data; the life cycle of the data based on the

age, purpose, and usage by an organization; additional data dependencies; and the value of the data. These factors will influence how the data is stored, processed, or managed by organizations based on the confidentiality and accessibility.

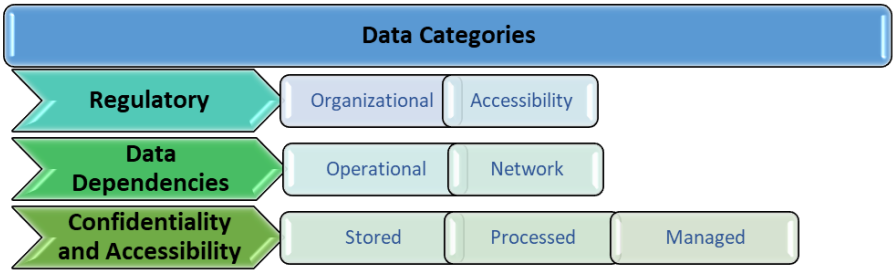

Figure 9. Example of Data Categories.

4.2.1.2 Intended Use

CBRN real-time data is used for decision-making to prevent or mitigate the effects of CBRN incidents and for public health and safety. Data information sharing among many strategic organizations (e.g., SLTT) may be different based on device use personnel, deployment location, and chain of command communications. Identifying the data that needs to be shared and how it shared among differing platforms informs how the data is handled and the sensitivities associated with network and operational data. The source of data collection can be geographically different depending on the device type, intended use, and network communications.

4.2.2 Operational Data

Operational data from CBRN devices is crucial for calibration, accuracy, safety, early warning signaling, and optimization. By analyzing operational data, organizations can detect patterns or anomalies that can indicate potential CBRN hazards. By understanding the data patterns, organizations can make informed decisions on response activities, but data sensitivities can be present that should not need to be provided to personnel outside current CWMD purview.

Determining the sensitivity of operational data from CBRN devices will depend on multiple factors such as the type of detector, manufacturer, device complexity, communication capabilities, and additional support equipment needed to operate a device. In general, device capabilities and features tend to increase going from mobile devices such as pagers to stationary devices like radiation portal monitors (RPMs). These increases in data output and the number of connected devices can lead to an expansion of additional connected devices that support the primary CBRN instrument, which leads to an increase in volume and variety of operational data streams.

Example operational datasets were randomly selected from CBRNDR and reviewed for information that could be considered sensitive. Additionally, any current guidance for operational data was internally reviewed. Information indicated that in 2022 there was a guidance document created through an interlay working group; however, a copy of that document was unable to be obtained.

The baseline operational data perimeters for manufacturer-agnostic devices can include the number of files generated, the file category type, the number of files, the class of the detection system, and the file extension that determines how to read the data.

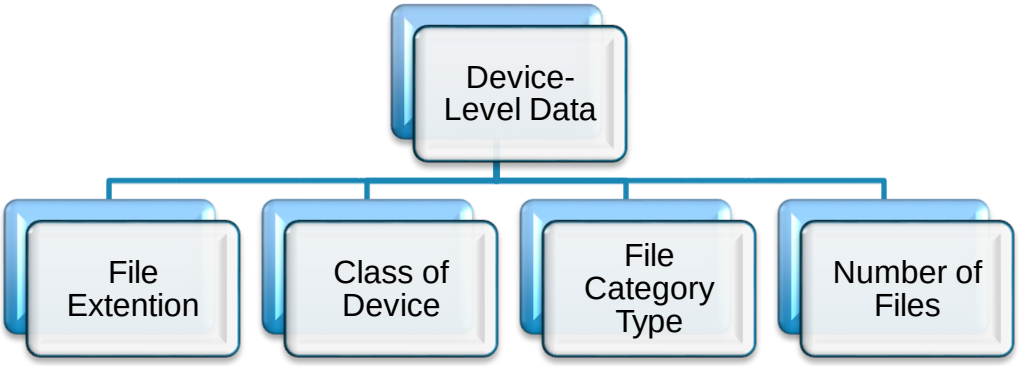

Figure 10. Device-Level Attributes.

Operational data for deployed/used devices can then be added to individual buckets for categorizing data sensitivities based on the data captured and generated by the devices in the field. These buckets can include information on users, location, measurements, and serial numbers of the devices. The sensitivity will be determined based on the combination of information shared within the devices' operational perimeters and the data generated during use.

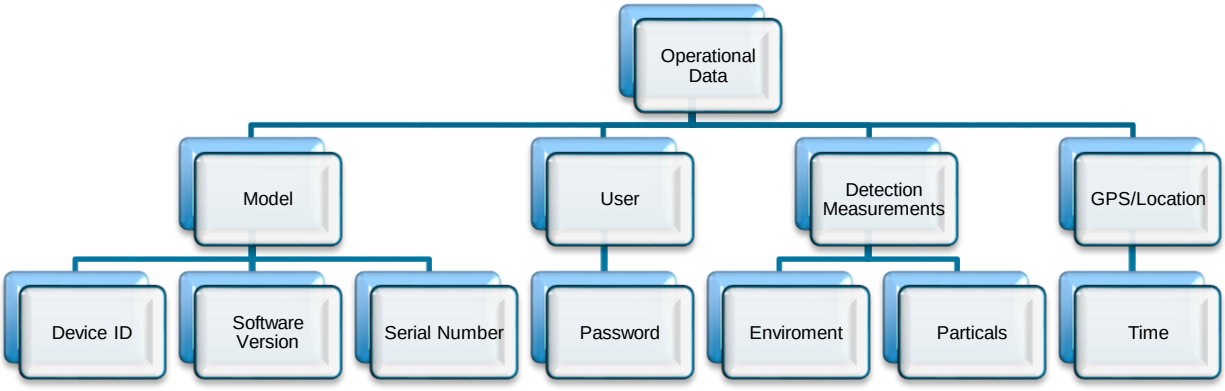

Figure 11. Device-Generated Data.

