A Smart City Infrastructure Ontology for Threats, Cybercrime, and Digital Forensic Investigation

Yee Ching Tok
Singapore Univ. of Tech. and Design
Singapore
yeeching_tok@sutd.edu.sg

Davis Yang Zheng
Singapore Univ. of Tech. and Design
Singapore
davis.zheng@owasp.org

Sudipta Chattopadhyay
Singapore Univ. of Tech. and Design
Singapore
sudipta_chattopadhyay@sutd.edu.sg

***Abstract*—Cybercrime and the market for cyber-related compromises are becoming attractive revenue sources for state-sponsored actors, cybercriminals and technical individuals affected by financial hardships. Due to burgeoning cybercrime on new technological frontiers, efforts have been made to assist digital forensic investigators (DFI) and law enforcement agencies (LEA) in their investigative efforts.**

Forensic tool innovations and ontology developments, such as the Unified Cyber Ontology (UCO) and Cyber-investigation Analysis Standard Expression (CASE), have been proposed to assist DFI and LEA. Although these tools and ontologies are useful, they lack extensive information sharing and tool interoperability features, and the ontologies lack the latest Smart City Infrastructure (SCI) context that was proposed.

To mitigate the weaknesses in both solutions and to ensure a safer cyber-physical environment for all, we propose the Smart City Ontological Paradigm Expression (SCOPE), an expansion profile of the UCO and CASE ontology that implements SCI threat models, SCI digital forensic evidence, attack techniques, patterns and classifications from MITRE.

We showcase how SCOPE could present complex data such as SCI-specific threats, cybercrime, investigation data and incident handling workflows via an incident scenario modelled after publicly reported real-world incidents attributed to Advanced Persistent Threat (APT) groups. We also make SCOPE available to the community so that threats, digital evidence and cybercrime in emerging trends such as SCI can be identified, represented, and shared collaboratively.

I. Introduction

Cybercrime and the market for cyber-related compromises are becoming attractive revenue sources for state-sponsored actors, cybercriminals and technical individuals affected by financial hardships. The expected economic impact of cybercrime will reach US$10.5 trillion by 2025 [1], making it an appealing choice compared to riskier ventures such as drug trafficking and piracy. Cybercrime and cyber threats were also critical concerns on the international front, ranking 8th out of 32 global risks in severity from short and long-term perspectives [2].

Due to burgeoning cybercrime on new technological frontiers, efforts have been made to assist digital forensic investigators (DFI) and law enforcement agencies (LEA) in their investigative efforts. Tools such as STITCHER [3] and smart city infrastructure (SCI) related threat models [4] have been proposed as possible solutions to correlate evidence, identify threats and aid in digital forensic investigations in these nascent areas. Meanwhile, in a bid to enhance information sharing and tool interoperability, ontologies such as the Unified Cyber Ontology (UCO) [5] and the Cyber-investigation Analysis Standard Expression (CASE) [6] were suggested. Although the previously mentioned tools and ontologies are helpful, some inadequacies exist. The tools lacked extensive information sharing and tool interoperability features, while the ontologies lacked the latest SCI context that was proposed [4].

To mitigate the weaknesses in both solutions highlighted previously and to ensure a safer cyber-physical environment for all, we propose the Smart City Ontological Paradigm Expression (SCOPE), an expansion profile of the UCO and CASE ontology that implements the research information obtained from the SCI threat model [4]. As a start, various cyber threats are categorized under Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service and Elevation of Privilege (STRIDE). The SCI data structures within SCOPE are technology-agnostic while adhering to international standards such as ISO37101:2016, ISO37120:2018, ISO37122:2019 and ISO37123:2019 [7]–[10]. Finally, digital forensic evidence related to SCI threats identified via threat modelling is also included in SCOPE.

The contributions of our research are summarized as follows:

1) We observed the data model used in UCO and CASE, and proposed a SCI-focused ontological profile known as SCOPE. Following the goals of UCO and CASE, SCOPE supports coordinated cyber investigations and tool interoperability. Furthermore, SCOPE also aligns with the cybersecurity industry and practitioners by integrating attack techniques, attack patterns and classifications from MITRE.
2) We showcase how SCOPE could be used to present complex data such as SCI-specific threats, cybercrime, investigation data, and incident handling workflows via an incident scenario modelled after publicly reported real-world incidents attributed to Advanced Persistent Threat (APT) groups. We also further compare SCI representations using SCOPE and UCO/CASE, showcasing the ease of usage if SCOPE is adopted into UCO and CASE.
3) We make available SCOPE to the community so that threats, digital evidence and cybercrime in emerging

trends such as SCI can be identified, represented, and shared in a collaborative manner.

For reproducibility and advancing the research in SCI, UCO and CASE, SCOPE and its associated contents are publicly available at: https://github.com/scopeProject.

The rest of this paper is organized as follows. In Section II, we present the background and motivation of the paper. In Section III, we showcase the structure of SCOPE and the various examples when SCOPE is used for digital forensic investigations. In Section IV, we evaluate SCOPE and showcase its differences from existing ontologies. In Section V, we highlight the limitations of our research and SCOPE. In Section VI, we summarize current related work. Finally, we conclude the paper in Section VII.

II. BACKGROUND AND MOTIVATION

Digital forensic investigations require not only technical training but also background knowledge related to the investigated platforms. This is to complete their investigations effectively. In a user study conducted by Tok et al. [3], it was observed that DFI who had undergone professional training or worked on prior hands-on investigations performed better when faced with the given domain-specific forensic scenario. However, SCI is an emerging area with multiple forms of interpretation of the area [4]. Freshly graduated DFI from Institutes of Higher Learning (IHL) may also need help in SCI investigation as the breadth and depth of teaching and assessments vary between various IHLs. This observation is further supported by the user study conducted by Tok et al. [3], where more than half of the participants that studied digital forensics in an IHL *failed to solve the given forensic scenario*. This presents a worrying scenario - DFI *might* be overwhelmed when attacks on SCI become prevalent, as they already face multiple challenges in their investigation work in traditional domains [3].

A. Sharing Forensic Data in Emerging Technologies

Other than digital forensic data, DFI and LEA also have to contend with different types of data, such as digital threats and types of cybercrime committed. Traditional commercial digital forensic tools such as EnCase [11] and FTK [12] dealt solely with data management and forensic processes on conventional platforms such as desktop computers, laptops and mobile devices. However, these tools do not include other data that would provide context. Current electronic discovery (e-Discovery) tools only work on traditional platforms, but no support for SCI-related platforms is mentioned [4].

Digital investigations may also involve multiple parties and evidence formats, especially in the case of complex systems such as SCI. As such, collaborative investigations, data-sharing and tool updates/innovations are inevitable requirements. File formats such as the Advanced Forensic Format (AFF4) could be used by tools to store some digital forensic metadata via the Resource Description Framework (RDF) [13]. Information-sharing frameworks such as Structured Threat Information Expression (STIX) and Trusted Automated Exchange of Intelligence Information (TAXII) have also been proposed but were focused more on cyber threat intelligence [14]. Meanwhile, digital forensic innovations and tool updates may take time. For example, a widely used tool in digital forensics, *bulk_extractor*, took longer than anticipated for a subsequent update [15]. However, the end result yielded impressive performance optimizations and significant improvements in code quality and reliability [15].

B. Representing Forensic Data in Emerging Technologies

There is a clear technical gap for representing and sharing data, particularly threats, cybercrime, and digital forensic information for emerging technologies such as SCI platforms. Some solutions have been proposed, such as the Unified Cyber Ontology (UCO) and the Cyber-investigation Analysis Standard Expression (CASE) ontology. UCO and CASE could be used to share threats, cybercrime and digital forensic information. A few example implementations for popular digital forensic tools such as Cellebrite, Magnet Forensics and MSAB XRY have been produced [16]–[18]. However, UCO and CASE ontologies are not geared towards SCI. In fact, CASE is an extension of UCO, which includes investigation-specific ontologies. Noting the challenges DFI are grappling with regarding investigation workloads and background knowledge, it would be hard to expect DFI to adapt UCO and CASE immediately to an SCI context. Therefore, we decided to build on our previous research data [4] and create a modular expansion that centers around SCI named Smart City Ontological Paradigm Expression (SCOPE). By creating this expansion, we offload the time-consuming process of creating an SCI investigation framework and associated SCI evidence artifact representation. DFI can spend more time investigating and clearing cases instead of grappling with uncertain variables such as ad-hoc forensic data representation in emerging technologies.

C. Supporting Cyber Defense in Emerging Technology Deployments

As emerging technologies and hardware are consolidated and integrated into SCI deployments on a city-to-national level, best practices such as product certifications and security testing would be expected. However, such certifications and security testing are merely product-specific and do not account for scenarios where various hardware components, protocols and frameworks are integrated on a city-to-national level. Governments looking to implement SCI would be concerned about the resiliency of SCI against cyberattacks from various adversaries (such as cybercriminals and APTs). Advanced security assessment activities such as adversary emulation [19] would be an excellent way to assess SCI deployments comprehensively. Such activities are driven by attack techniques cataloged by organizations (e.g., MITRE) and acknowledged by the cybersecurity industry. To facilitate cybercrime investigation on a city-to-national level, SCOPE has fully integrated the MITRE ATT&CK framework [20] and

MITRE Common Attack Pattern Enumerations and Classifications (CAPEC) [21]. By doing so, investigators and defenders can instantly map attack techniques and patterns utilized by adversaries and associated threats/digital forensic evidence during their investigation. Coupled with the ability to share attacks and digital forensic evidence, SCOPE provides cyber defenders opportunities to enhance detection and response capabilities as information recorded via SCOPE could be utilized to improve policies, procedures and technologies while adversary emulation activities are ongoing.

III. Smart City Ontological Paradigm Expression

The Smart City Ontological Paradigm Expression (SCOPE) extends the prior work of UCO and CASE, and is specified via Web Ontology Language OWL 2 (following the documentation from CASE [22]). UCO was designed to support information representation across multiple cyber domains, while CASE was an investigation-specific module extended from UCO [6]. SCOPE aims to provide a SCI-specific extension that can be used by any interested users while maintaining the collaborative and interoperability nature of UCO/CASE.

A. Motivations for the Creation of SCOPE

Initially, we envisioned using UCO/CASE based on its ontological structure and expanding it slightly for SCI-related digital forensic activities. However, it became evident that it would be challenging as smart city-specific terminologies, attack techniques, and patterns were absent within UCO/CASE. That meant extensive modification and expansion would be required if DFI adopted the usage of UCO/CASE. A one-off effort was possible, but it had its drawbacks and limitations. DFI need to know SCI intimately, along with associated threats, attack techniques, sources of evidence and related forensic artifacts. Unfortunately, based on prior research findings [3], [4], we have established that DFI are hard-pressed for time and unlikely to be able to spend time juggling casework and learning new knowledge concurrently.

To mitigate the impending problem, we concluded that a modular extension that abides by the ontological design of UCO/CASE but contains all the necessary information for SCI terminologies, attack techniques, patterns and forensic artifacts should be created. Table I shows the additional features offered by SCOPE by extending the prior work of UCO and CASE. By creating SCOPE, we offer DFI tasked with investigating SCI cybercrime substantial time savings (from repetitive literature review, research and conflicting opinions) and a comprehensive framework encompassing cyber threats, crimes, attack techniques and digital evidence.

