


Systematic Review

Knowledge Graphs and Semantic Web Tools in Cyber Threat Intelligence: A Systematic Literature Review

Charalampos Bratsas ^{1,2,†}, Efstathios Konstantinos Anastasiadis ^{1,2,*,†}, Alexandros K. Angelidis ^{1,2,*,†},
Lazaros Ioannidis ^{1,2}, Rigas Kotsakis ¹ and Stefanos Ougiaroglou ¹

¹ Department of Information and Electronic Engineering, International Hellenic University, 57400 Thessaloniki, Greece; cbratsas@iee.ihu.gr (C.B.); larjohn@gmail.com (L.I.); rkotsakis@gmail.com (R.K.); stoug@ihu.gr (S.O.)

² Open Knowledge Foundation Greece, 54352 Thessaloniki, Greece

* Correspondence: ek.anastasiad@gmail.com (E.K.A.); a.angelidis@ihu.gr (A.K.A.)

† These authors contributed equally to this work.

Abstract: The amount of data related to cyber threats and cyber attack incidents is rapidly increasing. The extracted information can provide security analysts with useful Cyber Threat Intelligence (CTI) to enhance their decision-making. However, because the data sources are heterogeneous, there is a lack of common representation of information, rendering the analysis of CTI complicated. With this work, we aim to review ongoing research on the use of semantic web tools such as ontologies and Knowledge Graphs (KGs) within the CTI domain. Ontologies and KGs can effectively represent information in a common and structured schema, enhancing interoperability among the Security Operation Centers (SOCs) and the stakeholders on the field of cybersecurity. When fused with Machine Learning (ML) and Deep Learning (DL) algorithms, the constructed ontologies and KGs can be augmented with new information and advanced inference capabilities, facilitating the discovery of previously unknown CTI. This systematic review highlights the advancements of this field over the past and ongoing decade and provides future research directions.

Keywords: semantic; ontologies; knowledge graph; cybersecurity; cyber threat intelligence; CTI; machine learning; deep learning; network



Citation: Bratsas, C.; Anastasiadis, E.K.; Angelidis, A.K.; Ioannidis, L.; Kotsakis, R.; Ougiaroglou, S. Knowledge Graphs and Semantic Web Tools in Cyber Threat Intelligence: A Systematic Literature Review. *J. Cybersecur. Priv.* **2024**, *4*, 518–545. <https://doi.org/10.3390/jcp4030025>

Academic Editor: Aniello Castiglione

Received: 28 May 2024

Revised: 25 July 2024

Accepted: 29 July 2024

Published: 1 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The number of cyber threats and cyber attacks is constantly increasing as the world faces socioeconomic changes governed by the surge of technological reliance, unexpected phenomena like the COVID-19 pandemic, and ongoing conflicts [1]. The financial impact of cyber-related incidents on the global economy is concerning. Cyber attacks are expected to cause damages topping USD 10.5 trillion by the end of 2024 [2]. To protect critical infrastructure and shield themselves from incoming future threats, organizations, governments, businesses, and individuals need to leverage available information regarding previous incidents and delve into offenders' motives, tactics, and patterns of behavior [3]. Knowledge of the vulnerabilities of both systems and devices is critical as well.

Cyber Threat Intelligence (CTI) is provided as a solution for assisting in decision-making while addressing security-related issues [4,5]. It can be briefly described as the process of collecting, processing, and sharing information that is valuable for analyzing and detecting potential cyber threats [6]. Within the CTI pipeline, unprocessed raw data stemming from numerous cybersecurity sources such as reports, manuals, and websites are used as input and then carefully manipulated to extract useful knowledge. By accessing available CTI, small- and medium-sized enterprises (SMEs) have the opportunity to enhance the protection of their infrastructure and reduce the cost of obtaining relevant data [3].

In order to effectively share, distribute, and use cyber threat-related information, a common knowledge representation format is needed in order to ensure that all newly

acquired unstructured data stemming from a wide range of heterogeneous sources are standardized according to predefined protocols [7]. MITRE's Structured Threat Information eXpression (STIX) [8] is an example of a representation used as a common language when sharing CTI; it contains definitions of indicators, events, targets, and tactics used during an attack [9,10].

Semantic web tools such as ontologies and Knowledge Graphs (KGs) appear to be particularly useful for establishing shared vocabularies and practices. In particular, ontologies and KGs are widely used within different and multidisciplinary scientific fields such as medicine [11–13], criminal justice [14,15], and sports science [16,17], among others [18–22]. Ontologies are also applicable in the cybersecurity domain, where preexisting representations such as STIX can be incorporated, enhanced, and extended with new entities and definitions.

Ontologies contain important concepts, formally defined as classes, of a specific domain along with their properties, all described within a formal and unambiguous context [23,24]. In computer science terms, an ontology resembles a carefully constructed tree of classes, relationships, and inheritance of restrictions and properties [23,25]. Within the cybersecurity landscape, ontological models allow analysts to explore the patterns and connections between different entities and calculate useful indicators and scores for assessing potential threats and vulnerabilities [26].

On the other hand, KGs usually leverage the structure imposed by the relevant ontology and represent actual data as a set of triples comprising pairs of entities and their relationships [27]. They are often stored as graph databases, and can be easily queried and visualized with various tools such as Neo4j [28,29].

In addition to establishing a common semantic framework, ontologies and KGs are capable of automating the process of registering new data and inferring new knowledge that may be difficult to be discovered manually. Especially when combined with Deep Learning (DL) methods for transforming unstructured data to meaningful knowledge, ontologies and KGs become an important asset in the CTI pipeline. The possibility of human error can be greatly mitigated, and analysis becomes more capable of dealing with the exponential generation of CTI data [30].

Contribution

The scope of this work is to provide a systematic review of the current state-of-the-art regarding the utilization of ontologies and KGs within the domain of CTI. Specifically, we focus on three aspects, namely, the creation, utilization, and enhancement of CTI-related ontologies and KGs. In the third topic, we explore how Machine Learning (ML) and DL can be used for extracting entities and relationships that result to the creation or extension of CTI ontologies and KGs.

The rest of this paper is structured as follows: Section 2 describes the procedure followed for collecting the relevant literature to be reviewed; Section 3 reviews the selected papers and categorizes them accordingly; and in Section 4 we offer a final discussion on the topic and provide suggestions for potential future research.

2. Literature Selection Strategy

Our review aligns with the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines for systematic reviews.

Relevant literature was collected from four distinct databases: Scopus, Web of Science, IEEE Xplore, and Science Direct. Preferences were shortlisted using the search terms (“knowledge graph” OR “semantic” OR “ontology” OR “ontologies”) AND “cyber threat” on each publication’s title, abstract, and keywords. The results were restricted within the time period of 2013 to 12 March 2024.

After applying the aforementioned preliminary criteria, 225, 21, 80, and 69 papers were obtained from the Scopus, Science Direct, IEEE Xplore, and Web of Science databases, respectively. Works not written in English were later removed, along with book chapters,

mini-reviews, early access papers, conference reviews, short surveys, notes, and editorials. Any gray literature, such as technical reports or theses, were also excluded. Thus, our collection consisted exclusively of peer-reviewed articles and conference papers submitted at reputable relevant conferences. After the removal of duplicates and the first stage of the screening process, 215 papers remained.

In the final stage of the screening process, we examined the 215 remaining papers and removed works that were not focused on either the construction or utilization of ontologies or KGs. ML-related works that resulted in the generation or expansion of ontologies and KGs were included in our final literature collection, while studies that applied ML and DL methods solely to entity and relationship extraction tasks without demonstrating the creation or extension of ontologies and KGs were dismissed. Five papers were removed due to accessibility issues. To ensure the reliability and consistency of our selection process, all accessible works that passed the initial stages of the screening process were independently reviewed by three researchers. Studies that received unanimous agreement (“yes”) from all three reviewers were immediately accepted. For studies that raised discordance, a second round of reviewing was conducted until consensus was reached. No formal statistical measures of inter-rater reliability were applied; instead, the final selection was made based on the collective agreement of all authors. After the screening process, we settled on a collection of 76 papers for review. The results of the literature collection procedure are presented in Figure 1, while Figure 2 shows the number of produced works per year satisfying our selection criteria.

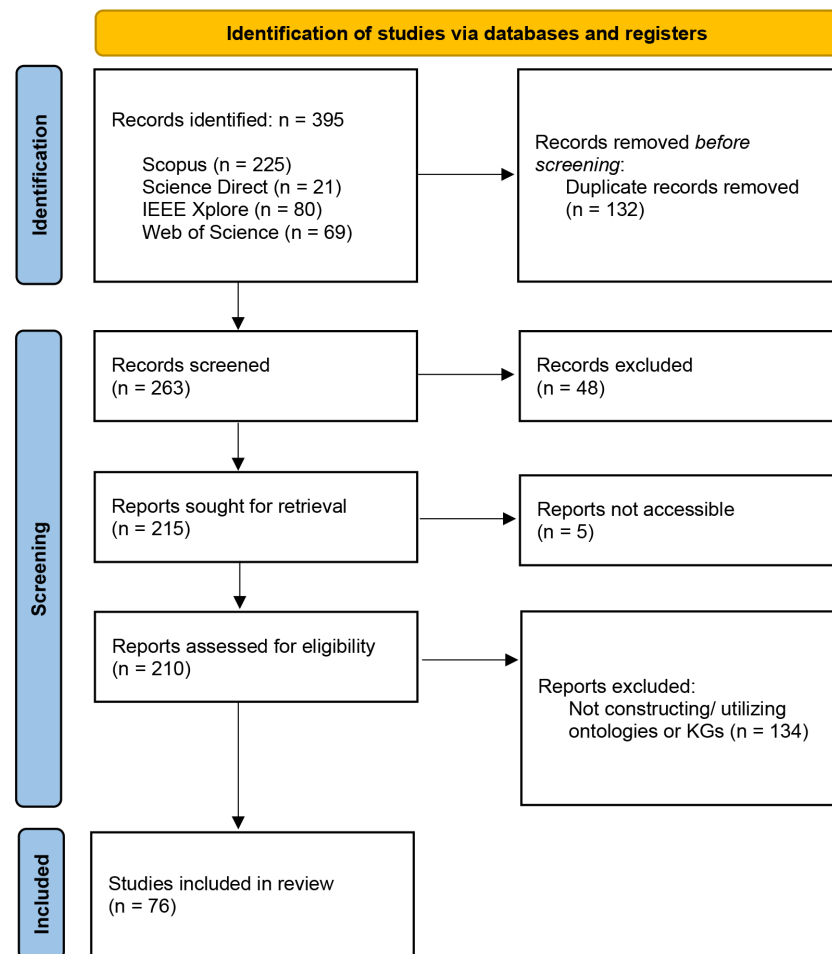


Figure 1. Number of records in each stage of the selection process.

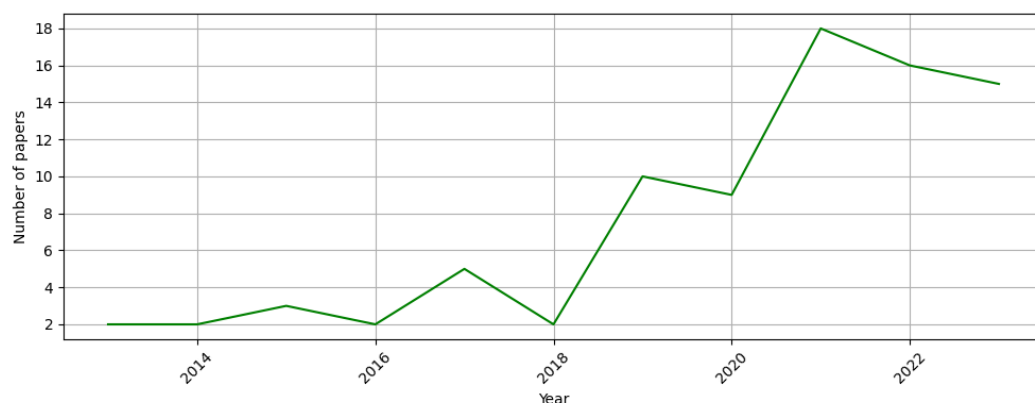


Figure 2. Amount of produced works related to CTI-related ontologies and KGs from 2013 to 2023. The significant number of conducted studies during the ongoing decade highlights the increasing popularity of semantic web technologies within the cybersecurity domain.

3. Results

The collected works that satisfied our selection criteria were divided into three main subcategories. The first subcategory (Section 3.1) contained all studies that explicitly focused on the development of a novel ontology or KG related to CTI, while the second (Section 3.2) encapsulated studies demonstrating the utilization of existing ontologies and KGs without describing the construction process. In the final subcategory (Section 3.3), we decided to separately gather all works that combined ML and DL methods with either construction or enhancement endeavors of ontologies and KGs. In each subcategory, the relevant works were reviewed in chronological order from oldest to latest. All studies reviewed in this chapter are also contained in aggregating tables in the corresponding subsection.