Device-level attributes and device-generated data can be grouped into meaningful categories based on common attributes and identified criteria. Categorization criteria are based on policies that dictate information-sharing processes and procedures, dependent on who the information is being shared with, which may change the classification of the data and how it is shared.

4.2.3 Network Data

Network data elements are frequently identified through packet analysis tools because these tools allow for deep inspection and analysis of network traffic at the packet level. When packet analysis tools are used in conjunction with CBRN operational data mechanisms, it allows for better details on network data elements such as headers, payloads, protocols, and metadata. The network data is important for categorizing all data elements to help identify potential security threats such as malware, intrusion attempts, or data exfiltration. Packet analysis tools can aid in compliance monitoring and auditing. By capturing and analyzing network traffic, organizations can ensure adherence to regulatory requirements and internal policies regarding data privacy, security, and governance. Table 3 summarizes the variability of these network data elements.

Table 3. Network Data Elements.

Field Name	Field Description	Risk
Source Address	• Media Access Control (MAC) address of the device. • Source Internet protocol (IP) address of the packet. Bluetooth Device Address Unique Identifier.	High
Destination Address	• MAC address of the device that is supposed to receive the frame. Destination port number used by the receiving application.	High
Company Identifier	Manufacturer Specific Data Advertising Data (AD) element within Bluetooth Low Energy (BLE) advertising packets.	High
Hypertext Transfer Protocol (HTTP)	Plaintext username and password.	High
Communication Protocol	Network traffic based on specific protocols.	Moderate
Packet Length	Packet lengths and related information.	Moderate
Time	Network packets based on their timestamp.	Low

Users should recognize the risks associated with incorrectly categorizing data as “High.” For instance, in certain scenarios, HTTP data can reveal plaintext usernames and passwords within the packet’s detail pane. Data related to HTTP traffic would fall under the “High” category and should be handled carefully to ensure that it is not improperly labeled and stored in an environment for open access.

4.3 Data Categorization Development

The operational data of CBRN detection devices must be properly categorized, identified, and secured throughout the device life cycle to reduce the threat surface to the mission of the CWMD program in critical environments. CBRN data that the device is capturing and/or

disseminating, including data sensitivities and thresholds, must be properly categorized to support structured data standards across disparate ownership platforms (e.g., SLTT, etc.).

Categorization based on the *type* of data involves multidimensional assessment of the underlying information model and the downstream application. To support a variety of CBRN use cases, we will build upon the fundamental big data characteristics defined by NIST[18]:

- Volume: This represents the amount of CBRN data available for analysis to extract valuable information. Volume is key in the design of a data management framework supporting storage, indexing, extraction, and data movement.
- Velocity: The rate at which CBRN data is generated by different types of detectors and sensors is crucial to identify the temporal scale of each stream and the classes of downstream applications.
- Variety: Classification based on multiple sources, domains, and data modalities is a crucial characteristic for CBRN data alignment, fusion, and automated data ingestion pipelines.
- Variability: Variability refers to expected changes in the data, and it affects the sensitivity of any downstream applications. For CBRN data, we need to categorize similar data sources based on the range of sensor measurements.

[18] NIST Big Data Public Working Group (NBD-PWG) (2019), NIST Big Data Interoperability Framework: Volume 1, Definitions, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.SP.1500-1r2 (Accessed May 1, 2024).

5.0 Next Steps

CWMD is tasked with protecting the United State from CBRN threats by utilizing CBRN detection device technologies. To effectively fulfill their mission, CWMD collects and analyzes vast amounts of data from diverse sources, including CBRN equipment in the field, intelligence agencies, law enforcement, customs, and state and local municipalities. Advanced data categorization techniques, such as DCMs and semantic analysis, are utilized to categorize the collected data into relevant categories. This may include patterns indicative of CBRN threats or security concerns. The captured data is analyzed for categorization to assess potential threats, emerging trends, and anomalous behaviors that may be indicative of imminent threats.

A DCM is used to integrate and cross-reference categorized data from multiple pieces of CBRN equipment to identify connections, networks, patterns, and relationships between intelligence agencies, law enforcement, customs, and state and local municipalities. The DCM can prioritize threats and risks based on their severity, likelihood, and potential impact on national security. This also enables the allocation of resources and efforts focused on addressing high-priority threats while considering strategic incident response activities. Decision support is paramount between CWMD and other entities to provide actionable intelligence and insights for supporting informed decision-making, resource allocation, incident response, and operational planning. The DCM presents findings through visualizations, reports, and dashboards to facilitate the understanding and communication of complex CBRN data. The DCM will enable the categorization and analysis of CBRN data sources and enable the detection of emerging threats and security vulnerabilities before they escalate, allowing for proper mitigation measures to be put in place. By identifying and prioritizing anomalous data categorization, CWMD can allocate resources, personnel, and additional CBRN needs more efficiently. By leveraging a DCM, CWMD will further enable the protection of the United States from CBRN threats.

The high complexity of data generated and captured from CBRN equipment may include measurement readings and network data, which can make it challenging to identify meaningful categories or labels. The diversity of data types, formats, and structures adds to the complexity, requiring specialized techniques to be effective. CBRN datasets have many features and variables that can present a challenge in terms of computational complexity, scalability, and the risk of overfitting when building a DCM. The proposed solution considers the dependencies, trends, and network patterns with deployed CBRN equipment. When categorizing datasets using a DCM solution, it will need to be scalable and efficient by enabling the processing, analyzing, and bucketing of massive volumes of data in real time or near real time. The PNNL team will take into consideration data captured from NIEM to influence the data models to include all metadata and data points identified.