When we explored UCO/CASE, we noticed some description entries utilized sources such as Wikipedia for some of the technical descriptions [23], [24]. SCOPE avoids this and pulls references from established sources such as ISO standards which are not as easily mutable to illegitimate changes. SCOPE also extends terminologies to cover emerging technology as an effort to reduce technological obsolescence.

TABLE I
Comparison of Features Between UCO, CASE and Scope

Features	UCO	CASE	SCOPE
Extensible	✓	✓	✓
Information Representation for Cyber Domains	✓	✓	✓
Supports Cyber Investigations	✗	✓	✓
Smart City Infrastructure Terminology Representation	✗	✗	✓
Smart City Infrastructure Threats Representation	✗	✗	✓
Smart City Infrastructure Cybercrime Classifications	✗	✗	✓
Smart City Infrastructure Data Sources	✗	✗	✓
Smart City Infrastructure Digital Evidence Sources	✗	✗	✓
MITRE ATT&CK Framework Representation	✗	✗	✓
MITRE CAPEC Framework Representation	✗	✗	✓

B. Overview of SCOPE

We provide an expanded ontology on SCI-focused cybercrime, digital evidence, data indicators, infrastructure, roles and threats. Users could leverage SCOPE to represent threats, cybercrime, and digital forensic information related to SCI. Table II shows the key elements of SCOPE Ontology.

C. SCOPE Developments

Since the preliminary version of SCOPE, as illustrated in the preceding section, additional amendments and extensions were incorporated during the iterative process of improving and scenario-based testing. Such an iterative process was followed as a natural step to provide coverage over any areas that were found lacking. Table III shows the added extensions to SCOPE after further refining our work.

During the conceptualization and further optimization of SCOPE, we ensure that the proposed ontology is developed logically and is universally acceptable. Based on our extensive literature review of ontology design principles, we followed the recommended steps and considerations outlined in prior work [25]. To further ensure the rigor of SCOPE, two researchers examined the design principles and steps outlined in the prior work of Noy and Mcguinness [25] and ensured that SCOPE (via mutual checks and consensus) was developed following the considerations. A graphical representation of the recommended steps and considerations is summarized and shown in Figure 1.

TABLE II
SCOPE ONTOLOGY SUMMARY

SCOPE v0.1.0	**Description**	**Justification**
scope-crime	An action which constitutes an offense and is punishable by law (in SCI context).	A common set of SCI-related cybercrime are listed here.
scope-evidence	Additional observables (evidence in our case) that is present in SCI and Internet-of-Things (IoT) devices that are currently not listed in uco-observable are placed here.	This ontology follows the definition of Smart City Infrastructure (SCI) presented in this paper [4].
scope-indicators	These are the data indicators used in SCI. For ease of reference, the indicators are specified by their respective ISO clause number and comments specifying the details of the indicator.	This ontology defines the indicators used in Smart City Infrastructure (SCI). For a more detailed explanation on how the indicators are grouped, please refer to the paper [4].
scope-infrastructure	This ontology defines infrastructure within a Smart City Infrastructure (SCI). The definition of SCI defined in this paper [4] is preferred, although additional infrastructure types can be added in future to suit deployment needs or as SCI gets more mature.	A technology-agnostic definition of SCI is provided.
scope-role	This ontology defines roles present in a Smart City Infrastructure (SCI) cybercrime investigation, identification of digital forensic opportunities and threat modeling.	These roles do not exist in UCO/CASE.
scope-threats	This ontology defines threats that are identified within Smart City Infrastructure (SCI).	These SCI-specific threats are not listed inside UCO/CASE.
scope-vocabulary	Vocabularies used in Smart City Infrastructure (SCI)	These SCI-related vocabulary and attacks are not listed inside UCO/CASE.

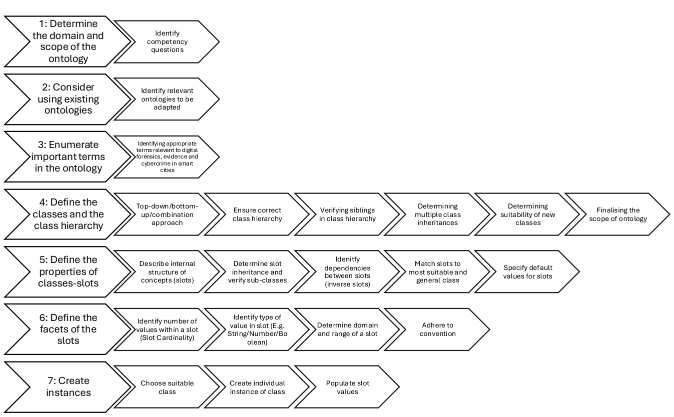

Fig. 1. SCOPE Ontology Design Steps (adapted from [25])

TABLE III
SCOPE ONTOLOGY EXTENSIONS

SCOPE v0.1.1	**Description**	**Extensions**
scope-attackpatterns	The Common Attack Pattern Enumeration and Classification (CAPEC) from MITRE was integrated to provide a publicly available catalog of common attack patterns that helps users understand how adversaries exploit weaknesses in applications and other cyber-enabled capabilities.	New entries that defines the attack patterns used in Smart City Infrastructure (SCI).
scope-indicators	Data Indicators used in SCI. Referencing from ISO standards as well as additional definitions from relevant government bodies.	Additional entries were further added to provide minute distinctions between the various indicators used.

In steps 1 and 2 (with reference to Figure 1), the primary objectives were to determine the domain and scope of the ontology, and also to consider using existing ontologies. The original structure of UCO/CASE was certainly robust and allowed further extensions, but a great deal of effort would be expanded to account for artifacts related to smart cities. As such, we determined an extension ontology (SCOPE) was the right way forward to benefit forensic investigators. In the next section, we discuss how the next design steps of SCOPE ontology were systematically followed.

D. SCOPE Construction and Protégé

In the development of SCOPE, we have employed Protégé, an open-source ontology editor and a framework for building intelligent systems [26]. We have chosen to use Protégé as it provides numerous benefits while building SCOPE.

First and foremost, Protégé delivers a comprehensive tool which allows for the easy development and visualization of numerous complex relationships between the various facets of SCOPE. This capacity enables us to derive a more complete and accurate representation from the complex world of SCI-based digital forensics, and facilitates steps 3 and 4 (with reference to Figure 1) of the ontology design steps.

Secondly, as Protégé supports the OWL 2 Web Ontology Language [27] and RDF specifications from the World Wide Web Consortium (W3C) such as Resource Description Framework Schema (RDFS) [28], Web Ontology Language (OWL) [29], Semantic Web Rule Language (SWRL) [30] and Turtle (TTL) [31]. The wide range of support ensures the compatibility and interoperability of SCOPE with other related ontologies such as UCO and CASE, thus enhancing the applicability of SCOPE.

Thirdly, the collaborative features provided by Protégé enable participation from all key stakeholders, allowing a participatory development approach (with reference to Figure 1, this would be steps 5 to 7). This enables experts from diverse backgrounds to join in the iterative development process, allowing them to lend their expertise to SCOPE's development, increasing its flexibility for various use cases.

Fourthly, based on the guidelines suggested by Gruber [32], the vocabulary and ontological commitments of SCOPE were validated, and Figure 2 shows a sample of ABox reasoning of SCOPE. Furthermore, two researchers performed the verification work individually and reached a consensus for the final output of SCOPE.

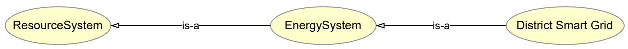

Fig. 2. SCOPE ABox Reasoning Example

Last but not least, Protégé's extensible and open-source nature provides an accessible plug-and-play environment. This makes it a flexible base for rapid prototyping and application development, allowing for customized extensions that cater to the specific needs of SCOPE development, including but not limited to specialized visualization tools. We leveraged Protégé to visualize the proposed SCOPE ontology, annotate various SCI components, threats and cybercrime, and built the conceptual map (shown in Figure 3). We further used SCOPE to annotate the evaluation scenarios used in Section IV, thus testing out the ABox reasoning for SCOPE. Additionally, as SCI matures, amendments and updating of SCOPE will be required due to the evolution of threats, cybercrime and digital evidence. Using both SCOPE and Protégé, users can leverage the synergy to construct an updated knowledge base that can assist with ABox reasoning. Such knowledge bases could accelerate cybercrime investigation by providing context to new investigators and insight into SCI's complexities.

In summary, the benefits mentioned in the preceding paragraph underpin the rationale of using Protégé, contributing to creating a scalable, interoperable, and comprehensive ontology that supports coordinated cyber-investigations in smart city infrastructure.

1) Conceptual Ontology Map: Using Protégé and with reference to Figure 3, we developed a conceptual map to provide a clear understanding and overview of the structure of SCOPE. This map visually represents the core classes, object properties, and relationships, highlighting the hierarchical organization and interdependencies among key concepts such as threats, vulnerabilities, and assets. By illustrating these elements and their interactions, the map serves to clarify the ontology's design, allowing for a more accessible and intuitive grasp of its complex structure and application within digital forensics and cybersecurity research.

Fig. 3. SCOPE Smart City Conceptual Map

2) Descriptive Statistics of Ontological Concepts: To provide a comprehensive overview of SCOPE 's structural elements, we include key descriptive statistics that detail its complexity. Table IV shows a brief summary of relevant ontological information about SCOPE. The ontology comprises 1,383 classes, each representing critical concepts in cybersecurity, with 180 object properties defining relational dynamics between these classes. Additionally, the ontology includes 587 data properties, enriching the model with specific characteristics for each class and 1,417 individuals, which instantiates specific entities within the ontology.

Regarding logical structure, the ontology contains 1,462 subclass axioms and three disjoint classes, ensuring logical consistency and explicit boundaries between mutually exclusive categories. Furthermore, the ontology includes extensive annotation properties, with 15,260 annotation assertions that provide supplementary metadata, including descriptions, labels, and sources for improved interpretability. The detailed SCOPE ontology is publicly available in our website: https://ontology.scopeontology.org.

TABLE IV
SCOPE ONTOLOGY METRICS TABLE

Metrics	
Axiom	26,621
Logical axiom count	3,748
Declaration axioms count	2,200
Class count	1,383
Object property count	180
Data property count	587
Individual count	1,417
Class Axioms	
SubClassOf	1,462
DisjointClasses	3
Object Property Axioms	
SubObjectPropertyOf	5
InverseObjectProperties	2
InverseFunctionalObjectProperty	1
TransitiveObjectProperty	2
IrrefexiveObjectProperty	1
ObjectPropertyDomain	6
ObjectPropertyRange	177
Data Property Axioms	
DataPropertyRange	586
Individual Axioms	
ClassAssertion	1,447
Annotation Axioms	
AnnotationAssertion	15,260

E. SCOPE *Examples*

In this section, we showcase some theoretical examples of how SCOPE can be used for SCI-related threats and investigations. Figure 4 shows an overview of the status quo and lists out the additional capabilities SCOPE can offer to DFI. For readability and illustrative purposes, the universally unique identifier (UUID) for each object is referenced with the associated system/object name (e.g., resource-system-uuid). We also import the UCO and CASE ontologies for use within our SCOPE ontology where needed.