3.1. Ontology and Knowledge Graph Construction

Khairkar et al. [31] constructed an ontology focusing on the representation of data extracted from security logs of web applications to improve the classification of attacks and the identification of related events. The ontology supports the extraction of semantic relations between computer attacks and intrusions in an Intrusion Detection System (IDS).

Strasburg et al. [32] introduced the Semantic Model of Automated IDSs (S-MAIDS), emphasizing the creation of an ontology for automated tuning, correlation, and response selection in Intrusion Detection and Response Systems (IDRS). S-MAIDS aims to address the limitations of current IDRS models by providing a more elaborate model that automatically captures the semantics of events, detectors, and responses in a common concept. The proposed ontology is formalized as a Web Ontology Language (OWL) ontology to ensure clarity and automation of the unification of existing IDRS models.

A taxonomy model for the classification of CTI information exchange technologies was proposed by Burger et al. [33]. The authors focused on categorizing the ontologies themselves, and used the proposed taxonomy model in the analysis of different CTI exchange ontologies. By using a taxonomy schema, the authors were able to decompose the ontologies in order to evaluate the dependency and interoperability within the CTI landscape.

Casey et al. [34] addressed the importance of a standardized representation of digital forensic information to enhance querying and analysis of digital evidence. The authors leveraged the Cyber Observable eXpression CybOX [35] to propose a new standard for representing and exchanging digital forensic information, named Digital Forensic Analysis eXpression (DFAX). DFAX defines actions associated with digital traces to improve the efficiency of forensic analysis, further supporting the representation of action patterns for categorization of behaviors and goals during digital investigations.

Meckl et al. [36] developed a detection framework for Advanced Persistent Threats (APTs) by leveraging an APT ontology and reasoning tree patterns with ontology-based applicability conditions to systematically represent APT detection models. The ontol-

ogy language they used was an extension of Resource Description Framework (RDFs), with additional features to promote learning and evidence representation.

Falk [37] described the process of creating a CTI ontology based on the Lockheed Martin kill chain model. The component and properties as well as the ontology itself were all defined according to the OWL standard.

Mohsin and Anwar [38] introduced an ontology-based framework for Internet of Things (IoT) security analytics, aiming to combat APTs by aligning attack kill chain phases with network semantics. The ontology creation process involved the extension of existing CTI ontologies with new concepts and their alignment with a novel IoT ontology by leveraging Protégé software for ontology development and the Pellet engine for inference and consistency checking purposes. During the ontology engineering phase, the authors introduced new concepts and designed meta-ontologies for IoT, security controls, and Common Vulnerabilities and Exposures (CVE). In addition, concepts from existing ontologies were aligned and consistency tests were performed to ensure completeness.

STIX-Analyzer, a threat analysis framework based on OWL, was created by Qamar et al. [39]. It specializes in formal specification, semantic reasoning, and contextual analysis for the derivation of network-associated threats from large volumes of shared data. The proposed ontology contains the structure of STIX [8], CybOX [35], network configurations, and CVE, and is populated with extracted information from descriptions of emerging threats such as Tactics, Techniques, and Procedures (TTP), indicators, observables, exploit targets, and CVEs. The investigation of cyber threats is automated. The framework classifies threat relevance, determines threat likelihood, and identifies exposed assets through formulated rules and inferences by applying logic-based deductive inference rules defined in the Semantic Web Rule Language (SWRL). Protégé and OWL-Manchester syntax are employed along with the open-source Pellet reasoner. The proposed framework was evaluated by assessing the ontology's structure, clarity, consistency, accuracy, and feature novelty.

Mavroeidis et al. [40] presented an automated threat assessment system (CTI ontology) relying on the analysis of continuous incoming feeds of Sysmon logs in order to classify software in different threat levels (high, medium, low, or unknown threat). A lookup engine is included to reduce the load of SPARQL queries. The proposed system outputs RDF graphs in triples.

Diederichsen et al. [41] proposed an approach to analyzing real-time network log data within a Network Security Monitoring (NSM) environment using a graph database. They attempted to address the problem of integrating information from different sources to identify the relationship between various network traffic entities in a real-time logging NSM environment. Using relatively simple queries, the user can find related data from different log files and ultimately reconstruct malicious cyber activity to effectively find different potentially related attack paths.

An ontology for cybersecurity assessment was constructed by Doynikova et al. [42]. The authors focused on defining concepts and relationships between the primary attributes of relevant security data. The ontology includes classes such as "Source", "SecurityInformation", "Infrastructure", and "Metric", along with several types of defined relationships such as class inheritance hierarchy, object properties, and data properties. The authors also proposed an inference mechanism for calculating ontology-based security metrics to represent security metrics as separate instances within the ontology, allowing for computation of comprehensive metrics describing the security state of a system.

Kim et al. [43] proposed an ontology knowledge base to recommend security requirements focusing on APTs and system domain knowledge. The proposed knowledge base is divided into three parts: APT ontology, general security knowledge ontology, and domain-specific knowledge ontology. The integration of the aforementioned parts renders the extraction of appropriate security requirements feasible through a recommendation process that considers both real attack cases and system specifications. The authors aimed to enhance the understanding of complex attacks such as APTs and improve the security response in real-world scenarios.

The works related to the construction of ontologies or KGs during the period 2013–2019 are summarized in Table 1.

Table 1. Works relevant to ontology or KG construction from 2013 to 2019. Records are arranged in ascending order based on their publication year, and alphabetically by author name when the publication years of two or more papers coincide.

Reference	Year Published	Brief Summary	Domain
[31]	2013	Proposed the use of Semantic Web and Ontology concepts to define an approach for analyzing security logs for security issues identification.	Web Attacks
[32]	2013	Introduced S-MAIDS, a semantic approach to IDRS modeling using OWL ontologies.	IDRS
[33]	2014	Proposed a taxonomy for the classification of existing CTI sharing technologies.	CTI
[34]	2015	Proposed the DFAX ontology for representing and exchanging digital information.	Digital forensic information
[36]	2015	Presented an ontology for the enhancement of cyber defense against APTs.	APT
[37]	2016	Proposed an ontology for the efficient organization of OSINT and threat indicators.	CTI
[38]	2017	Proposed an ontology-based framework for the IoT environment to safeguard against APTs.	APT in IoT Environments
[39]	2017	Defined an ontology for threat analytics.	CTI
[40]	2018	Presented a CTI ontology for the analysis and classification of Sysmon logs.	CTI
[41]	2019	Proposed an approach of analyzing real-time network log data within a NSM environment by using a graph database.	NSM
[42]	2019	Proposed an ontology of metrics for security management.	Cybersecurity Management
[43]	2019	Proposed an ontology knowledge base that can define APT attack patterns and recommend security requirements.	APT

Liu et al. [44] proposed a KG ontology modeling method for network security based on STIX [8]. They analyzed the overlap between STIX and network security knowledge and generated an ontology schema with minimal redundancy and a strong structural hierarchy, aiming to further understand the structure and the relationships of attack activities in order to encourage informed decision-making when addressing security incidents.

Piplai et al. [45] fused different representations of malware threat intelligence to extend cybersecurity-related knowledge. Malware samples were retrieved, executed, and ultimately integrated with OSINT to construct a cybersecurity KG that would allow professionals to leverage its reasoning capabilities for malware pattern tracking.

MALOnt, an open-source malware ontology leading to the construction of a threat intelligence KG from a corpus of annotated malware reports, was introduced by Rastogi et al. [23]. The ontology contains concepts referring to malware characteristics along with attack and attacker details. New knowledge is inferred through deduction and induction using an OWL-based reasoner. Analysis, detection, classification, and attribution of malware-related threats is made feasible using MALOnt. MALOnt was evaluated through SPARQL queries. Queries that extracted instances as a response indicated that the questions adhered to the aim of the ontology.

Sills et al. [46] focused on generating a CTI repository referring to IoT medical devices and their known vulnerabilities from heterogeneous online data sources. The collected information was used to enrich a cybersecurity KG. Using relevant knowledge from Wikidata [47] and the Food and Drug Administration's AccessGUDID database [48], the gener-

ated graph embeddings were further refined to allow the augmented KG to create node graph representations of higher quality. The authors observed a 31% increase in the Mean Average Precision (MAP) when implementing an information retrieval (IR) task on the CTI and Wikidata KG augmented with AccessGudid compared to performing the same task on the CTI-only KG and the CTI KG augmented only with Wikidata.

Dora and Nemoga [25] proposed an ontology structure for enhancing the defensive capabilities of web apps when facing Cross-Site-Scripting (XSS) attacks. The authors discussed the importance and the advantages of leveraging ontologies and concluded that the exclusive use of ontologies as a defensive mechanism was not adequate to improve the security of web apps.

An ontology-based framework for enhancing the resilience and security of Information Technology (IT) systems in maritime port ecosystems was developed by Hutschenreuter et al. [49]. The framework comprises three subcomponents that leverage ontologies and inference methods for the recovery of cyber-related incidents. Additional advanced security tools handle the detection of suspicious activity. The authors intended to validate their proposed framework empirically in future studies.

Mavroeidis et al. [50] constructed an ontology for threat actor type characterization and threat actor type inference based on the Threat Agent Library (TAL). They utilized controlled vocabularies to characterize threat actors and their operations and enhanced the CTI with higher contextual information. With the proposed structure, they aimed to reduce human bias in classification and decision-making situations by encoding domain knowledge within the ontology. The efficiency of the constructed ontology was demonstrated through a use case analysis in which they attempted to automatically infer actor types.

Merah and Kenaza [10] proposed an ontology-based CTI analysis leveraging the catalogued concepts of STIX [8] to provide important threat information based on reported security alerts. The effectiveness of the approach was tested in a cyber threat monitoring scenario. In another relevant study [51], the same authors developed an ontology structure for cyber risk monitoring by integrating CTI into cybersecurity solutions and aligning STIX-provided concepts. An existing ontology designed for Security Information Event Management (SIEM) was also extended. The aim was to provide a common understanding of objects, concepts, and relationships in the cybersecurity domain through descriptive logic. The proposed framework was tested on a use case involving cyber risk monitoring.

The works reviewed above are related to the construction of ontologies or KGs in 2020 and 2021 and are summarized in Table 2.

Table 2. Works relevant to ontology or KG construction from 2020 to 2021. Records are arranged in ascending order based on their publication year and alphabetically by author name when the publication years of two or more papers coincide.

Reference	Year Published	Brief Summary	Domain
[44]	2020	Proposed a STIX-based network security KG ontology modeling method for the analysis of the concepts of network security knowledge.	Network Security
[45]	2020	Proposed an enriched cybersecurity KG by merging information about malware behavior data.	CTI
[23]	2020	Proposed MALOnt, an ontology for malware threat intelligence.	Malware Threat Intelligence
[46]	2020	Constructed a KG that stores CTI regarding various medical devices.	Vulnerabilities of Medical Devices in IoT
[25]	2021	Proposed an ontology for XSS attacks.	XSS Attacks
[49]	2021	Proposed an ontology-based framework with logical inference for the detection of cyber attacks.	Cyber Attacks and Incidents

Table 2. Cont.

Reference	Year Published	Brief Summary	Domain
[50]	2021	Presented an ontological approach for the automatic inference of threat actor types based on a standardized set of attributes.	Threat Actor
[10]	2021	Proposed a STIX-based ontological reasoning approach for potential cyber threats	CTI
[51]	2021	Constructed an ontology for cyber risk monitoring	Cyber Risk Monitoring

Ammi et al. [52] designed a novel semantic-based methodology for cloud-native cyber incident response. They leveraged ontologies and KGs to facilitate the retrieval and analysis of data referring to security challenges within the cloud environment. The model was evaluated by running sample cypher queries on the Neo4j graph database based on the authors' developed graph model from various data sources and Amazon services.

Bromander et al. [53] presented the results from a questionnaire investigating the use of standards and sharing practices of CTI. They proposed the Semi-Automated CTI (ACT) data model based on objects and relationships to address the challenges of automation. The model uses a back-end based on Apache Cassandra and Elasticsearch. An Apache TinkerPop graph engine enables graph querying with Gremlin. Object types from STIX [8], the Detection Maturity Model, the Diamond Model, and OSINT were used to populate the ontology. The constructed model was evaluated by assessing its usefulness, expressiveness, consistency, semantic agreement, and reasoning capability.

Collen and Nijdam [54] focused on the dynamic and automated identification of ongoing attacks and evaluation of the associated risks. An IoT stack ontology was built based on an existing taxonomy of the smart home domain to encapsulate vulnerabilities, attack attributions, impact evaluations, and mitigation strategies within a smart home environment. The authors also proposed a Dynamic Risk Assessment Framework (DRAF) that leveraged the built ontology and comprised various risk modeling steps. The framework was validated in smart home testbeds spanning multiple European countries, allowing for dynamic adaptation of the risk weights used by the receptors of the model through decision-making feedback. The results revealed strong potential for automating the decision-making process while ensuring a balance between security and privacy concerns.