The DCM solution will be used to establish standardized formats and protocols for sharing categorized data to ensure consistency and interoperability between different CBRN systems and organizations. This will facilitate seamless data exchange and integration, reducing barriers to collaboration efforts across multiple agencies and municipalities.

5.1 Data Categorization Model

Data visualization is a key component in sharing categorized data with CBRN users and organizations for analysis and incident response activities. The DCM is designed as an easy to use, modular, extensible, and scalable user interface (UI) to summarize and explain the data categories, patterns, and correlations. The resulting UI follows a *dashboard* design methodology

that can integrate multiple visualizations from different data categories and analysis into a single workspace. The CBRN Data Categorization Dashboard will be developed using the sequential automated workflow in Figure 12 to comply with the data modality, sensitivity, and access controls defined by CWMD policies and share insights with organizations and incident responders. The workflow will be designed to analyze diverse datasets based on network and measurement data and generate a collection of insights, as described below.

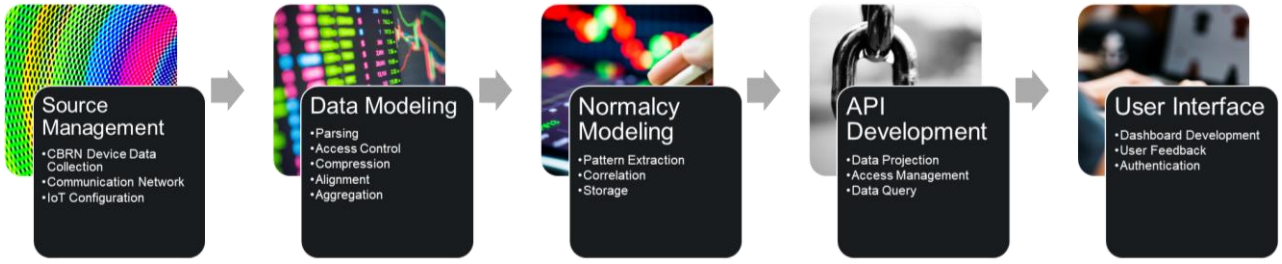

Figure 12. DCM Workflow.

5.1.1 CBRN Data Categorization Dashboard

To integrate all these visualization results, we will develop a data sensitivity and accept-level aware CBRN dashboard that can be extended as more analytical applications are developed. As shown in Figure 13, the dashboard categorizes data in six focus areas: “Assets,” “Communication,” “Topologies,” “Timeline,” “Risk,” and “Mitigation.” Each tab extracts the relevant insight from the data and access levels selected by the user. The portal will serve as a knowledge management system for CBRN data characterization with support for user access control, multifactor authentication, user sessions, information sharing, and report generation.

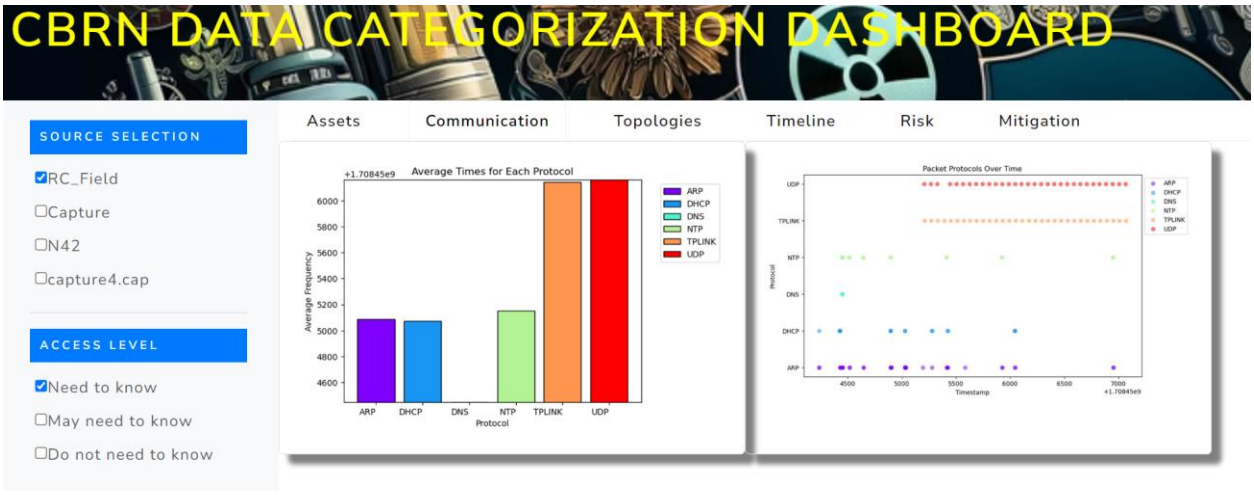

Figure 13. CBRN Data Categorization Dashboard.

5.1.2 CBRN Data Decision Tree Visualization

Working with section 2.2 (Data gathering process and methods) in this document, we will store the types of measurement information and network interactions that each CBRN device

produces. We will identify hierarchical data patterns in the measurement and network data and plan to use it to correlate data across different domains. For example, measurement data is exported using the “N42” standard that has a well-defined taxonomy. We will visualize the data using hierarchical graph layouts to assist users in understanding the variety of information that may be available to them.