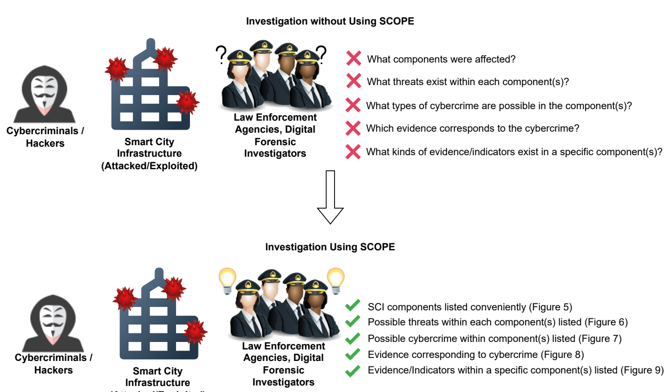

Fig. 4. Overview of Investigation With and Without Using SCOPE

With reference to Figure 5, we showcase an example of SCI represented in SCOPE format. The system layer (sometimes known as the context layer) components of a suggested SCI [4] are outlined for ease of future use. The identifiers can be used to tag forensic evidence to the corresponding systems affected by cybercrime or map discovered attack techniques on the component.

```
{
    "@id": "smart-city-infrastructure-uuid",
    "@type": "scope-infrastructure:SmartCityInfrastructure",
    "uco-core:name": "Sample_Smart_City_Infrastructure",
    "scope-infrastructure:SmartCityInfrastructure": [
        {
            "@id": "citizen-service-system-uuid"
        },
        {
            "@id": "essential-service-system-uuid"
        },
        {
            "@id": "livelihood-support-system-uuid"
        },
        {
            "@id": "resource-system-uuid"
        }
    ]
}
```

Fig. 5. Example of SCI in SCOPE Format

We can also use SCOPE to showcase the types of possible threats in SCI. For example, some threats applicable to the Resource System component of a SCI are illustrated in Figure 6. The component and threats are further assigned identifiers for ease of usage when evidence corresponding to threats and the Resource System component is discovered.

```
{
    "@id": "resource-systems-threat-assessment-uuid",
    "@type": "scope-threats:Threat",
    "uco-core:name": "resource-system-threats",
    "scope-infrastructure:ResourceSystem": {
        "@id": "resource-system-uuid"
    },
    "scope-threat:DataFlowInterruption": {
        "@id": "data-flow-threat-uuid"
    },
    "scope-threat:DataFlowSniffing": {
        "@id": "data-flow-sniffing-threat-uuid"
    },
    "scope-threat:DataStoreSpoofing": {
        "@id": "data-store-spoofing-threat-uuid"
    },
    "scope-threat:WeakAccessControl": {
        "@id": "weak-access-control-threat-uuid"
    }
}
```

Fig. 6. Types of Possible Threats Within Resource System Component of SCI

Using SCOPE, we then represent the types of cybercrime (with reference to Figure 7 and from prior research [4]) that could happen within a component of the SCI (Resource System component in this example). In this case, possible cybercrime, such as data and system interference and illegal access and interception, are shown.


```
{
    "@id": "resource-system-crime-identification-uuid",
    "@type": "scope-crime:Cybercrime",
    "uco-core:name": "resource-system-cybercrime",
    "scope-infrastructure:ResourceSystem": {
        "@id": "resource-system-uuid"
    },
    "scope-crime:DataInterference": {
        "@id": "data-interference-crime-uuid"
    },
    "scope-crime:IllegalAccess": {
        "@id": "illegal-access-crime-uuid"
    },
    "scope-crime:IllegalInterception": {
        "@id": "illegal-interception-crime-uuid"
    },
    "scope-crime:SystemInterference": {
        "@id": "system-interference-crime-uuid"
    }
}
```


Fig. 7. Types of Possible Cybercrime Within Resource System Component of SCI

We could also use SCOPE to map the evidence captured or needed for the respective cybercrime. For instance, it was discovered that a data interference cybercrime occurred within the Resource System of a SCI. Further investigation indicated that an internet-facing hardware firmware component was compromised, and the component's manufacturer and Media Access Control (MAC) address was recorded. The corresponding details are presented in Figure 8).


```
{
    "@id": "resource-system-evidence-identification-uuid",
    "@type": "scope-evidence:ResourceSystemEvidence",
    "uco-core:name": "resource-system-cybercrime",
    "scope-infrastructure:ResourceSystem": {
        "@id": "resource-system-uuid"
    },
    "scope-crime:DataInterference": {
        "@id": "data-interference-crime-uuid"
    },
    "uco-core:hasFacet": [
            {
                "@id": "resource-system-firmware-uuid",
                "@type": "scope-evidence:Firmware",
                "uco-observable:manufacturer": {
                    "@id": "manufacturer-uuid"
                }
            },
            {
                "@id": "resource-system-wifi-mac-uuid",
                "@type": "uco-observable:WifiAddressFacet",
                "uco-observable:addressValue": "aa:bb:cc:11:22:33"
            }
        ],
}
```


Fig. 8. Digital Evidence Mapped to Cybercrime Within Resource System Component of SCI

Finally, with reference to Figure 9, we can use SCOPE to collate the data indicators within a specified SCI system. All data indicators related to the Energy System (from the Resource System grouping of SCI and referencing ISO37120 [8]) are grouped and assigned individual identifiers. This allows a quick enumeration of all associated data indicators of a larger system, which could further assist investigators to ascertain about the affected data outputs when SCI is subject to an attack or cybercrime.


```
{
    "@id": "energy-system-data-indicator-query-uuid",
    "@type": "scope-infrastructure:EnergySystem",
    "uco-core:name": "energy-system-indicators",
    "scope-infrastructure:EnergySystem": {
        "@id": "energy-system-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointEightPointOne": {
        "@id": "indicator1-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointEightPointTwo": {
        "@id": "indicator2-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointFive": {
        "@id": "indicator3-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointFive": {
        "@id": "indicator4-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointFour": {
        "@id": "indicator5-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointOne": {
        "@id": "indicator6-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointSeven": {
        "@id": "indicator7-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointSix": {
        "@id": "indicator8-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointThree": {
        "@id": "indicator9-uuid"
    },
    "scope-indicator:ISO37120ClauseSevenPointTwo": {
        "@id": "indicator10-uuid"
    }
}
```


Fig. 9. Data Indicators Within Resource System Component (Energy System) of SCI

In summary, the examples using SCOPE demonstrate the additional SCI-related functionalities added to UCO/CASE. By adapting SCOPE into the UCO/CASE ontology instances, DFI and LEA can immediately represent SCI-related threats, cybercrime and data components without additional time to construct SCI-related components. In addition, the SCI represented within SCOPE is technology-agnostic, thus making it highly compatible with diverse SCI systems.

IV. EVALUATION OF SCOPE

For the purposes of determining the efficacy of SCOPE, we developed scenarios that emulate real-life investigations. Such scenarios would take place during a cybercrime investigation in a smart city, particularly those that address the aims and purposes of the United Nations Sustainable Development Goals and data indicators outlined by Tok and Chattopadhyay [4]. The large overarching scenario is split into three smaller scenarios representing key activities during a cybercrime investigation - ① initial overview of the scenario and its ontology representation, ② incident investigation with

Tactics, Techniques and Procedures (TTPs) identified during the examination and ③ containment and recovery using identified Indicators of Compromise (IoC).

These scenarios come with corresponding evidence necessary for the evaluation exercise. The evidence is manually mapped by researchers to the corresponding ontologies being evaluated. In the case of SCOPE, the researchers also manually mapped threats and evidence to the ATT&CK techniques and CAPEC attack patterns. To ensure consistency and to validate this mapping, two researchers independently performed the mapping and reached a consensus. The digital evidence in the scenarios is presented in both ontologies (UCO/CASE and SCOPE) to demonstrate investigation and collaboration workflow optimizations for DFIs. To compare SCOPE with a baseline, we manually extended the UCO/CASE ontology used in the evaluation scenario to incorporate SCI elements. Such efforts would require DFI involved in the investigation to be proficient in SCI threats, digital evidence types, attack techniques and classifications. Moreover, significant modifications and time are required to prepare this baseline. This reflects the need for a novel extension from UCO/CASE that covers SCI use cases, and DFI can directly adopt it for cybercrime-related incidents in SCI. We provide a visual summary of the evaluation exercise, along with salient points where SCOPE empowers DFI in Figure 10.

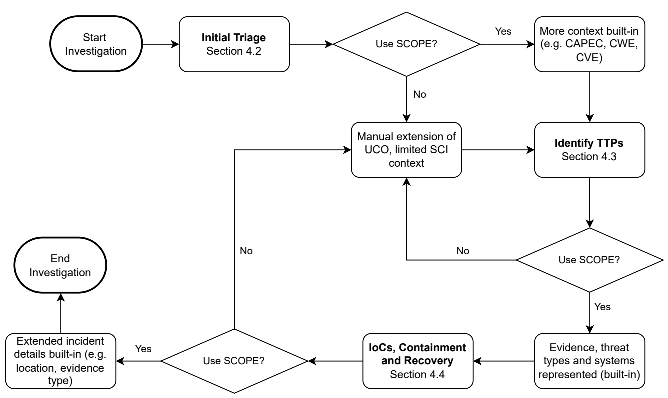

Fig. 10. Evaluation Scenario Flowchart

A. SCOPE *Scenario Overview*

Our proposed scenario is based on an actual APT Group (APT41, alternatively known as Brass Typhoon or Barium) and draws on several TTPs listed by MITRE [33]. The scenario is instantiated geographically in Singapore and in a township (Punggol) that has been announced as a smart district publicly (Punggol Digital District) [34]. APT41 was chosen as the reference adversary for the scenario as it targets high-technology industries, and some of its activities have been attributed to victims in Singapore [35].

In the context of the scenario, we have established a few parameters. Firstly, just like the current events, cybersecurity is a fast-paced and constantly evolving field with new technological developments and obsolescence occurring. Homonyms are increasingly common in computer science and cybersecurity, where exact definitions and standardization are essential for collaboration and investigation. Furthermore, tracking the growth of IoT development and the numerous sensors and devices being developed is increasingly difficult.

As such, the second parameter will be that first responders will find it challenging to identify and communicate what has already been discovered. This parameter is also exacerbated by the fact that the new Punggol Digital District had been built to develop and integrate smart solutions, resulting in an oversized amalgamation of interconnected technology systems that are built towards the envisioned Smart City as defined by Tok and Chattopadhyay [4].

In the final parameter, we presume that an incident investigation from a professional Cybersecurity Incident Response Team (CIRT) adheres to an established framework, such as those from the National Institute of Standards and Technology (NIST) [36], [37]. It usually follows the same process and procedures, such as evidence acquisition, chain of custody and investigations in an analysis facility while striving to achieve cybersecurity outcomes such as detection, response and recovery. Most CIRTs will maintain similar standard operating procedures with some workflow tweaks depending on their mission, goals and requirements.

1) Prologue: On 1st January 2100, the Singapore Computer Emergency Response Team (SingCERT) was notified from various organizations and companies situated at the Punggol Digital District that they were struck by a ransomware attack. APT Triple Dragon has claimed responsibility for the attack and has demanded a ransom payment of US$250 million to be made in 7 days.