Grigoriadis et al. [26] proposed the construction of a cybersecurity ontology for storing valuable information for the assessment of system security risks. The core elements of the ontology represented devices, networks, and human actors. The entities were filled using unstructured security-related data (reports, policies, and organization-specific cyber-information). The second layer of the ontology consisted of modeling information related to vulnerabilities, threats, exploits, and threat agents stemming from public resources. The authors used the ontology they constructed for Common Vulnerability Scoring System (CVSS) score prediction by applying logistic regression, and eventually discovered knowledge gaps that were later included in the National Vulnerability Database (NVD) [55].

A MITRE ATT&CK Enterprise Matrix-based ontology using OWL standards was built by Huang et al. [56]. CTI reports were parsed for extraction of essential knowledge, which was subsequently inserted into the ontology. The information extraction process was automated. Evaluation showed that the proposed rule-based information extraction method surpassed the performance of Neural Network (NN)-based methods in terms of precision.

Li et al. [57] proposed AttackKG for the automated extraction of attack behavior graphs from CTI reports. The performance of AttackKG was evaluated against 7373 procedures from 179 techniques stemming from MITRE ATT&CK and 1515 CTI reports collected from different CTI sources. AttackKG outperformed other existing frameworks such as EXTRAC-TOR [58] and TTPDrill [59] by identifying 28.272 attack techniques with 8393 Indicators of Compromise (IoC), achieving F1-scores equal to 0.887, 0.896, and 0.789 on attack-relevant entities, dependencies, and technique identification, respectively.

Rastogi et al. [60] proposed TINKER, a novel framework for capturing cyber threat information through a semi-supervised approach. TINKER transforms multimodal CTI into a structured format while preserving the context of the information. It leverages ontologies and information extraction models to capture CTI and integrates different data sources by structuring them using CTI KGs. The use of CTI-KG was demonstrated in two use cases involving malware family prediction and attack target inference.

Sharma and Kumar [61] advocated for the use of graph databases to monitor network logs in near-real-time by implementing a graph database easily queryable with tools such as Neo4j.

Yang et al. [28] focused on the construction of a KG based on kernel audit logs, aiming to efficiently organize large amounts of data through the use of graph databases such as Neo4j. The KG is used for enabling semantic queries in threat hunting activities, which involves two stages. In the first stage, the hunter formulates hypotheses from different semantic levels based on previous IoC and identified patterns. These hypotheses are later evaluated through graph query and visualization provided by Neo4j in order to assist the hunter in the elimination of false positives. The performance of the threat hunter KG was evaluated by assessing the time cost, required space occupation, and query delay.

The works related to the construction of ontologies or KGs in 2022 are summarized in Table 3.

Table 3. Works relevant to ontology or KG construction published in 2022. Records are arranged in an ascending alphabetical order.

Reference	Year Published	Brief Summary	Domain
[52]	2022	Proposed a cloud-native ontology capable of connecting security-related data from different cloud sources to enhance CTI.	Cloud-native cyber incident response
[53]	2022	Proposed a CTI ontology that enables the automation and analysis of the available threat intelligence.	CTI
[54]	2022	Developed a generic ontology for the representation of IoT objects to encapsulate vulnerabilities, attack attribution, impact evaluation and mitigation strategies within a smart home environment.	Risk Assessment in IoT Environments
[26]	2022	Proposed an ontology for risk assessment modeling.	Security Risk Information
[56]	2022	Proposed an ontology and an automated information extraction method, capable of integrating the parsed information from CTI reports into each instance.	CTI
[57]	2022	Proposed AttackKG, a KG used for the automated extraction of attack behavior graphs from CTI reports.	CTI
[60]	2022	Proposed TINKER, a framework that utilizes ontologies and KGs for capturing cyber threat information.	CTI
[61]	2022	Supported the use of graph databases for near real time network log monitoring.	Network Log Files
[28]	2022	Presented a KG based on kernel audit logs for the efficient organization of big data, enabling querying during threat hunting activities.	Cyber Threat Hunting

A unified ontology knowledge base for cybersecurity was constructed by Akbar et al. [62]. The ontology helps organizations to identify potential threats and design effective defensive methods. SPARQL queries allow for the extraction of relevant concepts for meaningful inference results, potentially leading to rapid responses when addressing a cybersecurity incident.

Compastié et al. [63] presented a security incident remediation strategy targeting cloud, edge, and on-premises environments for protection of SMEs. The authors introduced an ontology structure for the description of properties and relationships, which is necessary for selecting appropriate security measures. A mapping functionality complements the

ontology, aiming to identify the security mechanisms required in each case by leveraging relevant cyber threat knowledge.

To address the issue of blind SQL injection attacks against web application databases, Dora et al. [64] constructed OBSQL, an ontology for blind SQL weaknesses detection capable of providing relevant threat mitigation techniques. In their work, the authors additionally discussed the impact of blind SQL injections, different attack scenarios, detection and exploitation of vulnerabilities, and the usefulness of ontologies within the cybersecurity field.

Sánchez-Zas et al. [65] presented an ontology designed for real time risk management within a secured environment by organizing data on anomalies and cyber threats and applying established methodologies [66]. The ontology creation process required a broad understanding of the applied environment and technologies along with the extension of threat sources with anomalies captured by physical and logical sensors. Standardized methodologies such as the OQuRE framework [67] were used to evaluate the quality of the ontology to ensure its conformance to both functional and non-functional requirements. The OWL-based ontology integrates CTI and risk-management information and relates concepts through SWRL rules to improve system threat and risk assessment.

Wang et al. [24] proposed an event-based threat intelligence ontology for threat detection and response scenarios. The ontology uses the semantics of events to reorganize the elements of threat intelligence and simplifies the expression hierarchy, further improving the degree of structure. Additionally, the ontology combines the skeleton method with Formal Concept Analysis to ensure semi-automated construction and greater efficiency and formalization to assist in automated correlation analysis. The ontology was evaluated and validated through an example of specific instance data.

Zhang et al. [68] introduced the Attack and Defense Analysis of the Cybersecurity Ontology (ADACO) model, constructed by integrating data originating from multiple cybersecurity databases. Threat scenarios within the Threat Evolution Prediction Algorithm (TEPA) are represented as KGs, combining structural and textual features of entities for enhanced performance. TEPA enables the automatic detection of threats at device nodes and the correlation and mapping of multi-source information. The model also allows for the dynamic inference of threat evolution process, rendering the algorithm capable of making threat predictions.

Table 4 summarizes the works related to ontology and KGs construction that were published in 2023.

Table 4. Works relevant to ontology or KG construction published in 2023. Records are arranged in an ascending alphabetical order.

Reference	Year Published	Brief Summary	Domain
[62]	2023	Constructed a Unified Ontology encompassing APT techniques, weaknesses, vulnerabilities, and defense countermeasures.	APT
[63]	2023	Presented a security ontology for cloud, edge and on-premises environments and used it to determine the most appropriate security mechanisms.	CTI
[64]	2023	Proposed OBSQL, an ontology for detecting blind SQL weaknesses.	Blind SQL Injection
[65]	2023	Defined an ontology for the description of different types of anomalies and proposed an approach that merges the ontology with previously developed models for CTI.	Risk Management
[24]	2023	Proposed an event-based threat intelligence ontology for threat detection and response scenarios.	CTI
[68]	2023	Introduced the ADACO model, constructed by integrating data from multiple cybersecurity databases.	CTI

3.2. Utilization of Ontologies and Knowledge Graphs

Takahashi and Kadobayashi [69] proposed a method for generating RDF-based metadata to effectively manage cybersecurity-related information. The metadata structure adheres to the information types defined in a previously constructed cybersecurity information ontology [70]. The proposed RDF mechanism also manages to update obsolete information to mitigate potential security threats. The method was evaluated based on its extensibility, scalability, and credibility of information.

Lu and Kokar [71] developed a reasoning mechanism for cybersecurity queries when examining cyber situations. After proposing a formal definition of the term “cyber situation”, the authors used Barwise’s situation theory to extend the STIX ontology [8] for situation recognition. The proposed reasoning mechanism can recognize specific situation types and identify the smallest amount of information needed for answering a query automatically. Two queries were tested in a small part of the Skaion dataset, with the reported precision and recall scores exceeding 80% in both cases.

In another study, Asgarli and Burger [9] explored different CTI exchange standards such as STIX [8], Incident Object Description Exchange Format (IODEF) [72], and OpenIOC to evaluate the overlap between each other and other features using ontologies and techniques from library science. A comparison among the semantics of each format was conducted to highlight the potential advantages of RDF/OWL-based semantic exchange formats. The authors constructed an RDW/OWL ontology based on elements of STIX and IODEF, aiming to leverage the semantic power of RDF/OWL for the automation of cyber threat response systems.

Monteiro et al. [73] evaluated the security risks of networked systems by proposing a novel metric named “net vulnerability”. The authors leveraged an ontology they built in a prior work [74] and cataloged the vulnerabilities (CVE IDs and CVSS base vector) of the host of the network after gaining knowledge about the network’s inventory and mapping the relationships between the assets of the network. Two indices were computed for each vulnerability: the impact sub-score and the exploitability sub-score. Net vulnerability was evaluated on a fictitious network.

To enhance the knowledge regarding security issues within the IoT environment, Mozzaguardo and Jardim-Goncalves [75] utilized an ontology to gain relevant knowledge about the environment. To achieve this, the authors enabled the generation of automated alerts from sources such as IDS and firewalls or advanced correlations between different access attempts and network probes with location and time data. A case study approach for security requirement identification was also demonstrated.

Sikos [76] used a description logic formalism based on *SRIOQ* [77] to model fuzzy cyber-related knowledge and information, aiming to potentially enhance the automation of querying and reasoning when discovering CTI. By combining *P-SRIOQ* [78], *π -SRIOQ* [79], and *L-SRIOQ* [80], the author was able to infer probabilistic, possibilistic, and fuzzy cyber-related knowledge. The proposed method was evaluated on a case study involving the representation of vulnerabilities, attack patterns, rankings, and malware behavior, with certain, uncertain, crisp, and fuzzy axioms expressed as percentages of the captured semantics with respect to all the semantics of the model. The method was compared to alternative representations, and was able to capture 65.54% more knowledge than *ALC* logic or standard *SRIOQ*.

Avid and Wecl [81] used the STIX standard and the Maltego tool [82] to apply semantic methods in various CTI use cases. They emphasized the significant advantage of representing CTI within RDF repositories when dealing with previously undiscovered threats. Unlike relational databases, RDF repositories provide flexibility, as they do not require schematic alterations. The authors concluded that CTI knowledge could be sufficiently captured by STIX, Maltego, and threat hunting and logs, demonstrating that semantic technology was more advantageous for CTI than legacy relation databases.

An SIEM-based KG for relationship modeling among observed entities in proxy and DNS logs was introduced by Najafi et al. in [83]. In addition, the authors proposed a

graph algorithm for the estimation of maliciousness score called MalRank based on the associations a node had with other entities of the KG. The performance of MalRank was evaluated on a SIEM-based KG constructed from global enterprises' SIEMs data. MalRank scored 96% in AUC, outperforming Belief Propagation [84] in malicious Internet Protocols (IPs) and domain names detection.

Riesco and Villagra [66] developed a model that integrated risk intelligence along with CTI domains to demonstrate the advantages that ontologies offer in formal representation of concepts and relationships. The model consisted of different features, including data, business logic, services/applications, and visualization, and allowed for dynamic assessment of the security risks based on almost-real-time threat data to potentially refine its risk management capabilities.

A Cloud Forensic Readiness as a Service (CFRaaS) ontology-driven model aiming to gather digital evidence was proposed by Kebande et al. [85]. The ontological approach enables the classification and graphical visualization of different types of evidence in cloud-deployable models, resulting in a simplification of the semantics between different types of evidence.

Jung et al. [86] contributed with a security requirements recommendation tool for ATP attacks, using the Case-Based Problem Domain Ontology designed explicitly for APT attacks. The proposed tool analyzes an attack scenario and infers adjacent attacks according to similarity scores.

Shaked and Margalit [87] proposed OnToRisk, an automated cyber risk identification method integrated with information from different sources. OnToRisk relies on formal ontology concepts and relationships for analyzing organizational situations and providing suggestions for risk management. The proposed method was applied to a case study involving risk identification from software vulnerabilities, demonstrating its usefulness.

A domain-agnostic KG as a Service (KGaaS) framework for generating and maintaining domain-specific KGs for intelligent agent apps was proposed by Calyam et al. [88]. The authors presented a reference architecture comprising graph infrastructure tools and User Interfaces (UI) for generating large KGs within the domains of healthcare, power grids, and manufacturing. Data importers and semantic scripts are included within a customized knowledge curation pipeline to ensure rapid querying methods for cyber attack detection and defense mechanisms.