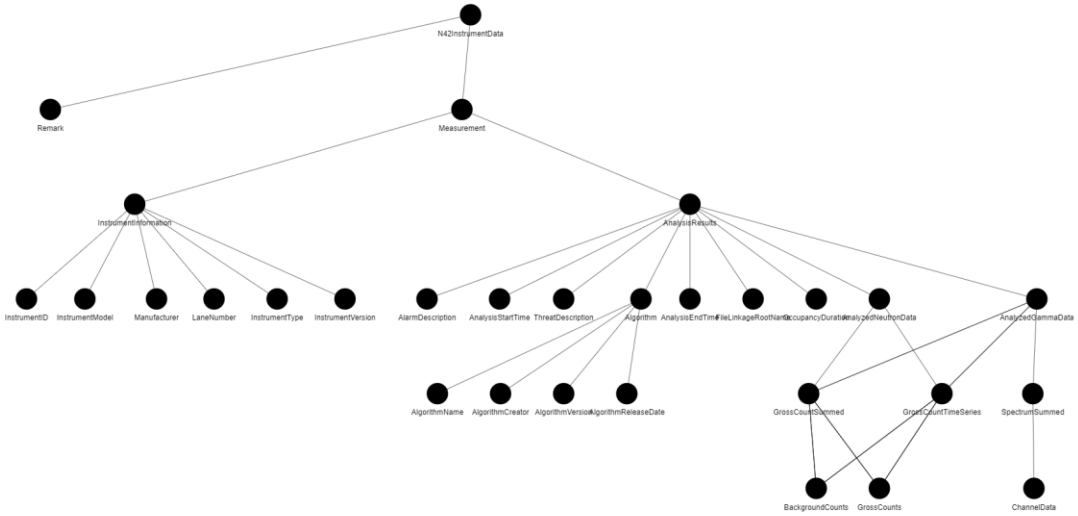

Figure 14. CBRN Data Decision Tree.

5.1.3 Time Series Data Visualization

Both CBRN devices and other networked devices generate a stream of data representing detection events and network traffic. For such data types, PNNL will develop a multiple temporal data visualizations including scatterplots, moving averages, etc. For example, we can highlight the protocols used in the network traffic packages to indicate the activation of new CBRN devices in the network and the applications installed on them.

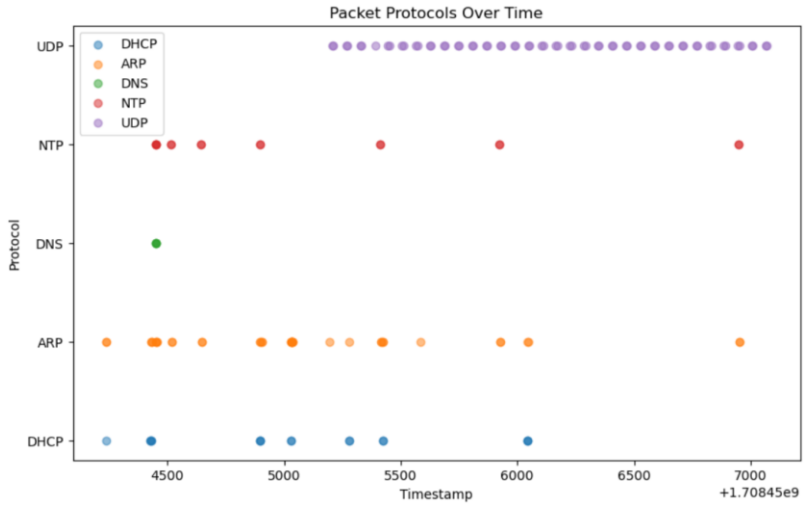

Figure 15. Temporal Pattern of Device Protocols Used.

5.1.4 Topological Visualization

Topological visualization is imperative to highlight insights about the interconnection and communication patterns of CBRN devices. It is also useful to generate additional context about CBRN measurements by connecting metadata and the categorical attributes available in the measurements. Additionally, we will also generate subgraph (a.k.a. *motif-based*) embeddings for topology characterization (Purohit et al. 2022) that can be used to correlate different entities in the network based on their interconnections and local behavior, as shown in Figure 16 and Figure 17.

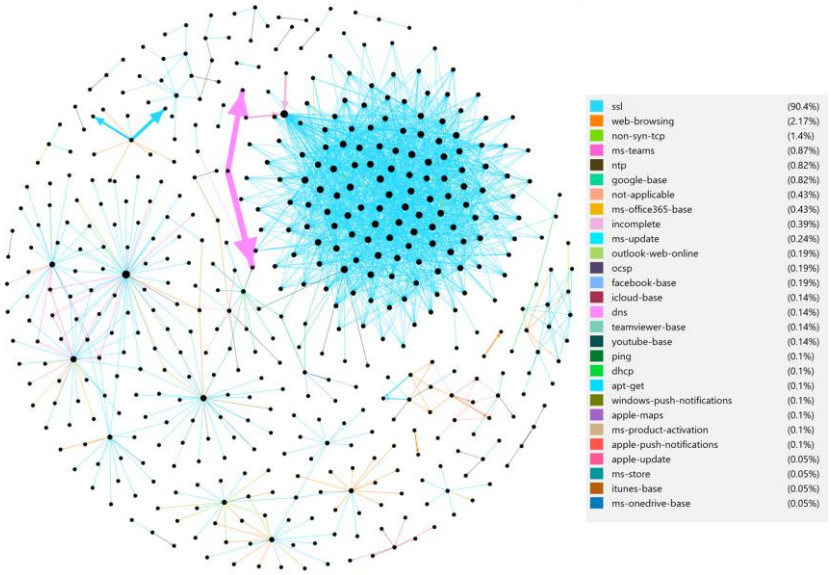

Figure 16. Graph-Based Visualization Highlighting Device Connections.

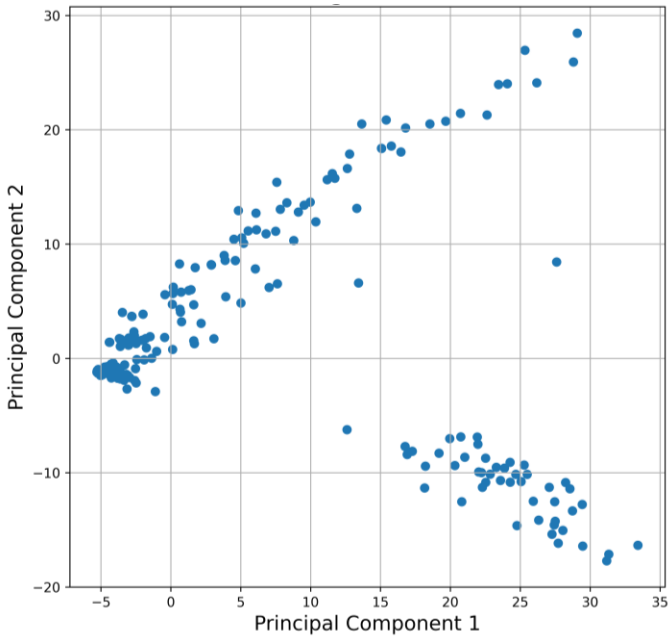

Figure 17. Device Similarity Based on Network Traffic.