B. *Scenario 1 - Initial Triage (1st January 2100)*

SingCERT dispatched a team of cyber incident responders equipped with jump kits and commenced initial triage of the incident. The incident responders performed the following actions:

1) **Imaging affected devices.** The incident responders used Tableu Forensic Imager TX1 to image the storage devices of infected computers for endpoint-based evidence.
2) **Initial analysis.** As a complete image of devices could take time, an initial assessment is also required to determine the impact and extent of damage. Using specially configured portable removable devices with triaging tools such as memory capture and assorted forensic artifact acquisition tools, the responders performed initial on-the-spot analysis.
3) **Network traffic analysis.** Relying on endpoint logs and device artifacts in a ransomware incident could be insufficient. Network packet capture and network traffic analysis provide a different perspective in investigations as valuable data, such as domains contacted and remote commands, could be captured to gain a deeper insight into the incident.
4) **Interim assessment.** Based on the previous steps, the incident responders gather, collaborate and corroborate

their findings to make an interim assessment of the incident.

The responders discovered that the extent of the infection was far larger than initially assessed, as the ransomware identified as Tiki could propagate via IoT systems. Thus, in an effort to be meticulous, SingCERT has directed the seizure of all systems that may be linked to the attack.

1) Lab Investigation: After seizure and retrieval of the impacted systems, logs collected were fed into a standalone computer and processed via log ingestion tools such as the Elastic, Logstash and Kibana (ELK) stack [38] and Splunk [39] for analysts to trawl through. A chain of custody was established and maintained for all seized digital evidence, ensuring the evidence was admissible in court. Additional bit-by-bit duplication was performed on the large data storage devices that could not be acquired during the initial triage due to their size. After the investigation efforts, analysts identified and extracted the ransomware from the infected systems, identifying it as Encryptor RaaS ransomware.

Throughout the attack, the perpetrator employed sophisticated, tailor-made, and elusive tools that successfully bypassed and circumvented the antivirus programs and traditional security measures. The malware instances examined by SingCERT turned out to be any of the following: 1) novel variants previously unidentified in the wild, 2) undetected by the conventional anti-malware systems utilized by the various security operation centers affected, 3) new variants of older malware, 4) custom tools never encountered previously or reported by security vendors or a combination of modified open-source tools designed to conceal the attacker's activities.

The attacker is evaluated as proficient and advanced, displaying traits indicative of an APT group, particularly taking into account the previous characteristics observed from the evidence.

2) Impact: Initial triage by the responders estimated the damage to be merely the infected on-site servers and end-user computers, with key hosts being ransomed. However, lab analysis of network traffic revealed that infection has spread much further than initially presumed. The incident responders identified malicious domains, and even IoT devices and cloud-based systems were revealed as being compromised, as they communicated with these malicious domains.

3) Timeline: The attack began in June 2099, as certain compromised hosts were identified to have made successful callbacks to malicious domains from early June. Host logs have determined that the point of entry was via spear phishing and an unsecured, exposed public-facing development website.

From June to December, the attacker was identified as propagating and spreading through the network and exfiltrating critical information from compromised data servers. Furthermore, it has been discovered that compromised game companies infected their in-development games with malicious backdoors and distributed them via video game digital distribution services and online storefronts to other victims.

4) Evaluation of Ontologies Used for Scenario 1: Figure 11 and Figure 12 represent the scenarios in resource description framework format using UCO & CASE and using SCOPE, respectively (differences are highlighted by the red box as shown in Figure 12).

```
@prefix uco-core: <https://ontology.unifiedcyberontology.org/uco/core/1.1.0> .
@prefix uco-observable: <https://ontology.unifiedcyberontology.org/uco/observable/1.1.0>
.
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

kb:bundle-PunggolSmartDistrictRansomware2024 a uco-core:Bundle ;
    uco-core:description "Investigation report for the ransomware attack on Punggol
Smart District, claimed by APT Triple Dragon on 1st Jan 2100." ;
    uco-core:object kb:cybercrime-incident .

kb:cybercrime-incident a uco-core:Cybercrime ;
    uco-core:name "Punggol_SD_Ransomware_Attack_2100" ;
    uco-core:description "APT Triple Dragon demanded $250 Million USD in ransom. The
attack affected various organizations within the Punggol Smart District." ;
    uco-core:threatActor "APT Triple Dragon" ;
    uco-observable:evidenceCollected kb:evidence-digital-communication ;
    rdfs:comment "This investigation is ongoing, and further evidence and indicators of
compromise are being examined." .

kb:evidence-digital-communication a uco-observable:DigitalEvidence ;
    uco-core:description "Digital communications and transactions related to the ransom
demand." ;
    uco-observable:ObservableObject kb:observable-email-communication, kb:observable-
cryptocurrency-transaction .
```

Fig. 11. UCO Representation of Incident

Both ontologies are capable of representing Scenario 1. However, the SCOPE ontology provides further smart city infrastructure-specific descriptions (e.g., *crimeType*, *Adversary*) which allow an increased level of granularity as well as adding in more information fields that could add useful context for cybercrime investigations. A greater distinction between SCOPE and UCO/CASE will be demonstrated in the subsequent sections and scenarios.

```
@prefix scope-crime: <https://ontology.scopeontology.org/scope/crime/0.1.1> .
@prefix scope-evidence: <https://ontology.scopeontology.org/scope/evidence/0.1.1> .
@prefix scope-indicators: <https://ontology.scopeontology.org/scope/indicators/0.1.1> .
@prefix scope-infrastructure:
<https://ontology.scopeontology.org/scope/infrastructure/0.1.1> .
@prefix scope-commonAttackpatterns:
<https://ontology.scopeontology.org/scope/infrastructure/0.1.1>
@prefix scope-role: <https://ontology.scopeontology.org/scope/role/0.1.1> .
@prefix scope-threats: <https://ontology.scopeontology.org/scope/threats/0.1.1> .
@prefix scope-vocabulary: <https://ontology.scopeontology.org/scope/vocabulary/0.1.1> .
@prefix uco-core: <https://ontology.unifiedcyberontology.org/uco/core/1.1.0> .
@prefix uco-observable: <https://ontology.unifiedcyberontology.org/uco/observable/1.1.0>
.
@prefix uco-role: <https://ontology.unifiedcyberontology.org/uco/role/1.1.0> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

kb:bundle-PunggolSmartDistrictRansomware2024 a uco-core:Bundle ;
    uco-core:description "Investigation report for the ransomware attack on Punggol
Smart District, claimed by APT Triple Dragon on 1st Jan 2100." ;
    uco-core:object kb:cybercrime-incident .

kb:cybercrime-incident a scope-crime:Cybercrime ;
    scope-master:name "Punggol_SD_Ransomware_Attack_2100" ;
    scope-crime:crimeType "Ransomware Attack" ;
    scope-master:description "APT Triple Dragon demanded $250 Million USD in ransom. The
attack affected various organizations within the Punggol Smart District." ;
    scope-threats:Adversary "APT Triple Dragon" ;
    scope-evidence:evidence kb:evidence-digital-communication ;
    rdfs:comment "This investigation is ongoing, and further evidence and indicators of
compromise are being examined." .

kb:evidence-digital-communication a scope-evidence:DigitalEvidence ;
    scope-evidence:description "Digital communications and transactions related to the
ransom demand." ;
    uco-observable:ObservableObject kb:observable-email-communication, kb:observable-
cryptocurrency-transaction .
```

Representation of incident

Threat

Fig. 12. SCOPE Representation of Incident

C. Scenario 2 - TTPs Identified

Based on industry best practices for determining and classifying TTPs used by adversaries (including APTs), the TTPs employed by APT Triple Dragon are categorized and classified according to the relevant MITRE ATT&CK techniques [20] and technique identifiers are outlined in the following sub-sections. The information will then be represented via UCO/CASE and SCOPE.

1) Initial Access:

Exploit Public Facing Application (T1190): The attackers initially targeted a vulnerable, public-facing gaming development website, leveraging an exposed GitLab console. This vulnerability, identified as CVE-2022-2884, allowed them to execute code remotely, gaining an initial foothold in the network.

Spearphishing Attachment (T1566.002): Concurrently, a spearphishing campaign was launched, targeting specific individuals within the organization. Malicious documents sent via email served as the initial dropper for malware, deceiving users into enabling macros or executing the embedded code, thus bypassing traditional email filters and security measures.

2) Discovery:

File and Directory Discovery (T1083): After gaining initial access (with reference to Section IV-C1), the attackers conducted reconnaissance to identify valuable files, directories, and configuration settings as seen in host logs. This allowed the attackers to identify the layout of the network, identifying targets for lateral movement, and locating sensitive data for exfiltration.

System Information Discovery (T1082): Gathering information about the operating systems, software installations, and network configurations would be crucial for tailoring further attacks, exploiting specific vulnerabilities, and avoiding detection.

3) Execution:

Command and Scripting Interpreter (T1059): Leveraging on information gleaned from initial access and discovery (referencing Section IV-C1 and IV-C2), the attackers utilized command-line interfaces and scripting to execute their payloads. This included the deployment of ransomware and Remote Access Trojans (RATs), allowing them to maintain control over compromised systems and perform further malicious activities undetected.

4) Persistence and Privilege Escalation:

External Remote Services (T1133): By exploiting the exposed GitLab console (from Section IV-C1), attackers ensured persistent access to the compromised web server. This access facilitated the lateral movement across the network and the compromise of additional hosts.

Valid Accounts (T1078): Through password spraying attacks, the attackers gained access to valid user accounts, escalating their privileges within the network. This allowed them to access critical resources and data, further entrenching their presence.

5) Lateral Movement:

Remote Services: SSH, RDP (T1021): With valid accounts at their disposal, attackers used services like SSH (Secure Shell) and RDP (Remote Desktop Protocol) to move laterally across the network, accessing and compromising additional systems.

Lateral Tool Transfer (T1570): To facilitate their movement and maintain access within the network, the attackers transferred tools or malware from one compromised host to another. This technique allowed for the execution of specific payloads tailored to each target system.

6) Defense Evasion:

Obfuscated Files or Information (T1027): The malware deployed during the attack was meticulously designed to evade detection by endpoint security solutions. Through obfuscation and packing techniques, including the use of virtual machine detection evasion tactics, the attackers minimized the malware's footprint and avoided triggering security alerts.

7) Collection :

Automated Collection (T1119): Malicious scripts and custom malware that automatically collects specified types of documents and data from the compromised systems were identified to have harvested important financial information.

Input Capture: Keylogging (T1056): Deploying keyloggers enabled the capture of credentials, sensitive information, and other inputs directly from the users' keystrokes, further compromising personal and organizational data.

8) Command and Control:

Application Layer Protocol: Web Protocols (T1071): The attackers utilized HTTP, HTTPS, or other web protocols for command and control (C2) communications which blended the malicious traffic with legitimate web traffic, and made detection more challenging.

Dynamic Resolution (T1568 & T1071.004): The attackers employed techniques such as Domain Generation Algorithms (DGA) and fast flux, making their C2 infrastructure more resilient and challenging to disrupt. Furthermore, they used the Domain Name System (DNS) application layer protocol to avoid detection filtering by blending in with existing traffic with expired domains that they had bought and could already pass through the network firewall. Commands to the remote system and the results of those commands were identified to be embedded within the protocol traffic between the client and server.

9) Exfiltration:

Exfiltration Over C2 Channel (T1041): Leveraging the established C2 channels to exfiltrate stolen data silently ensures that attackers maintain a low profile while continuously siphoning off sensitive information.

Scheduled Transfer (T1029): The attackers had set up automated exfiltration processes that operate at scheduled times can help avoid detection by blending in with standard network traffic patterns, especially during peak hours.

10) Impact:

Data Encrypted for Impact (T1486): The deployment of ransomware to encrypt critical systems and data disrupts operations and serves as a direct method for financial gain through ransom demands from the APT.