The works utilizing CTI ontologies or KGs in various scenarios are summarized in Table 5.

Table 5. Works strictly demonstrating the utilization of CTI ontologies or KGs for different purposes. Records are arranged in ascending order based on their publication year and alphabetically by author name when the publication years of two or more papers coincide.

Reference	Year Published	Brief Summary	Tools
[69]	2014	Proposed an RDF metadata generation mechanism for cybersecurity information management.	RDF metadata generation
[71]	2015	Developed a reasoning mechanism for query execution when addressing different cyber situations.	Reasoning, querying
[9]	2016	Explored the overlap between different information exchange standards using ontologies and library science techniques.	Library science methods
[73]	2016	Proposed "net vulnerability", a metric for measuring security risks of networked systems.	Novel metric
[75]	2017	Leveraged an ontology to gain knowledge about an IoT environment.	OWL
[76]	2018	Presented a novel description logic-based formalism to model fuzzy and uncertain cyber-related information.	Fuzzy logic

Table 5. Cont.

Reference	Year Published	Brief Summary	Tools
[81]	2019	Demonstrated the efficiency of semantic technology within CTI use cases by using the STIX representation and the Maltego tool.	Semantics
[83]	2019	Introduced a SIEM-based KG and MalRank, a graph-based inference algorithm for maliciousness score estimation.	Graph inference algorithm
[66]	2019	Developed a model relying on ontologies to enhance the risk management capabilities of organizations.	OWL, Reasoning
[85]	2020	Developed an ontology-driven CFRaaS model for digital evidence gathering.	Description logic
[86]	2021	Proposed an APT attacks recommendation tool for attack scenario analysis and inference.	Similarity measures
[87]	2022	Proposed OnToRisk, an ontology-driven risk management analysis method.	Protege, OWL
[88]	2023	Proposed the KGaaS framework for the generation and maintenance of domain-specific KGs.	Cypher, Natural Language, RDF
[89]	2023	Developed a novel framework for cyber insurance policy suggestion.	RDF, SPARQL, IR, AI

Sane and Joshi [89] automated the process of extracting coverage and exclusion key terms and rules of cyber insurance policies with the use of IR and artificial intelligence (AI) methods, particularly Semantic Web and Modal Logic. Through this proposed system, users can find the optimal cyber insurance policy according to specified coverage criteria. In addition, the authors implemented a KG querying method using SparqlWrapper through Python and different Flask libraries. A web interface using Angular was ultimately built to help users compare the coverages and exclusions provided by distinct policies.

3.3. Machine Learning and Deep Learning Applications

Despite collecting a large number of studies focusing on knowledge extraction from unstructured data sources using DL methods, we decided to omit several works that did not explicitly generate or leverage an ontology or KG after extraction, as already highlighted in Section 2. Therefore, we arranged the collected works into two subcategories. The “DL methods in data extraction” subcategory (Section 3.3.1) contains all studies related to entity and relationship extraction from different sources using DL models, while the “Other uses of ML and DL” subcategory focuses on works that applied either ML or DL methods to problems such as link prediction, among others.

3.3.1. Deep Learning Methods in Data Extraction

Husari et al. [59] introduced TTPDrill, a tool for the automated extraction of threat actions and TTP construction from unstructured threat reports. They developed a novel ontology-based approach to map these actions to known attack patterns and techniques, enabling comprehensive understanding of the attack cycle. They utilized a text mining approach combining Natural Language Processing (NLP) and IR techniques to extract threat actions based on semantic relationships rather than syntax, enhancing the accuracy of threat action extraction. TTPDrill’s use of a threat–action ontology demonstrated precise classification, achieving 84% precision and 82% recall.

Piplai et al. [27] created a knowledge extraction pipeline for obtaining important knowledge from After-Action Reports (AARs). The pipeline is based on a custom Named Entity Recognition (NER) tool called Malware Entity Extractor (MEE). MEE utilizes Conditional Random Fields (CRF), Gibbs sampling, and Relationship Extractor (RelExt), a DL framework for predicting relationships between the extracted entities. After extraction of

relevant knowledge, an AAR ontology was constructed. The authors demonstrated the efficiency of their method by comparing the quality of knowledge obtained from different AARs based on queries executed on fused KGs and non-fused KGs.

A novel method for improving the NER process was proposed by Wu et al [90]. Their model comprised a Bidirectional Long Short-Term Memory (BiLSTM) layer for capturing sentence-level features and a CRF layer for labeling sentence sequences. A CTI ontology was used for corrections during the NER task. The model was tested on a dataset of unstructured text data, achieving an F1-score of 85.27%.

Another framework for the automation of open-source CTI management was developed by Gao et al. in [91]. SECURITYKG utilizes a back-end system for data collection, knowledge extraction, and KG construction. A security ontology and NLP tools are leveraged for extraction of entities and relationships from various openly available CTI reports. The functionalities of the proposed framework are accessed through a UI.

Open-CyKG, a system for obtaining information from APT reports and using them for a KG construction, was created by Sarhan and Spruit [92]. The proposed system employs an attention-based Open Information Extraction architecture based on Gated Recurrent Unit (GRU) networks for the extraction of relation triples. An NER model is used to automatically annotate cybersecurity entities. The constructed KG can be used by analysts to retrieve vital information. The authors evaluated the model on the Microsoft Security Bulletins dataset and a malware dataset from CTI reports, observing that it outperformed several state-of-the-art and baseline models in terms of recall, precision, and F1-score.

Zongxun et al. [93] created a model based on Bidirectional Encoder Representations from Transformers (BERT) [94], BiLSTM networks, and CRF to automate the analysis of APTs. Extracted threat actions were used to construct a cyber threat ontology. The ontology can be used for IoC and TTP generation. The model was evaluated on a small sample-relevant dataset and outperformed other threat action extractors, scoring 97% and 96% in recall and accuracy, respectively.

Li et al. [29] proposed a DL-based entity and relationship extraction system for CTI. Unstructured threat intelligence text data are used as input, with BERT, BiLSTM, and Convolutional Neural Network (CNN) architectures used for the extraction process. The model ultimately returns a KG stored as a Neo4j graph database. Its performance was assessed against 227 threat intelligence documents and compared to different baseline models, where it achieved the highest F1-score on the entity extraction, co-reference resolution, and relation extraction tasks.

Sun et al. [95] created a framework for extracting CTI triples from ATP reports, which they named APTKG. The authors also constructed APTOnt, an ontology based on STIX representation. A BiLSTM model is used for entity extraction along with CRF, BERT, CNN, and StanfordCoreNLP [96], while a BiGRU is used for inferring relations between the extracted CTI entities. An APT group-focused KG is ultimately constructed. When compared to alternative DL models, APTKG achieved better precision, recall, and F1-scores in CTI entity recognition. The authors characterized the joint model for entity and relationship inference as “effective”, highlighting its relatively high F1-score (70.86%).

A method for automating the CTI extraction of threat actions from unstructured data was developed by Li et al. [97]. K-CTIAA detects professional knowledge in a CTI KG through knowledge queries and uses it as an input into pretrained models for better understanding of the semantics of security-related terms. Another feature of K-CTIAA is the suggestion of countermeasures for rapid response against cyber attacks. The proposed method was tested on APTNotes [98], an open-source APT report repository, and outperformed CTI analysis tools such as ActionMiner [99], TTPDrill [59], and different NN architectures, achieving high precision (0.931), recall (0.951), and F1-score (0.941).

Liu and Zhan [100] used a different approach in knowledge extraction, leveraging Large Language Models (LLMs) such as ChatGPT for CTI KG construction. They created a ChatGPT-based pipeline for the collection of metadata and appropriate relationships, which were later used for the generation of triples. A CTI ontology was also used in this

process. The created triples were ultimately parsed again through the LLM to ensure their quality and validity. The proposed method was tested on thirteen threat intelligence reports, outperforming AttackKG [57] and REBEL [101] in terms of recall and F1-score. However, the authors reported that ChatGPT occasionally generated relations that diverged from the structure of the ontology.

Finally, Ren et al. [102] created CSKG4APT, a platform based on a novel algorithm utilizing BERT for threat knowledge recognition from bilingual documents. The platform results in the construction of an APT KG that adheres to different threat intelligence standards. During evaluation, information from English and Chinese CTI reports was extracted. CSKG4APT outperformed other relevant models such as BERT-LSTM, BERT-BiLSTM, and BERT-BiLSTM-CRF in terms of both micro- and macro-evaluation metrics.

The works related to DL methods for knowledge extraction from text data combined with ontology or KG construction or usage are summarized in Table 6.

Table 6. Works relevant to DL methods for knowledge extraction from text data combined with ontology/KG construction or usage. Records are arranged in ascending order based on their publication year and alphabetically by author name when the publication years of two or more papers coincide.

Reference	Year Published	Brief Summary	Technology Used	Evaluation Scores
[59]	2017	Introduced TPPDrill, a tool that automates threat action knowledge extraction from unstructured reports.	NLP, IR	Precision = 84% Recall = 82%
[27]	2020	Created a knowledge extraction pipeline from AARs to fill a cybersecurity KG.	CRF, Gibbs' sampling, RelExt	Precision Recall F1-score
[90]	2020	Proposed a model for enhancing NER.	BiLSTM, CRF	F1-score = 85.27%
[91]	2021	Proposed SECURITYKG, a system for automating the processing of open-source CTI.	NLP	-
[92]	2021	Created Open-CyKG for presenting APT reports as a queryable KG.	BiGRU, CRF, XLM-RoBERTa	Dataset: Microsoft (CTI) Recall = 98.7% (80.8%) Precision = 99.2% (78.9%) F1-score = 98.9% (79.8%)
[93]	2021	Used a DL model for knowledge extraction and constructed a cyber threat ontology.	BERT, BiLSTM, CRF	Recall = 97% Accuracy = 96%
[29]	2022	Proposed a CTI extraction system using DL models and constructed a KG.	BERT, BiLSTM, CNN	<u>Entity extraction</u> Precision = 79.02% Recall = 77.22% F1-score = 78.11 Accuracy = 74.13 <u>Coreference resolution</u> Precision = 65.58% Recall = 72.25% F1-score = 68.75 <u>Relation extraction</u> Precision = 79.61% Recall = 48.59% F1-score = 60.35

Table 6. Cont.

Reference	Year Published	Brief Summary	Technology Used	Evaluation Scores
[95]	2022	Proposed APTKG, a framework that constructs automatically KGs from open-source APT reports.	BiLSTM, CRF, BERT, CNN, Stanford CoreNLP	F1-score = 70.86%
[97]	2023	Proposed K-CTIAA to automate CTI analysis with pre-trained models and KGs.	Pre-trained models	Precision = 93.1% Recall = 95.1% F1-score = 94.1%
[100]	2023	Leveraged LLMs to build a CTI KG.	LLM, ChatGPT	Entity recognition (Relation extraction) Precision = 76% (90%) Recall = 82% (60%) F1-score = 78% (56%)
[102]	2023	Created CSKG4APT for the organization of APT knowledge from bilingual documents.	BERT	English Macro (Micro) Precision = 77.67% (78.52%) Recall = 69.74% (69.74%) F1-score = 73.20% (73.87%)

3.3.2. Other Uses of Machine Learning and Deep Learning

Dhungana and Upadhyaya [103] sought to tackle ongoing cybersecurity issues potentially threatening Nepal by applying Random Forest (RF) and DL models to the Coburg Intrusion Detection dataset along with an ontology mapping to introduce adaptive learning into the CTI system they developed.

A tool for detecting attack patterns and methods named Attack Hypothesis Generator (AHG) was developed by Elitzur et al. [104]. AHG implements both supervised and unsupervised recommendation algorithms along with collaborative filtering. Several different algorithms were implemented: Projected Description (ProjD), Link Prediction on Projected Description (LPProjD), Projected Attack (ProjA), Collaborative Filtering (CF), and Supervised Link Prediction (SupLP). The evaluation of the tool on the ATT&CK knowledge base revealed that link prediction based on collaborative filtering with distance and topological graph features assisted the analysis more effectively regardless of the size of the examined KG.

Pingle et al. [105] introduced RelExt, a Feed-Forward Neural Network (FFNN) used for inferring nontrivial relationships between cybersecurity-related entities based on an assessment of the similarity of their contextual vectors. Relationships that do not follow the STIX 2.0 schema are deemed meaningless. The proposed model was evaluated on a corpus of cybersecurity entities that originated from previously unseen documents and achieved 96.61% accuracy, predicting more than 700 relationships from Dark Caracal and CrossRat malware descriptions.