5.2 Device System Categorization

PNNL has developed a Device System Categorization (DSC) process and analyzed the main policy-driven guidelines to provide CWMD with a holistic approach to cybersecurity processes and applications for their systems. The developed approach provides insight into the security-related controls needed for systems based on network connectivity, information processed and exchanged, and outside interactions with devices. This approach is outlined in two ways:

1. Adapt existing guidelines to CWMD systems for a holistic risk-based application focused on threats to CWMD systems.
2. Integrate system network communications and connections into the DSC process.

The CWMD DSC process is iterative and complements the existing DHS T&E Strategy for the Cyber Resilience process in Figure 18, which is a component of the overall system resilience and is supported in part by RMF cybersecurity activities and artifacts. The DSC process should be revisited as the system design and implementation continue.

CWMD DSC activities should be documented in DHS Cyber Resilience T&E artifacts (the Mission Need Statement, Operational Requirements Document, Concept of Operations, and the Test Evaluation Master Plan), and the results of the security categorization should be documented in RMF artifacts (Risk Assessment Report and the System Security Plan [SSP]).

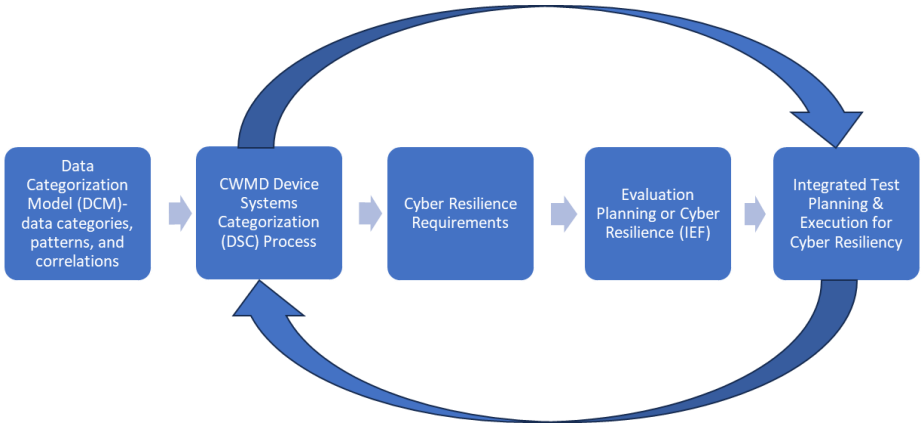

Figure 18. CWMD Device Systems Categorization T&E Process.

The CWMD DSC process provides an overview of the security requirements for CWMD systems and describes the controls that are planned or are in place for implementation (Appendix A). The security controls that are appropriate align with the information that is being transmitted, processed, or stored by the system. The categorization process also considers the system boundaries when determining what security controls need to be in place. If the security boundary changes, then the system in question will need to be reevaluated, and additional security may need to be implemented. Proper management is essential to assure that the confidentiality, integrity, and availability of the system are properly protected and that the proper security application is applied. The implementation also includes any requirements that are set forth by applicable laws, regulations, CWMD policies, procedures, and practices.

6.0 References

CNSSI No. 1253. 2014. *Security Categorization and Control Selection for National Security Systems*. Committee on National Security Systems (CNSS).

Countering Weapons of Mass Destruction Office (CWMD). 2018. *Solution Development Process (SDP) Guide*. U.S. Department of Homeland Security 100-CWMD-117550 v1.00. Washington, DC.

Countering Weapons of Mass Destruction Office (CWMD). 2020. *Operating Instruction-1: Test & Evaluation*. U.S. Department of Homeland Security 500-CWMD-123660, Version 2.06. Washington, DC.

Diyeb, I. A. I., A. Saif, and N. A. Al-Shaibany. 2018. "Ethical Network Surveillance Using Packet Sniffing Tools: A Comparative Study." *International Journal of Computer Network and Information Security* 10 (7): 12–22. https://doi.org/10.5815/ijcnis.2018.07.02.

Domestic Nuclear Detection Office (DNDO). 2013. *Acquisition and Commercial Engagement Strategy (DACES) Commercial First Guidebook*. U.S. Department of Homeland Security 100-DNDO-123150v1.00.

FIPS 199. 2004. *Standards for Security Categorization of Federal Information and Information Systems*. Gaithersburg, MD: National Institute of Standards and Technology. https://csrc.nist.gov/files/pubs/fips/199/final/docs/fips-pub-199-final.pdf.

Joint Task Force Transformation Initiative. 2010. *Guide for Applying the Risk Management Framework to Federal Information Systems: A Security Life Cycle Approach*. National Institute of Standards and Technology SP 800-37, Revision 1. Gaithersburg, MD. https://doi.org/10.6028/NIST.SP.800-37r1.

NIEM OPEN. Chemical, Biological, Radiological and Nuclear (CBRN). https://www.niem.gov/communities/cbrn

NIEM OPEN. 2022. NIEM 5.2 Release (Current). National Information Exchange Model. https://niem.github.io/niem-releases/5.2/

NIST Big Data Public Working Group (NBD-PWG). 2019. *NIST Big Data Interoperability Framework: Volume 1, Definitions*. National Institute of Standards and Technology SP 1500-1r2, Version 3. Gaithersburg, MD. https://doi.org/10.6028/NIST.SP.1500-1r2.