Endpoint Denial of Service (T1499): The identified attacks on individual high-value endpoints, such as resource hijacking and request floods, rendered devices unusable, further complicating recovery efforts.

11) Evaluation of TTPs Representation (Scenario 2): Represented in Figure 13 (UCO representation of tactics, techniques, and procedures) and Figure 14 (SCOPE representation of tactics, techniques, and procedures) were the two different resource description framework representations of the various tactics, techniques, and procedures identified to be used by the threat actor in the scenario. Both have their use cases and help convey essential knowledge about what the attacker has performed to other investigators. In the UCO representation, we demonstrated extending it using the MITRE ATT&CK framework (knowledge of the ATT&CK framework would be a prerequisite for DFI involved in the investigation). Meanwhile, in SCOPE, we further imbued the ontology with MITRE Common Attack Pattern Enumeration and Classification (CAPEC) [21].

The different approaches were due to the use case of the SCOPE ontology. MITRE ATT&CK provides a structured and comprehensive framework for understanding TTPs that adversaries may use; it focuses on specific actions that may happen at each stage of an attack lifecycle and is more detailed and commonly preferred by security professionals such as threat intelligence analysts. However, with the addition of CAPEC, DFI can enumerate and classify common attack patterns used by adversaries. This allows for a higher level of understanding and documentation of attacks, enabling more effective use in threat modelling and informing of defensive strategies for decision makers.

Throughout the entire attack, the attacker performs a variety of actions. These actions are reflected and captured using MITRE ATT&CK as seen in Figure 13 using UCO, which demonstrates and highlights the cyber adversaries' tactics and techniques. SCOPE, when adopted on top of UCO/CASE, will also highlight the common attack patterns the defenders observe. SCOPE also highlights the CWE and CVE as observed in Figure 14, noting the weaknesses that the attacker used. This allows the remediation team to take effective action and patch the specific vulnerabilities exploited.

```
@prefix uco-core: <https://ontology.unifiedcyberontology.org/uco/core/1.1.0> .
@prefix uco-observable: <https://ontology.unifiedcyberontology.org/uco/observable/1.1.0> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

kb:exploit-public-facing-application
    a uco-action:Action ;
    uco-core:name "Exploit Public Facing Application (T1190)" ;
    uco-core:description "The attackers targeted a vulnerable, public-facing gaming development
website leveraging an exposed GitLab console. Vulnerability CVE-2022-2884 allowed remote code
execution, gaining an initial foothold." .

kb:spearphishing-attachment
    a uco-action:Action ;
    uco-core:name "Spearphishing Attachment (T1566.002)" ;
    uco-core:description "A spearphishing campaign was launched, targeting individuals within the
organization with malicious documents via email, serving as the initial dropper for malware." .
kb:file-directory-discovery
    a uco-action:Action ;
    uco-core:name "File and Directory Discovery (T1083)" ;
    uco-core:description "Attackers conducted reconnaissance post-initial access to identify
valuable files and directories for lateral movement and data exfiltration." .

kb:system-information-discovery
    a uco-action:Action ;
    uco-core:name "System Information Discovery (T1082)" ;
    uco-core:description "Information about operating systems, software installations, and network
configurations was gathered to tailor further attacks and exploit vulnerabilities." .

kb:command-scripting-interpreter
    a uco-action:Action ;
    uco-core:name "Command and Scripting Interpreter (T1059)" ;
    uco-core:description "Attackers utilized command-line interfaces and scripting to execute
payloads, including ransomware and RATs, to maintain control over compromised systems." .

kb:external-remote-services
    a uco-action:Action ;
    uco-core:name "External Remote Services (T1133)" ;
    uco-core:description "The exposed GitLab console was exploited, ensuring persistent access and
facilitating lateral movement and additional host compromises." .


kb:exfiltration-over-c2-channel
    a uco-action:Action ;
    uco-core:name "Exfiltration Over C2 Channel (T1041)" ;
    uco-core:description "Stolen data was exfiltrated silently over established C2 channels." .

kb:data-encrypted-for-impact
    a uco-action:Action ;
    uco-core:name "Data Encrypted for Impact (T1486)" ;
    uco-core:description "Ransomware encrypted critical systems and data, disrupting operations
and demanding ransom." .
```

UCO Representation
ATT&CK TTP
Technique Description

Fig. 13. UCO Representation of Tactics, Techniques, and Procedures

TABLE V
MALWARE IOCS FOR SCENARIO 3

Type of Malware	**MD5 Hash**	**Malicious Domains**
Encryptor Malware	00c4c3946ec03c915cfe4cbddff e93da f84d54b351b7926106ef377b06 423734 762a96d79e747457e086e68128 16b0aa	5jua3omslrbkks4c[.] onion[.]link agegamepay[.]com ageofwuxia[.]com ageofwuxia[.]info ageofwuxia[.]net ageofwuxia[.]org

D. Scenario 3 - Indicators of Compromise, Containment and Recovery

IoCs that may prove helpful for defenders to denylist or further investigate are expected to be discovered during the investigation. During this fictional scenario, SingCERT discovers that the APT enlists the use of Encryptor Malware, which makes callbacks to various malicious domains. Table V shows the associated IoCs obtained after malware reverse-engineering was performed on the retrieved samples.

Meanwhile, the data in Table V were represented using UCO in Figure 15 and via SCOPE in Figure 16.

Throughout the investigation process, the IoCs gathered would be helpful for attribution and denial lists used by defenders. Figure 15 displays what the observed indicators are, using UCO/CASE, which is helpful in clearly stating what their origins are and what type of indicators they may be. SCOPE expands on this information by adding the infrastructure (e.g. *Digital/Operational Technology Layer*) that the indicator may affect and noting what specific types of threat they are and which systems are affected (e.g. *TelecommunicationSystem*) as seen in Figure 16. These details add an additional layer of granularity, which would speed up the investigation process and ensure that critical pieces of information are not overlooked when shared with other analysts during the investigation process.

As the ransomware ravaged the network and encrypted essential data on the hosts, the SingCERT team immediately mitigated the damage. The infected hosts were swiftly

```
@prefix scope-crime: <https://ontology.scopeontology.org/scope/crime/0.1.1> .
@prefix scope-evidence: <https://ontology.scopeontology.org/scope/evidence/0.1.1> .
@prefix scope-indicators: <https://ontology.scopeontology.org/scope/indicators/0.1.1> .
@prefix scope-infrastructure:
<https://ontology.scopeontology.org/scope/infrastructure/0.1.1> .
@prefix scope-commonAttackpatterns:
<https://ontology.scopeontology.org/scope/infrastructure/0.1.1>
@prefix scope-role: <https://ontology.scopeontology.org/scope/role/0.1.1> .
@prefix scope-threats: <https://ontology.scopeontology.org/scope/threats/0.1.1> .
@prefix scope-vocabulary: <https://ontology.scopeontology.org/scope/vocabulary/0.1.1> .
@prefix uco-core: <https://ontology.unifiedcyberontology.org/uco/core/1.1.0> .
@prefix uco-observable: <https://ontology.unifiedcyberontology.org/uco/observable/1.1.0>
.
@prefix uco-role: <https://ontology.unifiedcyberontology.org/uco/role/1.1.0> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

kb:exploit-public-facing-application
    a uco-action:Action ;
    uco-core:name "Exploit Public Facing Application (T1190)" ;
    uco-core:description "The attackers targeted a vulnerable, public-facing gaming
development website leveraging an exposed GitLab console. Vulnerability CVE-2022-2884
allowed remote code execution, gaining an initial foothold." .
    scope:CommonAttackPattern "CAPEC-253";
    scope-core:name "Exploitation of Remote Services" ;
    scope-core:description "Attackers exploit vulnerabilities in remote services to gain
unauthorized access to systems or networks." ;
    scope-core:hasCWE "CWE-94" ;
    scope-core:hasCVE "CVE-2022-2884" .

kb:spearphishing-attachment
    a uco-action:Action ;
    uco-core:name "Spearphishing Attachment (T1566.002)" ;
    uco-core:description "A spearphishing campaign was launched, targeting individuals
within the organization with malicious documents via email, serving as the initial
dropper for malware." .
    scope:CommonAttackPattern "CAPEC-163";
    scope-core:name "Spearphishing Attachment" ;
    scope-core:description "Attackers send malicious attachments via spearphishing
emails to trick recipients into executing malware." ;
    scope-core:hasCWE "CWE-451" .

kb:file-directory-discovery
    a uco-action:Action ;
    uco-core:name "File and Directory Discovery (T1083)" ;
    uco-core:description "Attackers conducted reconnaissance post-initial access to
identify valuable files and directories for lateral movement and data exfiltration." .

kb:system-information-discovery
    a uco-action:Action ;
    uco-core:name "System Information Discovery (T1082)" ;
    uco-core:description "Information about operating systems, software installations,
and network configurations was gathered to tailor further attacks and exploit
vulnerabilities." .
```

SCOPE CAPEC Extension
Common Attack Pattern
Attack Description
CWE
CVE

Fig. 14. SCOPE Representation of Tactics, Techniques, and Procedures

```
@prefix uco-core: <https://ontology.unifiedcyberontology.org/uco/core/1.1.0> .
@prefix uco-observable: <https://ontology.unifiedcyberontology.org/uco/observable/1.1.0>
.
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

kb:encryptor-malware
    a uco-observable:Artifact ;
    uco-core:name "Encryptor Malware" ;
    uco-observable:hash kb:hash-em-1, kb:hash-em-2, kb:hash-em-3 ;
    uco-core:description "MD5 hashes of Encryptor Malware identified." .

kb:hash-em-1
    a uco-types:Hash ;
    uco-types:hashMethod "MD5" ;
    uco-types:hashValue "00c4c3946ec03c915cfe4cbddffe93da" .

kb:hash-em-2
    a uco-types:Hash ;
    uco-types:hashMethod "MD5" ;
    uco-types:hashValue "f84d54b351b7926106ef377b06423734" .

kb:hash-em-3
    a uco-types:Hash ;
    uco-types:hashMethod "MD5" ;
    uco-types:hashValue "762a96d79e747457e086e6812816b0aa" .

kb:encryptor-malware-onion-link
    a uco-observable:URL ;
    uco-core:name "Encryptor Malware Onion Link" ;
    uco-observable:value "5jua3omslrbkks4c[.]onion[.]link" .

kb:malicious-domains-encryptor
    a uco-observable:DomainName ;
    uco-core:name "Encryptor Malware Associated Domains" ;
    uco-observable:value "agegamepay[.]com", "ageofwuxia[.]com", "ageofwuxia[.]info",
"ageofwuxia[.]net", "ageofwuxia[.]org" .
```