Mendsaikhan et al. [106] applied multiple ML and DL algorithms in a framework they proposed for quantitatively evaluating the relevance of cybersecurity text data. A custom NER model based on Stanford’s CRF Classifier and Google’s Universal Sentence Encoder (USE) was used, and several classification algorithms were employed. The authors generated a cybersecurity KG consisting of over 200,000 semantic triples from the Malware-TextDB [107], the CVE repository from the NVD, and from multiple forum discussions and news feeds related to cybersecurity. The results revealed that the logistic regression classifier performed the best when assessing the significance of text data, scoring 88% in classification accuracy.

Ding et al. [108] applied hierarchical clustering algorithms on the MITRE KG to mine hidden patterns of cybersecurity knowledge. A dictionary of hacker groups related to

cybersecurity technologies was built, to which the hierarchical clustering algorithm was applied to obtain clusters of cybersecurity technologies that shared a common hacker group feature. Experiments revealed that Peripheral Device Discovery, Data from Removable Media, and Junk Data attack methods were mostly common among different hacker groups. The authors asserted that the proposed method would be useful in predicting which technologies would be used in future cyber attacks.

In another work, Kriaa and Chaasbane [109] developed the SecKG schema. SecKG adheres to the information provided by MITRE ATT&CK and other open-source repositories. A KG that processes event logs is constructed, and a Knowledge Graph Convolutional Network (KGCN) is also employed for attack techniques prediction tasks based on each attack's neighborhood within the KG.

To verify the authenticity of CTI information primarily focused on malware or cyber attacks, Mitra et al. [110] proposed a framework based on NLP methods for examining the provenance of retrieved entities and relationships of a relevant KG. The verification is executed according to the Uniform Resource Locator (URL), publisher, and other important features of the data. A proposed provenance score quantifies the provenance of each datum and is later used to update the KG with provenance-specific classes and relations.

In an alternative use case of ML and DL application in CTI, Ranade et al. [111] used transformers such as GPT-2 to produce spurious CTI text data for later use in a data poisoning scenario with a KG and a cybersecurity corpus. The results showed that experts were not able to distinguish fake CTI from real data and were misguided to produce erroneous reasoning outputs, rendering the KG useless.

The effectiveness of cyber threat ontologies and Adversarial Machine Learning (AML) techniques in making predictions was studied by Yeboah-Ofori et al. [112]. APT attacks were modeled as a cyber threat ontology to accurately represent knowledge. Adversarial attacks based on RF and Gradient Boosting (GBoost) classifiers were subsequently applied. Analysis of the AML attacks showed that the use of ontological semantic reasoning capabilities was useful for validating a schema designed for vulnerability identification. In another study, Yeboah-Ofori et al. [113] utilized RF and GBoost algorithms along with cyber supply chain ontologies to map relationships between cyber attacks and cyber threat propagation and ultimately explore the cascading impact of different cyberattacks against the nodes of a supply chain network. The evaluation results revealed that the utilization of ontologies provided a better understanding of the correlations in the cyber supply chain security domain, achieving 80% accuracy in ML prediction of potential cyberattacks.

Wang et al. [114] proposed Relational Multi-Head Graph Attention Networks (R-MGAT), a model based on Graph Neural Networks (GNNs) for learning KG embeddings related to CTI with the aim of better understanding the semantics in text data. The authors also contributed a novel dataset of APT reports labeled using a text annotation tool (BRAT). R-MGAT achieved satisfying results in terms of Mean Rank, Mean Reciprocal Rank, and Hit@k in entity classification and link prediction tasks.

The Cyber threat indicators association prediction (Ctiap) model was created by Wang et al. in [115]; Ctiap combines matrix factorization and NN methods for the semantic and topological extraction processes, respectively. The authors created a graph dataset of threat indicators containing more than 20,000 samples and compared the performance of Ctiap against baseline methods that address the link prediction problem. Ctiap demonstrated the best accuracy and F1-score in mining threat indicator relationships.

Piplai et al. [116] demonstrated the effectiveness of pairing knowledge from cybersecurity KGs with offline Reinforcement Learning (RL) algorithms such as conservative Q-learning (CQL) for malware detection. The observed training time was lower when RL was guided by existing knowledge. For certain malware families, the proposed model achieved better performance when knowledge of these malware type existed; however, it was less efficient in generalizing the knowledge across all malware families.

Zhang et al. [117] proposed a CTI ontology and KG structure using data stemming from different sources, along with an inference model based on graph embedding algorithms

and reasoning rules (CTI-KGE). Their model applies link prediction tasks to augment the CTI landscape by automatically inferring the tail entities that potentially form relationships with the head entities. Using the proposed inference and reasoning rules, the analyst is able to automatically generate defense strategies. The method was evaluated in actual network system scenarios.

A model based on a GNN for edge information propagation was deployed by Zhang et al. [118]. The model was tested on three public datasets, achieving state-of-the-art results, and was later used for link prediction tasks on a constructed CTI KG (RCTI). A connectivity rate equal to 97% was observed between CTI and security requirements entities. The authors asserted that the model was suitable for management vulnerability detection.

Tables 7 and 8 contain all works related to ML and DL applications on ontologies or KGs besides entity and relationship extraction.

Table 7. Works related to various ML and DL applications to ontologies/KGs other than entity and relationship extraction. Records are arranged in ascending chronological order based on publication year and alphabetically by author name when the publication years of two or more papers coincide (continues in Table 8).

Reference	Year Published	Brief Summary	Technology Used	Evaluation Scores
[103]	2019	Used ML and DL methods with ontology mapping to address the cyber-related issues in Nepal.	RF, DL	RF (DL) Accuracy = 99.9% (99.93%) Precision = 99% (99.91%) Recall = 99.8% (99.95%)
[104]	2019	Proposed AHG, which applies link prediction techniques on CTI-derived KGs.	Link prediction, collaborative filtering	Algorithm Precision SupLP: 77% CF(k = 10): 66% LPPorjD: 66% CF(k = A): 0.59
[105]	2019	Introduced the RelExt system for predicting relationships between cybersecurity entities.	FFNN	Accuracy = 96.61%
[106]	2020	Proposed a novel ML- and DL-based mechanism for evaluating quantitatively the relevance of cybersecurity text data.	USE, CRFClassifier, Logistic Regression	Accuracy = 88%
[108]	2021	Used hierarchical clustering techniques to unveil hidden patterns in the MITRE KG.	Hierarchical clustering	-
[109]	2021	Created the SecKG schema, generated a KG and applied a KGCN for attack prediction.	KGCN	-
[110]	2021	Associated provenance with the entities and relations of a cybersecurity KG to ensure the authenticity of the data.	NLP	-
[111]	2021	Leveraged transformer-based models to produce fake CTI data and poison a cybersecurity KG and corpus.	Transformers, GPT-2	-

Table 8. Works related to various ML and DL applications on ontologies/KGs other than entity and relationship extraction. Records are arranged in ascending chronological order based on publication year and alphabetically by author name when the publication years of two or more papers coincide (continuation of Table 7).

Reference	Year Published	Brief Summary	Technology Used	Evaluation Scores
[112]	2021	Modelled APT attacks as an ontology and used supervised AML methods to deceive classifiers during training and testing.	AML	RF (GBoost) Before adversarial attacks: Accuracy = 72% (79%) After adversarial attacks: Accuracy = 22% (25%)
[113]	2021	Combined cyber supply chain security ontology concepts and ML for threat analysis and prediction.	RF, GBoost	Accuracy = 80%
[114]	2022	Proposed a GNN to learn KG embeddings and constructed a dataset from APT reports.	GNN	Entity Classification (Link Prediction) MR = 201 (223) MRR = 0.316 (0.345) Hits@1 = 0.232 (0.267) Hits@3 = 0.369 (0.394) Hits@10 = 0.607 (0.625)
[115]	2022	Proposed a Ctiap model that fuses semantic and topological features of a KG to predict relations.	Link prediction, NN	Accuracy = 93.08% Macro F1-score = 74.87%
[116]	2023	Examined the guidance of RL algorithms with KGs.	RL, CQL	Hits@10 > 65% for different malware families
[117]	2023	Proposed a CTI ontology and KG schema and used link prediction methods to infer new knowledge.	Link prediction	-
[118]	2024	Proposed EGNN for edge information propagation and applied link prediction on a constructed KG.	GNN	Dataset: FB15K-237, WN18RR, RCTI, WN18 hit@10: 0.532, 0.548, 0.324, 0.958 hit@3: 0.383, 0.489, 0.230, 0.953 hit@1: 0.258, 0.435, 0.160, 0.944 MR: 168, 2828, 3822, 250 MRR: 0.350, 0.474, 0.214, 0.950

4. Discussion

Our review identifies a noticeable imbalance between research focused exclusively on constructing CTI-based ontologies and KGs compared to studies exploring their utilization. While the number of the former works significantly surpassed the latter, this finding is not surprising. The vastness of the cyber threat domain requires more effort to map the existing knowledge into a formal representation; therefore, attention is primarily oriented

towards building novel ontologies. Despite observing a trend towards specialization over time (Tables 1–4), where ontologies and KGs focusing on specific subdomains such as malware [23], medical devices within IoT environments [46], cloud-based platforms [52,63], and smart-homes [54] are developed, there are other works that continue to develop more general CTI ontologies, highlighting the large amount of knowledge yet to be encoded.

A rise in studies that move beyond building knowledge bases and focus on creating functional frameworks capable of managing real-time tasks [24,28,41,61] is also observed. In addition, the constructed frameworks can be benefited from the advanced inference methods proposed in works such as Lu and Kokar [71], Mohsin and Anwar [38], and Najafi et al. [83], resulting in mechanisms capable of handling fuzzy and probabilistic knowledge [76].

Within the timeline we examined, ML and DL works related to ontologies and KGs emerge in 2017 for entity extraction tasks (NER for brevity) and in 2019 for other tasks such as link prediction. A close examination of Table 6 reveals that evaluation metrics for NER fluctuate depending on the task and dataset on which a proposed model was trained; however, the emerging use of pretrained models such as BERT has generally led to superior performance, with precision, recall, and F1-scores often exceeding 90% [92,93,97]. Notably, BERT has demonstrated satisfactory performance on NER tasks from Chinese reports, achieving scores greater than 80% in the aforementioned metrics. Thus, the use of pretrained models can not only reduce training times but also significantly enhance the performance of NER tasks across different languages.

An interesting pattern emerges when comparing the performance of models in entity extraction and relation extraction. According to reported results (Table 6), models generally perform better in NER than in relation extraction tasks, as indicated by the significant drop in recall and F1-score for relation extraction. However, precision in relation extraction is significantly higher than both recall and F1-score, and is either on par with [29] or higher than [100] in terms of precision in NER. This discrepancy is likely because relation extraction is a more complex task due to the difficulty of accurately capturing semantic relationships. Consequently, models become more conservative in predicting relations, reporting only those they deem most probable. This behavior reduces the occurrence of false positives, leading to higher precision, but results in missed true relations, leading to lower recall and consequently lower F1-scores. Conversely, the abundance of entities in cybersecurity text data makes identification of entities easier than the task of extracting their relationships.

A final and noteworthy remark concerns the influence of adversarial attacks on model performance and the validity of available data. As demonstrated by Yeboah-Ofori et al. [112], supervised classifiers such as RF and GBoost can be severely affected by adversaries in both the training and testing phases, raising concerns about the robustness of many proposed models. Furthermore, the systematic use of LLMs presents a double-edged sword. As Ranade et al. [111] have demonstrated, fake data that closely mimic real data can be generated to poison datasets, posing a significant threat to data integrity.

To conclude, despite the prominence of ML and DL in research on more elaborate inference techniques, works developing “less fancy” methods such as description logic-based or similarity measure-based inference remain appealing to researchers and produce considerable results. Finally, we would like to highlight the importance of unifying the proposed ontologies and KGs into a common representation that encapsulates all knowledge extracted from different studies. In this way, more complex and previously unseen patterns and relationships can be potentially unveiled when applying inference and link prediction techniques on all the knowledge encoded. Meanwhile, research should also aim to discover methods that shield the frameworks and the available data from possible adversarial attacks.

5. Concluding Remarks

5.1. Contribution

In this systematic review, we have explored the impact of semantic web tools as well as ML and DL methods on the CTI knowledge management domain. Several works demonstrate the importance of integrating those technologies into a common security analysis framework for organization, detection, classification, and prediction tasks [28,68,92,95,102,109].

In light of the enormous amount of vulnerability and cyber threat reports that are constantly being generated (and hopefully shared), the first layer of a potential security analysis platform would employ state-of-the-art DL models for entity and relationship extraction [29,92,93]. The extracted data would then be parsed into a standardized CTI-related ontology to ensure validity and verification [90]. Ideally, the ontology would be aligned [21] with different constructed CTI ontologies to capture all possible knowledge representations, promoting interoperability among different platforms. The analyst could then generate and inspect a relevant KG, augmented by advanced inference tools [95,115,117] and link prediction methods [117,118], to execute complex queries and potentially identify previously unseen patterns of behavior and relationships. A platform based on this scheme would allow security analysts to efficiently manage the massive influx of produced data, enhance the real time monitoring and assessment of cyber activities, and facilitate knowledge interchange among organizations.