OASIS OPEN PROJECTS. 2024. NIEM OPEN: NIEM Open Project. https://lists.oasis-open-projects.org/g/niemopen

Purohit, S., G. Chin, and L. B. Holder. 2022. "ITeM: Independent Temporal Motifs to Summarize and Compare Temporal Networks." *Intelligent Data Analysis* 26 (4): 1071–1096. https://doi.org/10.3233/IDA-205698.

Stine, K., R. Kissel, W. C. Barker, J. Fahlsing, and J. Gulick. 2008a. *Volume I: Guide for Mapping Types of Information and Information Systems to Security Categories*. National Institute of Standards and Technology SP 800-60 Volume I Revision 1. Gaithersburg, MD. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v1r1.pdf.

Stine, K. R. Kissel, W. C. Barker, A. Lee, and J. Fahlsing. 2008b. *Volume II: Appendices to Guide for Mapping Types of Information and Information Systems to Security Categories*. National Institute of Standards and Technology SP 800-60 Volume II Revision 1. Gaithersburg, MD. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v2r1.pdf.

U.S. Department of Homeland Security (DHS). 2016. *Acquisition Management Directive 102-01 Rev. 03*.

U.S. Department of Homeland Security (DHS). 2017. *Directive 026-06 Rev. 01, Test and Evaluation*.

U.S. Department of Homeland Security (DHS). 2019a. *Sensitive Systems Policy Directive 4300A, Version 13.1*.

U.S. Department of Homeland Security (DHS). 2019b. *Cyber Resilience Test and Evaluation (T&E) Supplemental Guidance, v2.0*. U.S. Department of Homeland Security. Washington, DC.

U.S. Department of Homeland Security (DHS). 2023. "DHS 4300A Sensitive Systems Attachments." U.S. Department of Homeland Security. Last Modified December 18, 2023. Accessed May 14, 2024. https://www.dhs.gov/publication/dhs-4300a-sensitive-systems-handbook.

Wireshark. n.d. "Display Filter Reference: Bluetooth Common." Wireshark. Accessed May 14, 2024. https://www.wireshark.org/docs/dfref/b/btcommon.html.

Wireshark Q&A. 2021. "How Source and Destination Is Identified?" Wireshark Q&A. Accessed May 14, 2024. https://osqa-ask.wireshark.org/questions/1863/how-source-and-destination-is-identified/.

Appendix A – Device System Categorization

The Countering Weapons of Mass Destruction Office (CWMD) recognizes that threats in the cyberspace domain continue to grow, which require chemical, biological, radiological, and nuclear (CBRN) detection devices and supporting systems to be both cybersecure and resilient to cyber threats. Similarly, the CWMD Device System Categorization (DSC) process supports existing U.S. Department of Homeland Security (DHS) Cyber Resilience[19] Test and Evaluation (T&E) activities, which consist of iterative processes, starting at the initiation of system acquisition and continuing throughout the entire acquisition life cycle.[20]

The CWMD DSC process outlined in this appendix is structured to correspond with DHS Cyber Resilience T&E activities[21] and existing processes and activities within the Acquisition Lifecycle Framework (ALF), ALF Acquisition Decision Events, the NIST Risk Management Framework (RMF)[22], and the Systems Engineering Life Cycle (SELC) phases.

CWMD's T&E effort follows a four-phase process for the planning, execution, analysis, and reporting of test events for CBRN detection systems[23] in accordance with DHS Acquisition Management Directive 102-01;[24] DHS Directive 026-06, Test and Evaluation;[25] the CWMD Solution Development Process;[26] and the CWMD Acquisition and Commercial Engagement Strategy.[27] Since CWMD uses a milestone review process to track progress and approve advancement through the four phases, the DSC process should occur in the early stages of the planning phase. In order to assure that T&E are executed effectively, the DSC process for categorization should include all aspects of the planning phase prior to completing the first testing event milestone. In support of DHS policy, directives, and strategy, this guidance builds upon the systematic approach to govern T&E activities outlined within CWMD Operating Instruction-1: Test and Evaluation best practices that are

- consistent with DHS T&E directives, policy, and guidance

[19] U.S. Department of Homeland Security (DHS). Science and Technology (S&T). Cyber Resilience Test and Evaluation (T&E) Supplemental Guidance, v2.0, October 2019.
[20] U.S. Department of Homeland Security (DHS). Science and Technology (S&T). Cyber Resilience Test and Evaluation (T&E) Supplemental Guidance, v2.0, October 2019.
Cyber resilience is the ability of an information system to continue to operate while under attack, even if in a degraded or debilitated state, and to rapidly recover operational capabilities for essential functions after a successful attack.
[21] U.S. Department of Homeland Security (DHS). Science and Technology (S&T) Cybersecurity Systems Engineering Implementation Guide, v1.0, August 1, 2019.
[22] NIST SP 800-37, Rev 1, "Guide for Applying the Risk Management Framework to Federal Information Systems: A Security Life Cycle Approach," February 2010, updated June 5, 2014.
[23] U.S. Department of Homeland Security (DHS) Countering Weapons of Mass Destruction Office (CWMD). Operating Instruction-1: Test and Evaluation, 500-CWMD-123660, Version 2.06, February 14, 2020.
[24] U.S. Department of Homeland Security (DHS). Acquisition Management Directive 102-01 Rev. 03, July 26, 2016.
[25] U.S. Department of Homeland Security (DHS). Directive 026-06 Rev. 01, Test and Evaluation, May 05, 2017.
[26] U.S. Department of Homeland Security (DHS) Countering Weapons of Mass Destruction Office (CWMD). Solution Development Process (SDP) Guide, 100-CWMD-117550 v1.00, September 27, 2018.
[27] U.S. Department of Homeland Security (DHS) Domestic Nuclear Detection Office (DNDO). Acquisition and Commercial Engagement Strategy (DACES) Commercial First Guidebook, 100-DNDO-123150v1.00, September 30, 2013.