UCO Evidence
Evidence Name
Items
Description

Fig. 15. UCO Representation of Evidence

```
@prefix scope-crime: <https://ontology.scopeontology.org/scope/crime/0.1.1> .
@prefix scope-evidence: <https://ontology.scopeontology.org/scope/evidence/0.1.1> .
@prefix scope-indicators: <https://ontology.scopeontology.org/scope/indicators/0.1.1> .
@prefix scope-infrastructure:
<https://ontology.scopeontology.org/scope/infrastructure/0.1.1> .
@prefix scope-commonAttackpatterns:
<https://ontology.scopeontology.org/scope/infrastructure/0.1.1>
@prefix scope-role: <https://ontology.scopeontology.org/scope/role/0.1.1> .
@prefix scope-threats: <https://ontology.scopeontology.org/scope/threats/0.1.1> .
@prefix scope-vocabulary: <https://ontology.scopeontology.org/scope/vocabulary/0.1.1> .
@prefix uco-core: <https://ontology.unifiedcyberontology.org/uco/core/1.1.0> .
@prefix uco-observable: <https://ontology.unifiedcyberontology.org/uco/observable/1.1.0>
.
@prefix uco-role: <https://ontology.unifiedcyberontology.org/uco/role/1.1.0> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

scope:encryptor-malware
    a scope-observable:Artifact ;
    scope-core:name "Encryptor Malware" ;
    scope-observable:hash scope:hash-em-1, scope:hash-em-2, scope:hash-em-3 ;
    scope-core:description "MD5 hashes of Encryptor Malware identified. This malware
targets systems responsible for communication within the Smart City Infrastructure." ;
    scope-commonattackpattern:attackpattern "Software exploitation";
    scope-evidence:type "software/malware";
    scope-infrastructure:layer "digital/operational technology"
    scope-threat:type "Ransomware";
    scope-core:targetSystem scope:TelecommunicationSystem .

scope:hash-em-1
    a scope-types:Hash ;
    scope-types:hashMethod "MD5" ;
    scope-types:hashValue "00c4c3946ec03c915cfe4cbddffe93da" .

scope:hash-em-2
    a scope-types:Hash ;
    scope-types:hashMethod "MD5" ;
    scope-types:hashValue "f84d54b351b7926106ef377b06423734" .

scope:hash-em-3
    a scope-types:Hash ;
    scope-types:hashMethod "MD5" ;
    scope-types:hashValue "762a96d79e747457e086e6812816b0aa" .

scope:encryptor-malware-onion-link
    a scope-observable:URL ;
    scope-core:name "Encryptor Malware Onion Link" ;
    scope-observable:value "5jua3omslrbkks4c[.]onion[.]link" .

scope:malicious-domains-encryptor
    a scope-observable:DomainName ;
    scope-core:name "Malicious Domains Associated with Encryptor Malware" ;
    scope-observable:value "agegamepay[.]com", "ageofwuxia[.]com", "ageofwuxia[.]info",
"ageofwuxia[.]net", "ageofwuxia[.]org" ;
    scope-core:targetSystem scope:TelecommunicationSystem .
```

SCOPE Evidence
Evidence Name
Items
Description
Evidence Type
Infrastructure
Threat Type
Target System

Fig. 16. SCOPE Representation of Evidence

quarantined, and the ransomware was removed to prevent further spread and damage. A patch was quickly developed and deployed with a custom script that detected and eradicated the ransomware to avoid reinfection.

To reduce further damage to the companies and organizations, restoration from backups was performed to restore business activity and ensure that affected victims could continue performing critical functions despite all that had happened.

While analyzing logs and evidence files from the event, the digital forensic team discovered a symmetric encryption key in transit around four days after the ransomware deployment. The key was found in a network packet capture recorded during the first day of the initial compromise when the attacker started deploying ransomware onto the hosts.

Due to fortunate circumstances, the infected hosts were all encrypted by the same strain of ransomware and thus used the same encryption key. The SingCERT team used the encryption key, decrypted all the encrypted files, and found that all the data was still intact. This allowed the response team to resolve the ransomware issue without paying a hefty sum to the attackers. The total time taken for recovery took only one day for patches to be deployed, and restoring backups to allow for business continuity, analysis, and encryption key discovery took another three days.

1) Evaluation of IoCs, Containment and Recovery Representation (Scenario 3): The swift recovery process and fortunate discovery of the ransomware encryption key facilitated the restoration of assets affected by the cybersecurity incident. Rapid deployment of patches allowed operations and businesses to resume activities swiftly. The recovery operations are represented by UCO in Figure 17, which highlights the actions taken and their descriptions. SCOPE is represented in Figure 18, which provides additional context for analysts, such as the location of the activity taken and the evidence type collected. The analysts are empowered to differentiate from SCI compared to other types of environments and systems they may encounter. (Note the red boxes detailing additional contextual information.)

With the advent of SCI, it has become essential to distinguish the various attack surfaces, from standard telecommunication systems to transportation systems or operational technology water and energy systems. With SCOPE, we can represent and share data from emerging technologies, especially complex systems such as smart city developments.

```
@prefix uco-core: <https://ontology.unifiedcyberontology.org/uco/core/1.1.0> .
@prefix uco-observable: <https://ontology.unifiedcyberontology.org/uco/observable/1.1.0>
.
@prefix case: <http://caseontology.org/core#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

kb:bundle-SGCirtRansomwareResponse a uco-core:Bundle ;
    uco-core:description "Investigation report for the ransomware attack by APT Triple
Dragon mitigated by the SingCERT team in 2100." ;
    uco-core:object kb:cybercrime-incident .

kb:cybercrime-incident a uco-core:Cybercrime ;
    uco-core:name "SingCERT_Ransomware_Response_2100" ;
    uco-core:description "The ransomware attack encrypted important data on various
hosts. The SingCERT team quarantined the infected hosts, removed the ransomware, and
deployed patches and custom scripts to prevent reinfection. Restoration from backups was
performed, and the symmetric encryption key was discovered in transit, allowing
decryption of the files." ;
    uco-core:threatActor "APT Triple Dragon" ;
    uco-observable:evidenceCollected kb:evidence-network-packet, kb:evidence-backup-logs
;
    rdfs:comment "The response team successfully resolved the ransomware issue without
paying the attackers, and operations resumed quickly." ;

kb:evidence-network-packet a uco-observable:DigitalEvidence ;
    uco-core:description "Network packet captured during the initial compromise
containing the symmetric encryption key." ;
    uco-observable:ObservableObject kb:observable-network-traffic .

kb:evidence-backup-logs a uco-observable:DigitalEvidence ;
    uco-core:description "Backup files used for restoration of the encrypted data." ;
    uco-observable:ObservableObject kb:observable-backup-files .

kb:action-quarantine a case:Action ;
    case:actionType "Quarantine" ;
    case:description "The infected hosts were swiftly quarantined to prevent further
spread and damage." .

kb:action-remove-ransomware a case:Action ;
    case:actionType "Removal" ;
    case:description "The ransomware was removed from the infected hosts." .

kb:action-deploy-patch a case:Action ;
    case:actionType "Deployment" ;
    case:description "A patch and custom script were deployed to detect and eradicate
the ransomware, preventing reinfection." .

kb:action-restore-backups a case:Action ;
    case:actionType "Restoration" ;
    case:description "Restoration from backups was performed to restore business
activity." .

kb:action-decrypt-files a case:Action ;
    case:actionType "Decryption" ;
    case:description "The encryption key was used to decrypt all the encrypted files,
ensuring data integrity." .
```

UCO Incident Details
Name
Description
Evidence
Evidence
Description
Evidence Type

Fig. 17. UCO Representation of Recovery

```
@prefix scope-crime: <https://ontology.scopeontology.org/scope/crime/0.1.1> .
@prefix scope-evidence: <https://ontology.scopeontology.org/scope/evidence/0.1.1> .
@prefix scope-indicators: <https://ontology.scopeontology.org/scope/indicators/0.1.1> .
@prefix scope-infrastructure: <https://ontology.scopeontology.org/scope/infrastructure/0.1.1> .
@prefix scope-commonAttackpatterns: <https://ontology.scopeontology.org/scope/infrastructure/0.1.1>
@prefix scope-role: <https://ontology.scopeontology.org/scope/role/0.1.1> .
@prefix scope-threats: <https://ontology.scopeontology.org/scope/threats/0.1.1> .
@prefix scope-vocabulary: <https://ontology.scopeontology.org/scope/vocabulary/0.1.1> .
@prefix uco-core: <https://ontology.unifiedcyberontology.org/uco/core/1.1.0> .
@prefix uco-observable: <https://ontology.unifiedcyberontology.org/uco/observable/1.1.0> .
@prefix uco-role: <https://ontology.unifiedcyberontology.org/uco/role/1.1.0> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

kb:bundle-SGCirtRansomwareResponse a uco-core:Bundle ;
    uco-core:description "Investigation report for the ransomware attack by APT Triple Dragon mitigated by the SingCERT
team in 2100." ;
    uco-core:object kb:cybercrime-incident .

kb:cybercrime-incident a uco-core:Cybercrime ;
    uco-core:name "SingCERT_Ransomware_Response_2100" ;
    uco-core:description "The ransomware attack encrypted important data on various hosts. The SGcirt team quarantined
the infected hosts, removed the ransomware, and deployed patches and custom scripts to prevent reinfection. Restoration
from backups was performed, and the symmetric encryption key was discovered in transit, allowing decryption of the
files." ;
    uco-core:threatActor "APT Triple Dragon" ;
    uco-observable:evidenceCollected kb:evidence-network-packet, kb:evidence-backup-logs ;
    rdfs:comment "The response team successfully resolved the ransomware issue without paying the attackers, and
operations resumed quickly." ;
    scope-infrastructure:layer "digital/operational technology" ;
    scope-threat:type "Ransomware" ;
    scope-core:targetSystem scope:TelecommunicationSystem .


kb:evidence-network-packet a uco-observable:DigitalEvidence ;
    uco-core:description "Network packet captured during the initial compromise containing the symmetric encryption
key." ;
    uco-observable:ObservableObject kb:observable-network-traffic ;
    scope-evidence:type "network/packet" .
    scope-infrastructure:layer "digital/network" .

kb:evidence-backup-logs a uco-observable:DigitalEvidence ;
    uco-core:description "Backup files used for restoration of the encrypted data." ;
    uco-observable:ObservableObject kb:observable-backup-files ;
    scope-evidence:type "log/backup" .
    scope-infrastructure:layer "digital/storage" .

kb:action-quarantine a case:Action ;
    case:actionType "Quarantine" ;
    case:description "The infected hosts were swiftly quarantined to prevent further spread and damage." ;
    scope-core:targetSystem scope:TelecommunicationSystem .

kb:action-remove-ransomware a case:Action ;
    case:actionType "Removal" ;
    case:description "The ransomware was removed from the infected hosts." ;
    scope-core:targetSystem scope:TelecommunicationSystem .
    scope-infrastructure:layer "digital/storage" .

kb:action-deploy-patch a case:Action ;
    case:actionType "Deployment" ;
    case:description "A patch and custom script were deployed to detect and eradicate the ransomware, preventing
reinfection." ;
    scope-core:targetSystem scope:TelecommunicationSystem .
    scope-infrastructure:layer "digital/network" .


kb:action-restore-backups a case:Action ;
    case:actionType "Restoration" ;
    case:description "Restoration from backups was performed to restore business activity." ;
    scope-core:targetSystem scope:TelecommunicationSystem .
    scope-infrastructure:layer "digital/storage" .

kb:action-decrypt-files a case:Action ;
    case:actionType "Decryption" ;
    case:description "The encryption key was used to decrypt all the encrypted files, ensuring data integrity." ;
    scope-core:targetSystem scope:TelecommunicationSystem .
    scope-infrastructure:layer "digital/storage" .
```

SCOPE Incident Details
Name
Description
Evidence
Location of Activity
Threat Type
Target System
Description
SCOPE Evidence type
SCOPE Affected Infrastructure Layer

Fig. 18. SCOPE Representation of Recovery

E. Scenario Evaluation Summary

Through this scenario (with reference to Section IV-B to Section IV-D), we have identified the key use cases that SCOPE would value add to cybercrime investigators, allowing users to add more granularity in details. An example is seen during the recovery phase of the scenario represented in Figure 18 as compared to Figure 17, which highlights the affected areas of infection and damage, allowing increased efficiency and rapid remediation. Complex technical details become easily accessible to other teams who may have a higher level requirement of the ontology, such as malware analysts who may require additional context during the investigation. They may appreciate the extra information such as the affected system and threat type as seen in Figure 16 compared to Figure 15, which during routine investigations may only capture the basic details of name, hash and filetype. SCOPE can support evidence and case data sharing through its interoperability with UCO/CASE.