Therefore, this work aims to serve as a reference for anyone interested in gaining a broad perspective on the available methods for processing semantic information within the domain of CTI, whether to explore further research directions or to incorporate the reviewed tools and methods in security analysis platforms.

5.2. Future Work

Future work can be guided in many directions. First, ontologies can be improved in quality, mainly by developing more elaborate evaluation criteria [24], as well as by refining the relationships and inference rules [24,65]. Additionally, methods for automatic adjustment of ontologies and KGs to incorporate recent knowledge can be researched [24,30,97]. Complex measures for assessing vulnerability scores, asset costs, or attack probability may be developed as well [117].

As stated in Section 4, the proposed ontologies can be merged into a common representation by matching the vocabularies and properties. This task can be quite difficult in cases where automated matching fails; therefore, it is essential to explore methods and frameworks that ease the ontology matching process, such as the Alignment platform, which offers the ability to collaboratively perform and validate the necessary matching [21].

The extraction process from different data sources can also be refined. Due to recent breakthrough involving LLMs, researchers can harness the continuous progress within this field to use more advanced models that can improve the accuracy of information extraction [100] and help in the fine-tuning of existing pretrained models [24]. Additionally, ontologies could serve as a robust baseline for correcting NER tasks [90]. By improving the extraction process, researchers will be able to expand the size of the datasets that can be used [119] and extend the developed frameworks to different languages and broader domains within the cybersecurity landscape [30,63], potentially resulting to more robust models with better ability to generalize [119]. Furthermore, defense mechanisms should be developed to mitigate the effect of adversarial attacks; for instance, analysts might track the source of CTI-related texts and assign provenance scores that quantify the validity of information [110,111]. An interesting approach for defending against suspicious CTI data generated by LLMs could be to detect linguistic mistakes [111] or phrases that are quite uncommon in relevant CTI sources.

Ultimately, all proposed frameworks and methods should be scaled to handle large amounts of data [30,64] and should be tested and evaluated based on realistic situations [30,62,120]. Therefore, collaboration between scholars and practitioners in the field

should be encouraged in order to bridge the gap between industry and purely academic work [30].

Author Contributions: Conceptualization, C.B., E.K.A., A.K.A. and L.I.; methodology, C.B., E.K.A., A.K.A. and L.I.; formal analysis, E.K.A. and A.K.A.; investigation, E.K.A. and A.K.A.; data curation, E.K.A. and A.K.A.; writing—original draft preparation, E.K.A. and A.K.A.; writing—review and editing, C.B., E.K.A., A.K.A., L.I., R.K. and S.O.; visualization, E.K.A. and A.K.A.; supervision, C.B., R.K. and S.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to acknowledge the ongoing INTERSOC project, funded by the European Union’s Cybersecurity and Trust Programme under Grant Agreement no. 101145853. Acknowledgments are also extended to the reviewers, whose remarks significantly improved the quality of this work.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AARs	After Action Reports
ACT	Semi-Automated Cyber Threat Intelligence
ADACO	Attack and Defense Analysis of Cybersecurity Ontology
AHG	Attack Hypothesis Generator
AI	Artificial Intelligence
AML	Adversarial Machine Learning
APT	Advanced Persistent Threat
BERT	Bidirectional Encoder Representations from Transformers
BiLSTM	Bidirectional Long Short-Term Memory
CFRaaS	Cloud Forensic Readiness as a Service
CNN	Convolutional Neural Network
CRF	Conditional Random Fields
CTI	Cyber Threat Intelligence
CVE	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Scoring System
CyBOX	Cyber Observable eXpression
DFAX	Digital Forensic Analysis eXpression
DL	Deep Learning
DRAF	Dynamic Risk Assessment Framework
FFNN	Feed-Forward Neural Network
GBoost	Gradient Boosting
GNN	Graph Neural Network
GRU	Gated Recurrent Unit
IDRS	Intrusion Detection and Response Systems
IDS	Intrusion Detection System
IoC	Indicators of Compromise
IODEF	Incident Object Description Exchange Format
IoT	Internet of Things
IP	Internet Protocol
IR	Information Retrieval
IT	Information Technology
KG	Knowledge Graph
KGaaS	Knowledge Graph as a Service
KGCN	Knowledge Graph Convolutional Network
LLM	Large Language Model
MAP	Mean Average Precision
MEE	Malware Entity Extractor

ML	Machine Learning
NER	Named Entity Recognition
NLP	Natural Language Processing
NN	Neural Network
NSM	Network Security Monitoring
NVD	National Vulnerability Database
OSINT	Open-Source Intelligence
OWL	Web Ontology Language
R-MGAT	Relational Multi-Head Graph Attention Network
RDF	Resource Description Framework
RelExt	Relationship Extractor
RF	Random Forest
RL	Reinforcement Learning
S-MAIDS	Semantic Model of Automated Intrusion Detection Systems
SIEM	Security Information Event Management
SMEs	Small- and Medium-sized Enterprises
STIX	Structured Threat Information eXpression
SWRL	Semantic Web Rule Language
TAL	Threat Agent Library
TEPA	Threat Evolution Prediction Algorithm
TTP	Tactics, Techniques and Procedures
UI	User Interfaces
URL	Uniform Resource Locator
USE	Universal Sentence Encoder
XSS	Cross-Site Scripting

References

1. The Latest 2024 Cyber Crime Statistics (Updated March 2024). Available online: <https://aag-it.com/the-latest-cyber-crime-statistics/> (accessed on 2 April 2024).
2. Top Concerns Industry Leaders Have about Cyberattacks in 2024 and beyond. Available online: <https://www.ibm.com/blog/top-concerns-industry-leaders-have-about-cyberattacks-in-2024-and-beyond/> (accessed on 2 April 2024).
3. Sun, N.; Ding, M.; Jiang, J.; Xu, W.; Mo, X.; Tai, Y.; Zhang, J. Cyber Threat Intelligence Mining for Proactive Cybersecurity Defense: A Survey and New Perspectives. *IEEE Commun. Surv. Tutorials* **2023**, *25*, 1748–1774. [CrossRef]
4. What Is Cyber Threat Intelligence? Available online: <https://www.microsoft.com/en-us/security/business/security-101/what-is-cyber-threat-intelligence> (accessed on 5 April 2024).
5. What Is Cyber Threat Intelligence? Available online: <https://www.cisco.com/c/en/us/products/security/what-is-cyber-threat-intelligence.html> (accessed on 5 April 2024).
6. Saeed, S.; Suayyid, S.A.; Al-Ghamdi, M.S.; Al-Muhaisen, H.; Almuhaideb, A.M. A Systematic Literature Review on Cyber Threat Intelligence for Organizational Cybersecurity Resilience. *Sensors* **2023**, *23*, 7273. [CrossRef] [PubMed]
7. Mavroidis, V.; Bromander, S. Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontologies within Cyber Threat Intelligence. In Proceedings of the 2017 European Intelligence and Security Informatics Conference (EISIC), Athens, Greece, 11–13 September 2017; pp. 91–98. [CrossRef]
8. Available online: <https://stixproject.github.io/> (accessed on 10 May 2024).
9. Asgarli, E.; Burger, E. Semantic ontologies for cyber threat sharing standards. In Proceedings of the 2016 IEEE Symposium on Technologies for Homeland Security (HST), Waltham, MA, USA, 10–11 May 2016. [CrossRef]
10. Merah, Y.; Kenaza, T. Proactive Ontology-based Cyber Threat Intelligence Analytic. In Proceedings of the 2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI), Tebessa, Algeria, 21–22 September 2021. [CrossRef]
11. Bratsas, C.; Kapsas, G.; Konstantinidis, S.; Koutsouridis, G.; Bamidis, P.D. A semantic wiki within moodle for Greek medical education. In Proceedings of the 2009 22nd IEEE International Symposium on Computer-Based Medical Systems, Albuquerque, NM, USA, 2–5 August 2009; pp. 1–6. [CrossRef]
12. Bratsas, C.; Koutkias, V.; Kaimakamis, E.; Bamidis, P.; Maglaveras, N. Ontology-Based Vector Space Model and Fuzzy Query Expansion to Retrieve Knowledge on Medical Computational Problem Solutions. In Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007; pp. 3794–3797. [CrossRef]
13. Antoniou, P.E.; Chondrokostas, E.; Bratsas, C.; Filippidis, P.M.; Bamidis, P.D. A Medical Ontology Informed User Experience Taxonomy to Support Co-creative Workflows for Authoring Mixed Reality Medical Education Spaces. In Proceedings of the 2021 7th International Conference of the Immersive Learning Research Network (iLRN), Eureka, CA, USA, 17–10 June 2021; pp. 1–9. [CrossRef]

14. Spyropoulos, A.Z.; Kornilakis, A.; Makris, G.C.; Bratsas, C.; Tsiantos, V.; Antoniou, I. Semantic Representation of the Intersection of Criminal Law & Civil Tort. *Data* **2022**, *7*, 176. [[CrossRef](#)]
15. Spyropoulos, A.Z.; Bratsas, C.; Makris, G.C.; Garoufallo, E.; Tsiantos, V. Interoperability-Enhanced Knowledge Management in Law Enforcement: An Integrated Data-Driven Forensic Ontological Approach to Crime Scene Analysis. *Information* **2023**, *14*, 607. [[CrossRef](#)]
16. Filippidis, P.M.; Dimoulas, C.; Bratsas, C.; Veglis, A. A unified semantic sports concepts classification as a key device for multidimensional sports analysis. In Proceedings of the 2018 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), Zaragoza, Spain, 6–7 September 2018; pp. 107–112. [[CrossRef](#)]
17. Filippidis, P.M.; Dimoulas, C.A.; Bratsas, C.; Veglis, A.A. A Multimodal Semantic Model For Event Identification On Sports Media Content. *J. Media Crit.* **2018**, *4*, 295–306.
18. Bratsas, C.; Chrysou, D.E.; Eftychiadou, E.; Kontokostas, D.; Bamidis, P.D.; Antoniou, I. Semantic Web Game Based Learning: An I18n approach with Greek DBpedia. In Proceedings of the LiLe@WWW, Lyon, France, 17 April 2012.
19. Kontokostas, D.; Bratsas, C.; Auer, S.; Hellmann, S.; Antoniou, I.; Metakides, G. Internationalization of Linked Data: The case of the Greek DBpedia edition. *J. Web Semant.* **2012**, *15*, 51–61. [[CrossRef](#)]
20. Lange, C.; Ion, P.; Dimou, A.; Bratsas, C.; Sperber, W.; Kohlhase, M.; Antoniou, I. Bringing Mathematics to the Web of Data: The Case of the Mathematics Subject Classification. In *Semantic Web: Research and Applications*; Simperl, E., Cimiano, P., Polleres, A., Corcho, O., Presutti, V., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 763–777.
21. Karampatakis, S.; Bratsas, C.; Zamazal, O.; Filippidis, P.M.; Antoniou, I. Alignment: A Hybrid, Interactive and Collaborative Ontology and Entity Matching Service. *Information* **2018**, *9*, 281. [[CrossRef](#)]
22. Bratsas, C.; Chondrokostas, E.; Koupidis, K.; Antoniou, I. The Use of National Strategic Reference Framework Data in Knowledge Graphs and Data Mining to Identify Red Flags. *Data* **2021**, *6*, 2. [[CrossRef](#)]
23. Rastogi, N.; Dutta, S.; Zaki, M.; Gittens, A.; Aggarwal, C. MALOnt: An Ontology for Malware Threat Intelligence. In Proceedings of the First International Workshop, MLHat 2020, San Diego, CA, USA, 24 August 2020; Volume 1271, pp. 28–44. [[CrossRef](#)]
24. Wang, P.; Dai, G.; Zhai, L. Event-Based Threat Intelligence Ontology Model. In Proceedings of the 5th International Conference, SciSec 2023, Melbourne, VIC, Australia, 11–14 July 2023; Volume 14299, pp. 261–282. [[CrossRef](#)]
25. Dora, J.; Nemoga, K. Ontology for Cross-Site-Scripting (XSS) Attack in Cybersecurity. *J. Cybersecur. Priv.* **2021**, *1*, 319–339. [[CrossRef](#)]
26. Grigoriadis, C.; Berzovitis, A.; Stellios, I.; Kotzanikolaou, P. A Cybersecurity Ontology to Support Risk Information Gathering in Cyber-Physical Systems. In Proceedings of the CyberICPS, SECPRE, ADIoT, SPOSE, CPS4CIP, and CDT&SECOMANE, Darmstadt, Germany, 4–8 October 2022; Volume 13106, pp. 23–39. [[CrossRef](#)]
27. Piplai, A.; Mittal, S.; Joshi, A.; Finin, T.; Holt, J.; Zak, R. Creating Cybersecurity Knowledge Graphs from Malware after Action Reports. *IEEE Access* **2020**, *8*, 211691–211703. [[CrossRef](#)]
28. Yang, F.; Han, Y.; Ding, Y.; Tan, Q.; Xu, Z. A flexible approach for cyber threat hunting based on kernel audit records. *Cybersecurity* **2022**, *5*, 11. [[CrossRef](#)]
29. Li, Y.; Guo, Y.; Fang, C.; Liu, Y.; Chen, Q. A Novel Threat Intelligence Information Extraction System Combining Multiple Models. *Secur. Commun. Netw.* **2022**, *2022*, 8477260. [[CrossRef](#)]
30. Ahmed, K.; Khurshid, S.K.; Hina, S. CyberEntRel: Joint extraction of cyber entities and relations using deep learning. *Comput. Secur.* **2024**, *136*, 103579. [[CrossRef](#)]
31. Khairkar, A.; Kshirsagar, D.; Kumar, S. Ontology for Detection of Web Attacks. In Proceedings of the 2013 International Conference on Communication Systems and Network Technologies, Gwalior, India, 6–8 April 2013; pp. 612–615. [[CrossRef](#)]
32. Strasburg, C.; Basu, S.; Wong, J. S-MAIDS: A semantic model for automated tuning, correlation, and response selection in intrusion detection systems. In Proceedings of the 2013 IEEE 37th Annual Computer Software and Applications Conference, Kyoto, Japan, 22–26 July 2013; pp. 319–328. [[CrossRef](#)]
33. Burger, E.; Goodman, M.; Kampanakis, P.; Zhu, K. Taxonomy model for cyber threat intelligence information exchange technologies. In Proceedings of the 2014 ACM Workshop on Information Sharing & Collaborative Security, Scottsdale, AZ, USA, 3 November 2014; pp. 51–60. [[CrossRef](#)]
34. Casey, E.; Back, G.; Barnum, S. Leveraging CyBOX™ to standardize representation and exchange of digital forensic information. *Digit. Investig.* **2015**, *12*, S102–S110. [[CrossRef](#)]
35. Available online: <https://cyboxproject.github.io/> (accessed on 21 May 2024).
36. Meckl, S.; Tecuci, G.; Boicu, M.; Marcu, D. *Towards an Operational Semantic Theory of Cyber Defense against Advanced Persistent Threats*; George Mason University: Fairfax, VA, USA, 2015; Volume 1523, pp. 58–65.
37. Falk, C. An Ontology for Threat Intelligence. In Proceedings of the European Conference on Cyber Warfare and Security, Munich, Germany, 7–8 July 2016.
38. Mohsin, M.; Anwar, Z. Where to Kill the Cyber Kill-Chain: An Ontology-Driven Framework for IoT Security Analytics. In Proceedings of the 2016 International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan, 19–21 December 2017; pp. 23–28. [[CrossRef](#)]
39. Qamar, S.; Anwar, Z.; Rahman, M.A.; Al-Shaer, E.; Chu, B.T. Data-driven analytics for cyber-threat intelligence and information sharing. *Comput. Secur.* **2017**, *67*, 35–58. [[CrossRef](#)]