- applicable to all CWMD acquisition programs
- repeatable to assure standardized, sufficient rigor and implementation of T&E best practices.

Cybersecurity data categorization plays a key role early in the acquisition life cycle and supports cyber resilience T&E activities prior to execution in the acquisition and development timeline. Figure A.1 presents a high-level overview of the nexus between DHS Cyber Resilience T&E activities, the four phases of the ALF, the nine activities of the SELC, and the six steps of the RMF. The RMF steps and related cybersecurity activities are the cornerstone of the CWMD DSC process and a component of overall system resilience. The DHS Sensitive Systems Policy Directive 4300A, Version 13.1[28,29] specifically states components, such as CWMD, "should assure that the (Department) RMF is applied to designed critical infrastructure for improving and ensuring adequate critical infrastructure cybersecurity, as recommended in the NIST Framework for Improving Critical Infrastructure Cybersecurity" (i.e., The Cybersecurity Framework).[30] Therefore, CWMD's application of the RMF for categorization—namely, Step 1, Categorize System data—is an imperative step to assure appropriate mission needs and alignment are established prior to the initiation of DHS Cyber Resilience T&E activities. Furthermore, the CWMD DSC process follows RMF Step 1 guidance, whereby the categorization of asset information types and system/subsystem types is determined in accordance with NIST 800-60 Volumes I[31] and II[32] for sensitive but unclassified systems and with the Committee on National Security Systems Instruction (CNSSI) 1253[33] for national security systems, where applicable. Finally, the program manager (PM) and information owner should categorize the system data by determining the appropriate security objectives and impact level in accordance with FIPS 199, "Standards for Security Categorization of Federal Information and Information Systems."[34]

[28] Note: This Policy implements DHS Management Directive 140-01, "Information Technology Security Program."
[29] U.S. Department of Homeland Security (DHS) Sensitive Systems Policy Directive 4300A, Version 13.1, July 27, 2017.
[30] U.S. Department of Homeland Security (DHS). Sensitive Systems Policy Directive 4300A, Version 13.1, October 2, 2019.
[31] National Institute of Standards and Technology (NIST) Special Publication (SP) 800-60, Volume I: Guide for Mapping Types of Information and Information Systems to Security Categories, Revision 1, August 2008. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v1r1.pdf.
[32] National Institute of Standards and Technology (NIST) Special Publication (SP) 800-60, Volume II: Appendices to Guide for Mapping Types of Information and Information Systems to Security Categories, Revision 1, August 2008. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-60v2r1.pdf.
[33] Committee on National Security Systems (CNSS) Instruction 1253, Security Categorization and Control Selection for National Security Systems, March 27, 2014. https://www.cnss.gov/CNSS/openDoc.cfm?n4wWvsdOQfNKeOqvvE1XSQ==.
[34] Federal Information Processing Standards (FIPS) Publication 199, "Standards for Security Categorization of Federal Information and Information Systems," February 2004. https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.199.pdf.

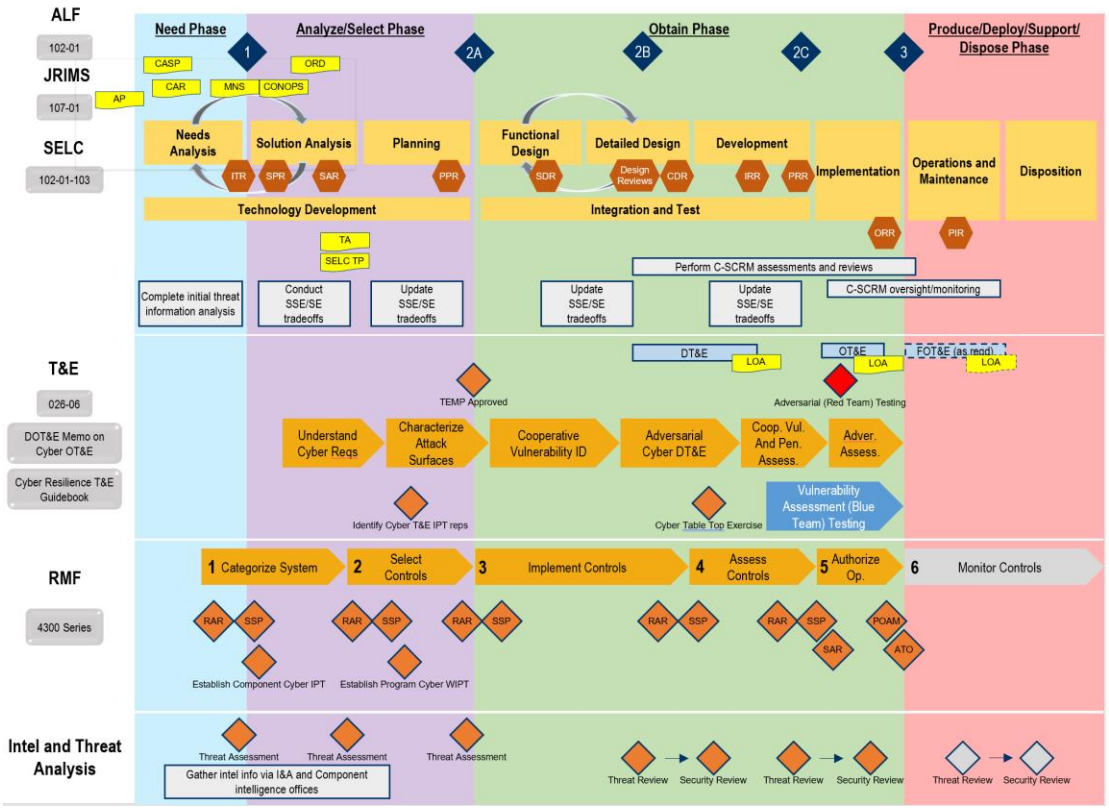

Figure A.1. DHS Cybersecurity and Acquisition Lifecycle Tool.