SCOPE expands on SCI-focused cybercrime, providing specific terminology based on ISO backed definitions which ensures that newer innovations will be covered while retaining older technologies which may still be in use as not all cities in the world develop at the same pace.

V. Limitations

This section states the limitations of SCOPE and how we ensured reasonable contributions and innovation to the field.

A. Industry Adoption

Cybersecurity developments in academia are generally created independently of industry trends, with the industry preferring research that may generate revenue compared to purely academic research that may not yield results or are not profitable. Additionally, there may be parallel research being done with industry preferring their own standards over other proposed standardizations, such as using STIX (Structured Threat Information Expression) and TAXII (Trusted Automated Exchange of Intelligence Information) over UCO/CASE, which performed similarly. These frameworks and protocols were developed for cybersecurity information sharing and providing machine readability for automation. However, the industry has indicated a preference towards developments led by leading cybersecurity organizations and companies.

Despite the lack of adoption of ontologies in the cybersecurity industry, MITRE has also recently developed its own set of cybersecurity-themed ontologies [40] as more and more organizations have recognized the importance of accurately representing and sharing information. Another notable gap is that SCOPE aims to provide coverage over SCI, where the technology is still in developmental stages. Thus, there may be no notable adoption or usage in the industry any time soon.

B. Scoping

Due to the specific pain point we are trying to solve, SCOPE only covers cybercrime in SCI. It thus may lack particular definitions found in other scenarios, such as physical security and social engineering. While this allows us to drill down and ensure coverage of pertinent information, it also leaves the different specialties of cybersecurity unaddressed. Additionally, as SCI is a new and upcoming field, we acknowledge that further developments will occur and render previous efforts useless. Thus, we have tried our best to ensure that SCOPE is technology-agnostic, allowing it to continue providing coverage as technology develops. Moreover, we have provided SCOPE as an open framework for further development.

C. Threats to Validity

As an ontology, a possible gap that SCOPE faces may be content validity. SCOPE may not contain all relevant concepts and relationships since SCI is a continuously developing field with constant progress in cybersecurity. We can mitigate these issues by conducting extensive and in-depth literature reviews, engaging with experts in the domain to ensure proper coverage as well as continuous and iterative updating of SCOPE.

Additionally, to ensure that SCOPE would strike a delicate balance of being granular enough to represent critical points in SCI and broad enough to cover possible scenarios, we have turned to scenario-based testing to ensure that SCOPE would work in a real-life example.

Another caveat on external and construct validity factors would be that SCOPE was designed for smart cities in urban environments with advanced technological infrastructure. This may not be representative of rural cities or others that may have a different socioeconomic status. Currently, there are relatively few cities that may be termed smart cities or even cities with the infrastructure that supports developments to become smart cities.

D. Future Developments and Work

For the future development of SCOPE, some fields were not utilized that we desire to address in future iterations. Since this was the first version of SCOPE, fields such as *owl:priorVersion* and *owl:backwardCompatibleWith* were not used, and we have yet to develop a corresponding Python API like CASE. Nonetheless, our work is compatible with UCO/CASE projects, as we adhered to the corresponding ontology requirements.

Further future developments include aiming to extend Instance-Level Relationship Representation. While ontologies offer robust frameworks for representing structured knowledge in complex domains, they however inherently limit direct representation of relationships between individual instances, such as linking distinct cyber incidents. This gap can be addressed by integrating ontology rule languages like Semantic Web Rule Language (SWRL) or Datalog, which allow advanced reasoning to define and infer relationships at the instance level within the ontology.

We also intend to incorporate SWRL rules or Datalog-based reasoning within SCOPE to enhance its capability for capturing inter-instance relationships. Such an extension would enable us to model complex patterns, identify recurring behaviors, and establish incident linkages that are essential for forensic analysis. This would advance SCOPE's utility as a dynamic tool for investigating and analyzing cyber incidents within interconnected systems, including Smart City Infrastructure. Incorporating these rule languages is a significant step toward realizing a fully interconnected knowledge base that supports in-depth, context-aware incident analysis in rapidly evolving cyber-physical environments.

Finally, we plan to conduct a user study with digital forensic professionals from the public and private sectors with respect to the usage of SCOPE in their investigation processes related to SCI to determine how we can further improve SCOPE for industry usage. We will also explore creating open-source tools that focus on using SCOPE as a mechanism for investigation and data sharing.

VI. Related Work

Digital forensics investigation has seen significant advances in the methodologies, tools, and techniques used to extract features and correlate evidence. These tools and techniques have various degrees of adoption and perform their objectives to different levels of success. Additionally, to improve collaboration between organizations, efforts have been made to propose a variety of formats that can help standardize the information shared.

A. Feature Extraction and Correlation

One notable example of innovative developments in feature extraction would be Garfinkel's cross-drive analysis [41], where he introduces the technique of examining multiple hard drives simultaneously to capture patterns and identify connections that may not be visible in isolation. This approach is beneficial in extensive complex investigations that may span multiple systems. Flaglien et al.'s [42] cross-evidence correlation research on malware also provides a comprehensive framework for malware detection by leveraging cross-evidence correlation to identify malware traces across different hosts. Their proposition increases the speed and accuracy of malware identification and allows efficient processing of large datasets, which may bog down investigators. Another relevant research would be the FACE framework by Case et al. [43], which, in addition to automating the process of discovering and correlating digital evidence from multiple data sources, also provides a suite of tools and techniques to analyze information from a myriad of sources and standardizes the extracted data as eXtensible Markup Language (XML) and Attribute Relation File Format (ARFF).

Despite these advances in feature extraction and correlation in digital forensics, a few gaps still need to be addressed. One major limitation would be the need for standardized methods of ingesting data from more diverse sources, such as s Garfinkel's [41] cross-drive analysis, which has robust support but lacks the capability of ingesting network and application logs. Similarly, while efficient at identifying malware, Flaglien et al.'s [42] research does not support the holistic depth needed in digital forensics. Another issue would be the need for more real-time analysis in the FACE framework [43], paramount in a time-sensitive incident response scenario. These gaps signify the importance of digital forensic frameworks that can integrate well into the existing ecosystem of tools with industry-standard formats such as RDF and the significance of analyzing various sources in real time.

One recent development by Kougioumtzidou et al. [44] used a neural network-assisted framework to build and update cybersecurity taxonomies and ontologies in cyber threat intelligence. Kougioumtzidou et al. [44] constructed the proposed taxonomy by identifying relevant entities via Natural Language Processing (NLP) techniques and regular expressions from a custom dataset. The resulting entities were further distilled via Term Extraction and Relation Extraction, piping into a Python library to construct the ontology. Finally, the taxonomies and ontologies were updated through semantic searches. While it was a refreshing take on the aspect of feature extraction and relationship correlation, the proposed framework lacked further ontology validation work, such as ABox reasoning. Moreover, since the custom datasets did not contain information about SCI, the scope of the approach was constrained compared to our work.

SCOPE aims to address the gaps identified by developing a comprehensive ontology that could be abstracted and express common constructs across the different branches of cybersecurity and to provide a unified standard for representing collected evidence aligned with the cybersecurity industry's widely used formats.

B. Data Representation in Cybersecurity

Data representation is crucial in cybersecurity, where collaborative investigation is paramount among blue teams. Additionally, a standardized format helps to streamline the analysis of data between various tools. However, due to this lack of standardization, existing tools and forensic frameworks do not integrate well with each other.

Additionally, numerous new smart devices need to be standardized due to the burgeoning field of IoT and the inception of smart cities. Stoyanova et al. [45] highlight that the origins of IoT data are often unverified, and the data itself typically lacks metadata, making it challenging to identify potential evidence. Furthermore, these IoT nodes are often in continuous operation and generate large amounts of data; if the data is dirty and unusable, it is just noise. Garfinkel's work on digital forensics XML [46] was aimed at creating a standardized format to ensure interoperability among different forensics tools; however, with no digital forensics XML standard and the lack of fixed schema, there needed to be more adoption. Another significant contribution to the standardization of digital forensics is the Cyber-investigation Analysis Standard Expression (CASE) framework [6], developed by an international community of forensic practitioners and researchers. It provides a detailed and structured method of representing information from cyber incidents. While comprehensive, many organizations and companies may have proprietary environments incompatible with CASE [6]. The Digital Forensics framework by Baguelin et al. [47] was another noteworthy attempt at standardizing digital forensics, collecting multiple accolades and publications during its popularity. It was open source and supported on both Linux and Windows. However, user and developer interest shifted away, resulting in its gradual demise.

In the realm of cybersecurity countermeasures, Sánchez-García et al. [48] conducted a systematic mapping review of countermeasures and their taxonomies for risk treatment. Their work analyzes 26 taxonomies and catalogs of cybersecurity countermeasures, revealing trends, gaps, and the evolving focus of risk treatment strategies. Notable frameworks such as ISO27002 and NIST SP800-53 are highlighted for their widespread application. However, Sánchez-García et al. identified gaps in addressing risks associated with rapidly evolving domains, such as IoT and advanced persistent threats [48]. The study stresses the importance of holistic, standardized approaches integrating residual risk assessments while promoting interoperability across diverse systems. Despite its depth, the review broadly categorizes existing frameworks without proposing concrete mechanisms for their alignment or adaptation to emerging technologies.

SCOPE seeks to address these lapses by building on top of existing frameworks such as UCO/CASE ontology to ensure that it is interoperable and follows the standardized formats of ontology while staying technology agnostic to withstand the march of development.

C. Cybersecurity Ontologies

Ontologies are commonly used in academia, where formal naming and definitions are essential to establish the same baseline knowledge for research; its usefulness in the standardization of expertise has not been unnoticed by cybersecurity professionals. One of the most notable cybersecurity ontologies is UCO (Unified Cyber Ontology) [5], which provides a comprehensive list of definitions and support for various scenarios such as cybercrime. Another popular ontology would be the D3FEND Digital Artifact Ontology (DAO) from MITRE [40], developed as part of its STIX framework; it is adept at facilitating the sharing of threat intelligence and supports widespread integration as part of STIX. IEEE also has its standard core ontologies, with a Common Core Cyber Ontology (C30) developed by its cyber ontology working group [49]. IEEE's C30 is designed as an overarching ontology covering aspects of cyberspace, serving as a domain-level representation of the various activities of cybersecurity.

Kahvedžić and Kechadi [50] introduced the DIALOG framework, which offers a structured methodology for modeling, analyzing, and reusing digital forensic knowledge. By addressing the need for a comprehensive and reusable forensic knowledge repository, DIALOGUE reduces investigative process redundancy while enabling systematic, knowledge-driven investigations [50]. Its modular design and scalability are particularly effective for addressing contemporary challenges, such as integrating forensic workflows across diverse technological and organizational environments. While DIALOG provides a robust foundation, the applicability to emerging fields like IoT and distributed environments such as SCI remains underexplored, limiting its utility in addressing modern digital forensic complexities.

With various organizations developing their cybersecurity ontologies, a few gaps have not been addressed during their development. A key issue is the inability to address new and upcoming cybersecurity developments, such as UCO providing a general, well-rounded cybersecurity ontology that can represent many scenarios; however, they lack the newer terms and definitions due to IoT and SCI developments. Another critical gap is the over specialization of ontologies. The MITRE DAO is highly adept at threat intelligence sharing but does not have a comprehensive vocabulary that could represent the myriad of possible scenarios. Widespread adoption is also critical for an ontology to be successful. With a lack of implementation in cybersecurity tools, IEEE's C30 faces issues with industry usage since it has not been used with any of the typical cybersecurity applications, unlike UCO or DAO, which has seen usage in tools such as Autopsy [51].

SCOPE extends from UCO, ensuring that it is interoperable and can also be used with tools that already incorporate UCO. SCOPE also extends the terminology to cover emerging technologies, thus providing the representation of cutting-edge cybercrime and investigation scenarios. By integrating these advances, SCOPE presents a forward-looking solution that addresses the unique challenges posed by emerging cybercrime scenarios and technologies.

VII. Discussion and Conclusion

With the danger of cybercrime in the constantly developing field of technology, it is paramount that the LEA and DFI are suitably equipped to deal with this burgeoning issue. Additionally with the advent of smart cities with technology interwoven at every layer of the system, communicating and sharing key vital information is critical to defeating these adversaries.

We proposed SCOPE, a modular expansion profile of the Unified Cyber Ontology (UCO) and the Cyber-investigation Analysis Standard Expression (CASE) ontology. Noting that building a SCI ontology just by relying on UCO and CASE could be a laborious and semantically challenging endeavor, we proposed SCOPE as an expansion. SCOPE adheres to the original design requirements of UCO and CASE, and also achieves tool interoperability and the ability for collaborative investigations.

SCOPE is technology-agnostic while also adhering to international standards such as the ISO standards. Additionally, it contains enough granularity to allow users to pinpoint key information while also ensuring it is able to capture abstract definitions that can cover emerging technologies. For extensibility and further research in cybercrime and digital forensics, we have made the entire SCOPE ontology available in the following URL:

https://ontology.scopeontology.org

VIII. Acknowledgments

This research is partially supported by Infocomm Media Development Authority under its Future Communications Research & Development Program (Award number FCP-SUTD-RG-2022-017) and National Research Foundation, Singapore, under its National Satellite of Excellence Programme "Design Science and Technology for Secure Critical Infrastructure: Phase II" (Award No: NRF-NCR25-NSOE05-0001). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the respective funding agencies.

References

[1] Steve Morgan. Cybercrime to cost the world $10.5 trillion annually by 2025, 2020. (Accessed 5 January 2023).

[2] World Economic Forum. The Global Risks Report 2023: 18th Edition, 2023. (Accessed 13 January 2023).

[3] Yee Ching Tok, Chundong Wang, and Sudipta Chattopadhyay. Stitcher: Correlating digital forensic evidence on internet-of-things devices. *Forensic Science International: Digital Investigation*, 35:301071, 2020.

[4] Yee Ching Tok and Sudipta Chattopadhyay. Identifying threats, cybercrime and digital forensic opportunities in Smart City Infrastructure via threat modeling. *Forensic Science International: Digital Investigation*, 45:301540, 2023.

[5] Zareen Syed, Ankur Padia, Tim Finin, Lisa Mathews, and Anupam Joshi. UCO: A unified cybersecurity ontology. In *Workshops at the thirtieth AAAI conference on artificial intelligence*, pages 195 – 202, 2016.
[6] Eoghan Casey, Sean Barnum, Ryan Griffith, Jonathan Snyder, Harm van Beek, and Alex Nelson. Advancing coordinated cyber-investigations and tool interoperability using a community developed specification language. *Digital Investigation*, 22:14 – 45, 2017.
[7] ISO. Sustainable development in communities — Management system for sustainable development — Requirements with guidance for use. ISO 37101:2016, International Organization for Standardization, Geneva, Switzerland, 2016.
[8] ISO. Sustainable cities and communities — Indicators for city services and quality of life. ISO 37120:2018, International Organization for Standardization, Geneva, Switzerland, 2018.
[9] ISO. Internet of Things (IoT) - Reference Architecture. ISO 37122:2019, International Organization for Standardization, Geneva, Switzerland, 2019.
[10] ISO. Internet of Things (IoT) - Reference Architecture. ISO 37123:2019, International Organization for Standardization, Geneva, Switzerland, 2019.
[11] OpenText. *OpenText Encase Forensic*, 2024. (Accessed 5 January 2024).
[12] Exterro. FTK Forensics Toolkit - Digital Forensics Software Tools | Exterro, 2024. (5 January 2024).
[13] Michael Cohen, Simson Garfinkel, and Bradley Schatz. Extending the advanced forensic format to accommodate multiple data sources, logical evidence, arbitrary information and forensic workflow. *Digital Investigation*, 6:S57–S68, 2009. The Proceedings of the Ninth Annual DFRWS Conference.
[14] OASIS Cyber Threat Intelligence Technical Committee. Cyber threat intelligence technical committee, 2023. (Accessed 3 January 2023).
[15] Simson Garfinkel and Jon Stewart. Sharpening your tools. *Commun. ACM*, 66(8):44–52, jul 2023.
[16] CASE Ontology. Github - casework/case-implementation-ufed-xml, 2024. (Accessed 2 February 2024).
[17] CASE Ontology. Github - casework/case-implementation-axiom: Parser for magnet axiom, 2024. (Accessed 2 February 2024).
[18] CASE Ontology. Github - casework/case-implementation-xry: Implementation scripts for converting msab xry reports into case, 2024. (Accessed 2 February 2024).
[19] MITRE. Adversary Emulation Plans | MITRE ATT&CK, 2024. (Accessed 5 April 2024).
[20] MITRE. MITRE ATT&CK, 2024. (Accessed 5 April 2024).
[21] MITRE. CAPEC - Common Attack Pattern Enumeration and Classification, 2024. (Accessed 5 April 2024).
[22] Cyber-investigation Analysis Standard Expression Ontology. Downloads, 2024. (Accessed 30 September 2024).
[23] Unified Cyber Ontology. Ontospy > uco-1.3.0-docs, 2024. (Accessed 2 February 2024).
[24] Cyber-investigation Analysis Standard Expression Ontology. Ontospy > case-1.3.0-docs, 2024. (Accessed 2 February 2024).
[25] N. Noy and Deborah Mcguinness. Ontology Development 101: A Guide to Creating Your First Ontology. *Knowledge Systems Laboratory*, 32, 01 2001. (Accessed 30 September 2024).
[26] Stanford Center for Biomedical Informatics Research. protégé, 2024. (5 January 2024).
[27] W3C. OWL 2 Web Ontology Language Document Overview (Second Edition), 2012. (Accessed 5 April 2024).
[28] W3C. RDF Schema 1.1, 2014. (Accessed 5 April 2024).
[29] W3C. Owl web ontology language document overview, 2012. (Accessed 5 April 2024).
[30] W3C. SWRL: A Semantic Web Rule Language Combining OWL and RuleML, 2004. (Accessed 5 April 2024).
[31] W3C. RDF 1.1 Turtle, 2014. (Accessed 5 April 2024).
[32] Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.
[33] MITRE. APT41, Wicked Panda, Brass Typhoon, BARIUM, Group G0096 | MITRE ATT&CK, 2024. (Accessed 5 April 2024).
[34] JTC. About | Punggol Digital District, 2022. (Accessed 6 April 2024).
[35] Mandiant. APT41 (Double Dragon): A Dual Espionage and Cyber Crime Operation | Mandiant, 2022. (Accessed 5 April 2024).
[36] National Institute of Standards and Technology. Computer security incident handling guide. *NIST*, 2012.
[37] National Institute of Standards and Technology. Incident response recommendations and considerations for cybersecurity risk management. *NIST*, 2024.
[38] Elastic. Elastic Stack: (ELK) Elasticsearch, Kibana & Logstash | Elastic, 2024. (Accessed 5 April 2024).
[39] Splunk. Splunk Products | Splunk, 2024. (Accessed 5 April 2024).
[40] Peter E. Kaloroumakis and Michael J. Smith. D3fend: A knowledge graph of cyber defense techniques. https://d3fend.mitre.org/resources/D3FEND.pdf, 2021. Accessed: 2024-07-07.
[41] Simson L. Garfinkel. Forensic feature extraction and cross-drive analysis. *Digital Investigation*, 3:71–81, 2006. The Proceedings of the 6th Annual Digital Forensic Research Workshop (DFRWS '06).
[42] Anders Flaglien, Katrin Franke, and André Årnes. Identifying Malware Using Cross-Evidence Correlation. In *Advances in Digital Forensics VII*, volume 361, pages 169–182. Springer, 01 2011.
[43] Andrew Case, Andrew Cristina, Lodovico Marziale, Golden G. Richard, and Vassil Roussev. Face: Automated digital evidence discovery and correlation. *Digital Investigation*, 5:S65 – S75, 2008. The Proceedings of the Eighth Annual DFRWS Conference.
[44] Anna Kougioumtzidou, Angelos Papoutsis, Dimitrios Kavallieros, Thanassis Mavropoulos, Theodora Tsikrika, Stefanos Vrochidis, and Ioannis Kompatsiaris. An end-to-end framework for cybersecurity taxonomy and ontology generation and updating. In *2024 IEEE International Conference on Cyber Security and Resilience (CSR)*, pages 247–254, 2024.
[45] Maria Stoyanova, Yannis Nikoloudakis, Spyridon Panagiotakis, Evangelos Pallis, and Evangelos K. Markakis. A Survey on the Internet of Things (IoT) Forensics: Challenges, Approaches, and Open Issues. *IEEE Communications Surveys & Tutorials*, 22(2):1191–1221, 2020.
[46] Simson Garfinkel. Digital forensics XML and the DFXML toolset. *Digital Investigation*, 8(3):161–174, 2012.
[47] Frédéric Baguelin, Solal Jacob, Christophe Malinge, and Jérémy Mounier. Digital forensics framework (dff). https://github.com/arxsys/dff, 2010. Accessed: 2024-07-07.
[48] Isaac D. Sánchez-García, Tomás San Feliu Gilabert, and Jose A. Calvo-Manzano. Countermeasures and their taxonomies for risk treatment in cybersecurity: A systematic mapping review. *Journal of Systems and Software*, 195:111555, 2023.
[49] Alexander P. Cox, Brian Haugh, Casey Rock, David Kasmier, Donald Pellegrino, James Schoening, J. Neil Otte, Laurie Tyzenhaus, Mark Jensen, and Ronald Rudnicki. Cyber ontology releases. https://opensource.ieee.org/cyber-ontology-working-group/cyber-ontology-releases, 2022. Accessed: 2024-07-07.
[50] Damir Kahvedžić and Tahar Kechadi. Dialog: A framework for modeling, analysis and reuse of digital forensic knowledge. *Digital Investigation*, 6:S23–S33, sep 2009.
[51] Basis Technology. *Autopsy User Documentation: Reporting*, 2024. (Accessed 4 April 2024).