40. Mavroeidis, V.; Josang, A.; ACM. Data-Driven Threat Hunting Using Sysmon. In Proceedings of the 2nd International Conference on Cryptography, Security and Privacy, Guiyang, China, 16–19 March 2018; pp. 82–88. [[CrossRef](#)]
41. Diederichsen, L.; Choo, K.K.; Le-Khac, N.A. A Graph Database-Based Approach to Analyze Network Log Files. In Proceedings of the 13th International Conference, NSS 2019, Sapporo, Japan, 15–18 December 2019; Volume 11928, pp. 53–73. [[CrossRef](#)]
42. Doynikova, E.; Fedorchenko, A.; Kotenko, I. Ontology of metrics for cyber security assessment. In Proceedings of the 14th International Conference on Availability, Reliability and Security, Canterbury, UK, 26–29 August 2019. [[CrossRef](#)]
43. Kim, M.; Dey, S.; Lee, S.W. Ontology-driven security requirements recommendation for APT attack. In Proceedings of the 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW), Jeju, Republic of Korea, 23–27 September 2019; pp. 150–156. [[CrossRef](#)]
44. Liu, Z.; Sun, Z.; Chen, J.; Zhou, Y.; Yang, T.; Yang, H.; Liu, J. STIX-based Network Security Knowledge Graph Ontology Modeling Method. In Proceedings of the 2020 3rd International Conference on Geoinformatics and Data Analysis, Marseille, France, 15–17 April 2020; pp. 152–157. [[CrossRef](#)]
45. Piplai, A.; Mittal, S.; Abdelsalam, M.; Gupta, M.; Joshi, A.; Finin, T. Knowledge Enrichment by Fusing Representations for Malware Threat Intelligence and Behavior. In Proceedings of the 2020 IEEE International Conference on Intelligence and Security Informatics (ISI), Arlington, VA, USA, 9–10 November 2020; pp. 47–52. [[CrossRef](#)]
46. Sills, M.; Ranade, P.; Mittal, S. Cybersecurity Threat Intelligence Augmentation and Embedding Improvement—A Healthcare Use Case. In Proceedings of the 2020 IEEE International Conference on Intelligence and Security Informatics (ISI), Arlington, VA, USA, 9–10 November 2020; pp. 62–67. [[CrossRef](#)]
47. Vrandečić, D.; Krötzsch, M. Wikidata: A free collaborative knowledgebase. *Commun. ACM* **2014**, *57*, 78–85. [[CrossRef](#)]
48. Available online: <https://accessgudid.nlm.nih.gov/> (accessed on 22 May 2024).
49. Hutschenreuter, H.; Çakmakçı, S.; Maeder, C.; Kemmerich, T. Ontology-based Cybersecurity and Resilience Framework. In Proceedings of the 7th International Conference on Information Systems Security and Privacy, Virtual, 11–13 February 2021; pp. 458–466. [[CrossRef](#)]
50. Mavroeidis, V.; Hohimer, R.; Casey, T.; Jesang, A. Threat Actor Type Inference and Characterization within Cyber Threat Intelligence. In Proceedings of the 2021 13th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 25–28 May 2021; pp. 327–352. [[CrossRef](#)]
51. Merah, Y.; Kenaza, T. Ontology-based Cyber Risk Monitoring Using Cyber Threat Intelligence. In Proceedings of the 16th International Conference on Availability, Reliability and Security, Vienna, Austria, 17–20 August 2021. [[CrossRef](#)]
52. Ammi, M.; Adedugbe, O.; Alharby, F.; Benkhelifa, E. Leveraging a cloud-native architecture to enable semantic interconnectedness of data for cyber threat intelligence. *Clust. Comput.* **2022**, *25*, 3629–3640. [[CrossRef](#)]
53. Bromander, S.; Swimmer, M.; Muller, L.; Jøsang, A.; Eian, M.; Skjøtskift, G.; Borg, F. Investigating Sharing of Cyber Threat Intelligence and Proposing A New Data Model for Enabling Automation in Knowledge Representation and Exchange. *Digit. Threat. Res. Pract.* **2022**, *3*, 1–22. [[CrossRef](#)]
54. Collen, A.; Nijdam, N. Can I Sleep Safely in My Smarthome? A Novel Framework on Automating Dynamic Risk Assessment in IoT Environments. *Electronics* **2022**, *11*, 1123. [[CrossRef](#)]
55. National Vulnerability Database. Available online: <https://nvd.nist.gov/> (accessed on 17 May 2024).
56. Huang, C.C.; Huang, P.Y.; Kuo, Y.R.; Wong, G.W.; Huang, Y.T.; Sun, Y.; Chang Chen, M. Building Cybersecurity Ontology for Understanding and Reasoning Adversary Tactics and Techniques. In Proceedings of the 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 17–20 December 2022; pp. 4266–4274. [[CrossRef](#)]
57. Li, Z.; Zeng, J.; Chen, Y.; Liang, Z. AttackKG: Constructing Technique Knowledge Graph from Cyber Threat Intelligence Reports. In Proceedings of the 27th European Symposium on Research in Computer Security, Copenhagen, Denmark, 26–30 September 2022; Volume 13554, pp. 589–609. [[CrossRef](#)]
58. Satvat, K.; Gjomemo, R.; Venkatakrishnan, V. Extractor: Extracting Attack Behavior from Threat Reports. In Proceedings of the 2021 IEEE European Symposium on Security and Privacy (EuroS&P), Vienna, Austria, 6–10 September 2021; pp. 598–615.
59. Husari, G.; Al-Shaer, E.; Ahmed, M.; Chu, B.; Niu, X. TTPDrill: Automatic and accurate extraction of threat actions from unstructured text of CTI Sources. In Proceedings of the 33rd Annual Computer Security Applications Conference, Orlando, FL, USA, 4–8 December 2017; Volume F132521, pp. 103–115. [[CrossRef](#)]
60. Rastogi, N.; Dutta, S.; Gittens, A.; Zaki, M.; Aggarwal, C. TINKER: A framework for Open source Cyberthreat Intelligence. In Proceedings of the 2022 IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Wuhan, China, 9–11 December 2022; pp. 1569–1574. [[CrossRef](#)]
61. Sharma, K.; Kumar, A. A Graph Database-Based Method for Network Log File Analysis. In Proceedings of the 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 16–17 December 2022; pp. 545–550. [[CrossRef](#)]
62. Akbar, K.; Rahman, F.; Singhal, A.; Khan, L.; Thuraisingham, B. The Design and Application of a Unified Ontology for Cyber Security. In Proceedings of the 19th International Conference, ICISS 2023, Raipur, India, 16–20 December 2023; Volume 14424, pp. 23–41. [[CrossRef](#)]
63. Compastí, M.; López Martínez, A.; Fernández, C.; Gil Pérez, M.; Tsarsitalidis, S.; Xylouris, G.; Mlakar, I.; Kourtis, M.; Šafran, V. PALANTIR: An NFV-Based Security-as-a-Service Approach for Automating Threat Mitigation. *Sensors* **2023**, *23*, 1658. [[CrossRef](#)]
64. Dora, J.; Hluchý, L.; Nemoga, K. Ontology for Blind SQL Injection. *Comput. Inform.* **2023**, *42*, 480–500. [[CrossRef](#)]

65. Sánchez-Zas, C.; Villagrà, V.; Vega-Barbas, M.; Larriva-Novo, X.; Moreno, J.; Berrocal, J. Ontology-based approach to real-time risk management and cyber-situational awareness. *Future Gener. Comput. Syst.* **2023**, *141*, 462–472. [[CrossRef](#)]
66. Riesco, R.; Villagrà, V.A. Leveraging cyber threat intelligence for a dynamic risk framework. *Int. J. Inf. Secur.* **2019**, *18*, 715–739. [[CrossRef](#)]
67. Duque-Ramos, A.; Fernandez-Breis, J.; Stevens, R.; Aussenac-Gilles, N. OQuARE: A SQuARE-based approach for evaluating the quality of ontologies. *J. Res. Pract. Inf. Technol.* **2011**, *43*, 159–176.
68. Zhang, S.; Su, X.; Shi, P.; Du, T.; Han, Y. Threat Modeling and Application Research Based on Multi-Source Attack and Defense Knowledge. *Comput. Mater. Contin.* **2023**, *77*, 349–377. [[CrossRef](#)]
69. Takahashi, T.; Kadobayashi, Y. Mechanism for linking and discovering structured cybersecurity information over networks. In Proceedings of the 2014 IEEE International Conference on Semantic Computing, Newport Beach, CA, USA, 16–18 June 2014; pp. 279–284. [[CrossRef](#)]
70. Takahashi, T.; Kadobayashi, Y.; Fujiwara, H. Ontological approach toward cybersecurity in cloud computing. In Proceedings of the International Conference on Security of Information and Networks, Taganrog, Russia, 7–11 September 2010.
71. Lu, S.; Kokar, M. A Situation Assessment Framework for Cyber Security Information Relevance Reasoning. In Proceedings of the 2015 18th International Conference on Information Fusion (Fusion), Washington, DC, USA, 6–9 July 2015; pp. 1459–1466.
72. Takahashi, T.; Landfield, K.; Kadobayashi, Y. *An Incident Object Description Exchange Format (IODEF) Extension for Structured Cybersecurity Information*; RFC 7203; RFC Editor: Marina del Rey, CA, USA, 2014. [[CrossRef](#)]
73. Monteiro, M.; Sarmiento, T.; Barreto, A.; Costa, P. A holistic approach to evaluate cyber threat. *STIDS* **2016**, *1788*, 64–68.
74. de Barros Barreto, A. Cyber-ARGUS Framework—Measuring Cyber-Impact on the Mission. Ph.D. Thesis, Instituto Tecnológico de Aeronáutica, São José dos Campos, São Paulo, Brazil, 2013.
75. Mozzaquatro, B.; Jardim-Goncalves, R.; Agostinho, C. Situation awareness in the Internet of Things. In Proceedings of the 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC), Madeira, Portugal, 27–29 June 2017; pp. 982–990. [[CrossRef](#)]
76. Sikos, L.F. Handling Uncertainty and Vagueness in Network Knowledge Representation for Cyberthreat Intelligence. In Proceedings of the 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–6. [[CrossRef](#)]
77. Bourai, S.B.; Mokhtari, A.; Khellaf, F. Poss–SROIQ (D): Possibilistic description logic extension toward an uncertain geographic ontology. *New Trends Databases Inf. Syst.* **2014**, *241*, 277–286.
78. Klinov, P.; Parsia, B. Understanding a probabilistic description logic via connections to first-order logic of probability. In Proceedings of the International Workshop on Uncertainty Reasoning for the Semantic Web, Karlsruhe, Germany, 26 October 2008; pp. 41–58.
79. Bal-Bourai, S.; Mokhtari, A. SROIQ (D): Possibilistic Description Logic for Uncertain Geographic Information. In Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, Morioka, Japan, 2–4 August 2016; pp. 818–829.
80. Bobillo, F.; Straccia, U. Reasoning with the finitely many-valued Łukasiewicz fuzzy description logic SROIQ. *Inf. Sci.* **2011**, *181*, 758–778. [[CrossRef](#)]
81. Aviad, A.; Wecel, K. Cyber Treat Intelligence Modeling. In Proceedings of the 22nd International Conference, BIS 2019, Seville, Spain, 26–28 June 2019; Volume 353, pp. 361–370. [[CrossRef](#)]
82. Available online: <https://www.maltego.com/> (accessed on 20 May 2024).
83. Najafi, P.; Mühle, A.; Pünter, W.; Cheng, F.; Meinel, C. MalRank: A Measure of Maliciousness in SIEM-based Knowledge Graphs. In Proceedings of the 35th Annual Computer Security Applications Conference, San Juan, PR, USA, 9–13 December 2019; pp. 417–429. [[CrossRef](#)]
84. Available online: <https://github.com/HewlettPackard/sandpiper> (accessed on 21 May 2024).
85. KEBande, V.; Karie, N.; Ikuesan, R.; Venter, H. Ontology-driven perspective of CFRaaS. *Wiley Interdiscip. Rev. Forensic Sci.* **2020**, *2*, e1372. [[CrossRef](#)]
86. Jung, J.W.; Park, S.H.; Lee, S.W. A Tool for Security Requirements Recommendation using Case-Based Problem Domain Ontology. In Proceedings of the 2021 IEEE 29th International Requirements Engineering Conference (RE), Notre Dame, IN, USA, 20–24 September 2021; pp. 438–439. [[CrossRef](#)]
87. Shaked, A.; Margalit, O. Sustainable Risk Identification Using Formal Ontologies. *Algorithms* **2022**, *15*, 316. [[CrossRef](#)]
88. Callyam, P.; Kejriwal, M.; Rao, P.; Cheng, J.; Wang, W.; Bai, L.; Siddhardh Nadendla, V.S.; Madria, S.; Das, S.K.; Chadha, R.; et al. Towards a Domain-Agnostic Knowledge Graph-as-a-Service Infrastructure for Active Cyber Defense with Intelligent Agents. In Proceedings of the 2023 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), St. Louis, MO, USA, 27–29 September 2023; pp. 1–8. [[CrossRef](#)]
89. Sane, K.; Joshi, K.; Mittal, S. Semantically Rich Framework to Automate Cyber Insurance Services. *IEEE Trans. Serv. Comput.* **2023**, *16*, 588–599. [[CrossRef](#)]
90. Wu, H.; Li, X.; Gao, Y. An Effective Approach of Named Entity Recognition for Cyber Threat Intelligence. In Proceedings of the 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China, 12–14 June 2020; pp. 1370–1374. [[CrossRef](#)]

91. Gao, P.; Liu, X.; Choi, E.; Soman, B.; Mishra, C.; Farris, K.; Song, D. A System for Automated Open-Source Threat Intelligence Gathering and Management. In Proceedings of the 2021 International Conference on Management of Data, Virtual, 20–25 June 2021; pp. 2716–2720. [CrossRef]
92. Sarhan, I.; Spruit, M. Open-CyKG: An Open Cyber Threat Intelligence Knowledge Graph. *Knowl.-Based Syst.* **2021**, *233*, 107524. [CrossRef]
93. Li, Z.; Li, Y.; Zhang, H.; Li, J. Construction of TTPS from APT Reports Using Bert. In Proceedings of the 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 17–19 December 2021; pp. 260–263. [CrossRef]
94. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the North American Chapter of the Association for Computational Linguistics, Mineapolis, MN, USA, 3–5 June 2019.
95. Sun, L.; Li, Z.; Xie, L.; Ye, M.; Chen, B. APTKG: Constructing Threat Intelligence Knowledge Graph from Open-Source APT Reports Based on Deep Learning. In Proceedings of the 2022 5th International Conference on Data Science and Information Technology (DSIT), Shanghai, China, 22–24 July 2022; pp. 1–6. [CrossRef]
96. Manning, C.D.; Surdeanu, M.; Bauer, J.; Finkel, J.R.; Bethard, S.; McClosky, D. The Stanford CoreNLP natural language processing toolkit. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, MD, USA, 23–25 June 2014; pp. 55–60.
97. Li, Z.; Li, Y.; Liu, Y.; Liu, C.; Zhou, N. K-CTIAA: Automatic Analysis of Cyber Threat Intelligence Based on a Knowledge Graph. *Symmetry* **2023**, *15*, 337. [CrossRef]
98. kbandla/APTnotes: Various Public Documents, Whitepapers and Articles about APT Campaigns. Available online: <https://github.com/aptnotes/data> (accessed on 16 May 2024).
99. Husari, G.; Niu, X.; Chu, B.; Al-Shaer, E. Using Entropy and Mutual Information to Extract Threat Actions from Cyber Threat Intelligence. In Proceedings of the 2018 IEEE International Conference on Intelligence and Security Informatics (ISI), Miami, FL, USA, 9–11 November 2018; pp. 1–6. [CrossRef]
100. Liu, J.; Zhan, J. Constructing Knowledge Graph from Cyber Threat Intelligence Using Large Language Model. In Proceedings of the 2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy, 15–18 December 2023; pp. 516–521. [CrossRef]
101. Huguet Cabot, P.L.; Navigli, R. REBEL: Relation Extraction By End-to-end Language generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*; Moens, M.F., Huang, X., Specia, L., Yih, S.W.T., Eds.; Association for Computational Linguistics: Punta Cana, Dominican Republic, 2021; pp. 2370–2381. [CrossRef]
102. Ren, Y.; Xiao, Y.; Zhou, Y.; Zhang, Z.; Tian, Z. CSKG4APT: A Cybersecurity Knowledge Graph for Advanced Persistent Threat Organization Attribution. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 5695–5709. [CrossRef]
103. Dhungana, S.; Upadhyaya, P. Optimal Cyber Threat Intelligent System for Nepal. In Proceedings of the 2019 Artificial Intelligence for Transforming Business and Society (AITB), Kathmandu, Nepal, 5 November 2019. [CrossRef]
104. Elitzur, A.; Puzis, R.; Zilberman, P. Attack hypothesis generation. In Proceedings of the 2019 European Intelligence and Security Informatics Conference (EISIC), Oulu, Finland, 26–27 November 2019; pp. 40–47. [CrossRef]
105. Pingle, A.; Piplai, A.; Mittal, S.; Joshi, A.; Holt, J.; Zak, R. Relext: Relation extraction using deep learning approaches for cybersecurity knowledge graph improvement. In Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Vancouver, BC, USA, 27–30 August 2019; pp. 879–886. [CrossRef]
106. Mendsaikhan, O.; Hasegawa, H.; Yamaguchi, Y.; Shimada, H. Quantifying the significance and relevance of cyber-security text through textual similarity and cyber-security knowledge graph. *IEEE Access* **2020**, *8*, 177041–177052. [CrossRef]
107. Lim, S.K.; Muis, A.O.; Lu, W.; Ong, C.H. MalwareTextDB: A Database for Annotated Malware Articles. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Vancouver, BC, Canada, 30 July–4 August 2017; Barzilay, R., Kan, M.Y., Eds.; Association for Computational Linguistics: Vancouver, BC, Canada, 2017; pp. 1557–1567. [CrossRef]
108. Ding, Z.; Cao, D.; Liu, L.; Yu, D.; Ma, H.; Wang, F. A Method for Discovering Hidden Patterns of Cybersecurity Knowledge Based on Hierarchical Clustering. In Proceedings of the 2021 IEEE Sixth International Conference on Data Science in Cyberspace (DSC), Shenzhen, China, 9–11 October 2021; pp. 334–338. [CrossRef]
109. Kriaa, S.; Chaabane, Y. SecKG: Leveraging attack detection and prediction using knowledge graphs. In Proceedings of the 2021 12th International Conference on Information and Communication Systems (ICICS), Valencia, Spain, 24–26 May 2021; pp. 112–119. [CrossRef]
110. Mitra, S.; Piplai, A.; Mittal, S.; Joshi, A. Combating Fake Cyber Threat Intelligence using Provenance in Cybersecurity Knowledge Graphs. In Proceedings of the 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, 15–18 December 2021; pp. 3316–3323. [CrossRef]
111. Ranade, P.; Piplai, A.; Mittal, S.; Joshi, A.; Finin, T. Generating Fake Cyber Threat Intelligence Using Transformer-Based Models. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 18–22 July 2021. [CrossRef]
112. Yeboah-Ofori, A.; Ismail, U.; Swidurski, T.; Opoku-Boateng, F. Cyber Threat Ontology and Adversarial Machine Learning Attacks: Analysis and Prediction Perturbance. In Proceedings of the 2021 International Conference on Computing, Computational Modelling and Applications (ICCMA), Brest, France, 14–16 July 2021; pp. 71–77. [CrossRef]

113. Yeboah-Ofori, A.; Mouratidis, H.; Ismai, U.; Islam, S.; Papastergiou, S. Cyber Supply Chain Threat Analysis and Prediction Using Machine Learning and Ontology. In Proceedings of the 17th IFIP WG 12.5 International Conference, AIAI 2021, Hersonissos, Greece, 25–27 June 2021; Volume 627, pp. 518–530. [[CrossRef](#)]
114. Wang, X.; Chen, R.; Song, B.; An, J.; Jiang, J.; Wang, J.; Yang, P. Learning Cyber Threat Intelligence Knowledge Graph Embedding with Heterogeneous Relation Networks Based on Multi-Head Relational Graph Attention. In Proceedings of the 2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta), Haikou, China, 15–18 December 2022; pp. 1796–1803. [[CrossRef](#)]
115. Wang, Y.; Lang, B.; Xiao, N.; Chen, Y. Cyber Threat Indicators Association Prediction Based on Weighted Fusion of Semantic and Topological Information. In Proceedings of the 2022 5th International Conference on Algorithms, Computing and Artificial Intelligence, Sanya, China, 23–25 December 2022. [[CrossRef](#)]
116. Piplai, A.; Joshi, A.; Finin, T. Offline RL+CKG: A hybrid AI model for cybersecurity tasks. In Proceedings of the AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023), San Francisco, CA, USA, 27–29 March 2023; Volume 3433.
117. Zhang, S.; Li, S.; Chen, P.; Wang, S.; Zhao, C. Generating Network Security Defense Strategy Based on Cyber Threat Intelligence Knowledge Graph. In Proceedings of the First International Conference, ICENAT 2022, Shenzhen, China, 15–17 November 2023; Volume 1696, pp. 507–519. [[CrossRef](#)]
118. Zhang, Y.; Chen, J.; Cheng, Z.; Shen, X.; Qin, J.; Han, Y.; Lu, Y. Edge propagation for link prediction in requirement-cyber threat intelligence knowledge graph. *Inf. Sci.* **2024**, *653*, 119770. [[CrossRef](#)]
119. Liu, C.; Wang, B.; Wang, Z.; Tian, J.; Luo, P.; Yang, Y. TCFLTformer: TextCNN-Flat-Lattice Transformer for Entity Recognition of Air Traffic Management Cyber Threat Knowledge Graphs. *Aerospace* **2023**, *10*, 697. [[CrossRef](#)]
120. Kaiser, F.K.; Dardik, U.; Elitzur, A.; Zilberman, P.; Daniel, N.; Wiens, M.; Schultmann, F.; Elovici, Y.; Puzis, R. Attack Hypotheses Generation Based on Threat Intelligence Knowledge Graph. *IEEE Trans. Dependable Secur. Comput.* **2023**, *20*, 4793–4809. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.