DHS Cyber Resilience T&E activities assess a system's resilience to cyberattacks. Figure A.2 presents the DHS Cyber Resilience T&E Strategy for Cyber Resilience[35], consisting of four primary activities and several supporting activities. An essential part of planning for Cyber Resilience T&E is understanding the operational requirements. While the categorization of systems is not mentioned in the first primary activity, it is important that PMs and information owners understand the RMF Step 1 activities for categorization and consider including these in the early stages of the T&E planning phase (e.g., Understanding Cyber Resilience Requirements). The program requirements sponsor should also be fully integrated into the development of cyber resilience requirements, especially to address mission requirements for systems that must operate in contested environments. Ideally, the Understanding Cyber Resilience Requirements phase should align with the ALF and begin earlier in the "Need" phase. As it currently stands, the DHS T&E Strategy for Cyber Resilience sets "Understanding Cyber Resilience Requirements" in alignment with the "Analyze/Select" phase of the ALF, coinciding with RMF-related artifacts that should have already been documented. Understanding Cyber Resilience Requirements is an iterative activity, and it may continue throughout the Analyze/Select and Obtain ALF Phases.

[35] U.S. Department of Homeland Security (DHS). Science and Technology (S&T). Cyber Resilience Test and Evaluation (T&E) Supplemental Guidance, v2.0, October 2019.

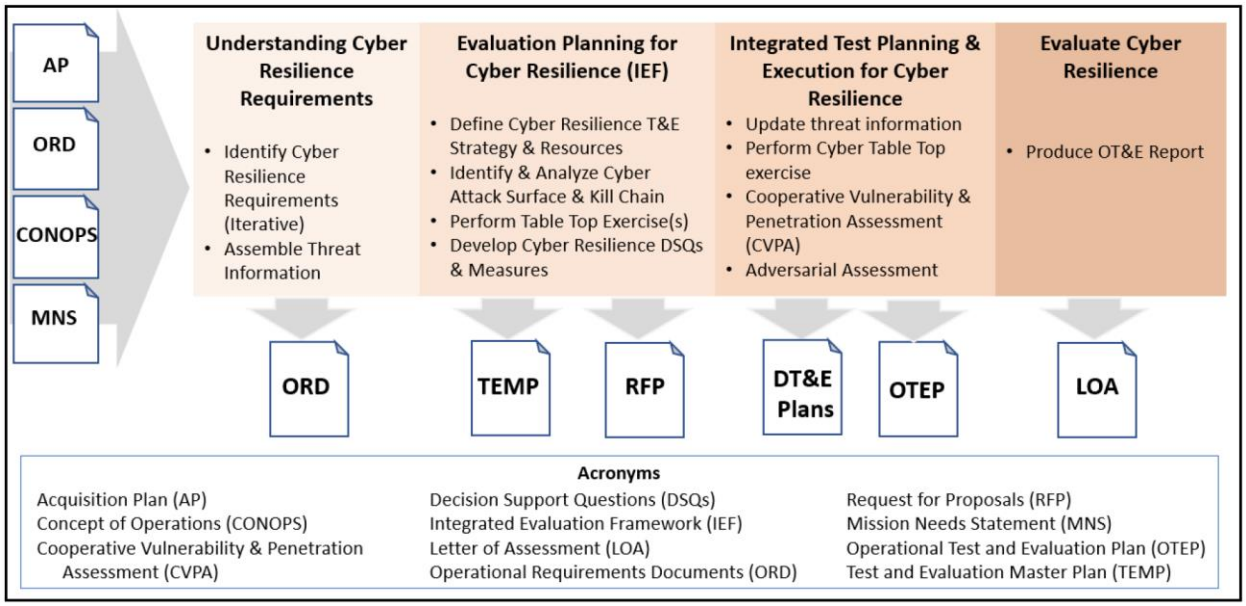

Figure A.2. DHS T&E Strategy for Cyber Resilience.

CWMD is tasked with protecting the United States from CBRN threats by utilizing CBRN detection device technologies. To effectively fulfill their mission, CWMD collects and analyzes vast amounts of data from diverse sources, including CBRN equipment in the field, intelligence agencies, law enforcement, customs, and state and local municipalities. Advanced data categorization techniques, such as data categorization models (DCMs) and semantic analysis, are utilized to categorize the collected data into relevant categories. This may include patterns indicative of CBRN threats or security concerns. The captured data is analyzed for categorization to assess potential threats, emerging trends, and anomalous behaviors that may be indicative of imminent threats. A DCM is used to integrate and cross-reference categorized data from multiple pieces of CBRN equipment to identify connections, networks, and patterns and to convey the relationships to intelligence agencies, law enforcement, customs, and state and local municipalities. The DCM can prioritize threats and risks based on their severity, likelihood, and potential impact on national security. This also enables the allocation of resources and focused efforts on addressing high-priority threats while considering strategic incident response activities. Decision support is paramount between CWMD and other entities to provide actionable intelligence and insights for supporting informed decision-making, resource allocation, and operational planning. The DCM presents findings through visualizations, reports, and dashboards to facilitate the understanding and communication of complex CBRN data. The DCM will enable the categorization and analysis of CBRN data sources and enable the detection of emerging threats and security vulnerabilities before they escalate, allowing for proper mitigation measures to be put in place. By identifying and prioritizing anomalous data categorization, CWMD can allocate resources, personnel, and additional CBRN needs more efficiently. By leveraging a DCM, CWMD will further enable the protection of the United States from CBRN threats.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov