DISINFOX: A Threat Intelligence sharing platform for disinformation incidents

Master in Cybersecurity

Master's Thesis

Author:
Felipe Sánchez González
Thesis Advisors:
Javier Pastor Galindo
José Antonio Ruipérez Valiente

January 20, 2025

Facultad Informática Universidad Murcia

DISINFOX: A Threat Intelligence sharing platform for disinformation incidents

Using Cyber Threat Intelligence standards for interoperable modeling, sharing, and investigation of disinformation threats

Autor
Felipe Sánchez González

Tutor/es
Javier Pastor Galindo
Department of Computer Systems, Polytechnic University of Madrid
José Antonio Ruipérez Valiente
Department of Information and Communications Engineering, University of Murcia

Facultad de Informática

Master in Cybersecurity

UNIVERSIDAD DE MURCIA

Murcia, January 20, 2025

Acknowledgments

Me gustaría *destapar* esta sección de agradecimientos dando las gracias a mis compañeros de máster con los que he compartido un año y medio. No solo me han enseñado muchos conceptos técnicos y teóricos que desconocía, sino que he compartido con ellos alguna que otra experiencia que ha ganado un valor superior gracias a su compañía.

También me gustaría destacar este año y medio que he pasado dentro del equipo del CyberDataLab. En especial, al *autodenominado* grupo DISINFO, al cual le guardo un cariño especial, dado todo el compañerismo y comprensión mostrados. Este equipo ha servido como un hilo conductor perfecto para este TFM: cada reunión y charla informal ha inspirado muchas de las ideas que se plasman en este documento. En el futuro, a la vista del potencial y el talento entre sus filas, preveo algún que otro éxito, seguramente promovido por alguno de los proyectos que ya *asoman la pata* internamente. Mucha suerte a los estudiantes de doctorado, máster y grado que habitan el grupo. Os deseo lo mejor. También, gracias a Gregorio, Javier y José por confiar en mí y en mi trabajo desde el momento en que les propuse mi incorporación.

A Javier y José, también me gustaría agradecer la ayuda que me han brindado a la hora de realizar el TFM. No solo hablo de la redacción, donde me han transmitido la rigurosidad necesaria para la elaboración de un documento de estas características, sino también de la libertad que me han dado a la hora de proponer la línea de investigación que intuí más cómoda y potente para mis habilidades y gustos. Gracias al *expertise* que suman ambos, he podido abordar muchos desafíos que este trabajo ha generado. Os deseo mucha suerte en la gestión de DISINFO.

Por último, quiero agradecer a todo mi grupo de amigos la compañía durante esta época, estando en las buenas y en las malas, y siempre tendiéndome su mano cuando lo he necesitado. Este ha sido un año de cambios para todos en el que, sin duda, hemos crecido profesional y personalmente. Ojalá podamos seguir vinculados mucho más tiempo. Gracias por todo.

Finalmente, gracias a Silvia por estar ahí siempre, a mi lado, ante cualquier adversidad y revés que he tenido. Estoy convencido de que este paso habría sido más difícil sin tu apoyo y cariño. Gracias de corazón.

A Silvia.

Por regalarme tu amor y valentía, compañeros en cada reto, duda y logro hasta la culminación de este trabajo.

A Juan, Blasa, María Isabel y Consuelo.

Por sus consejos y constante apoyo durante mi vida.

Abstract

Cyber Threat Intelligence (CTI) has empowered cybersecurity teams worldwide by improving the quality and speed of their analysis for cybersecurity incidents through the establishment standards and specialized tools. These tools and frameworks facilitate correlation and collaboration across global communities, helping organizations stay informed about the evolving cyber threat landscape.

Despite its success in cybersecurity, CTI has yet to be leveraged for the systematic exchange and management of knowledge about disinformation threats, which are often described in unstructured natural language.

This thesis introduces DISINFOX, an open-source threat intelligence sharing platform designed to enable the interoperable exchange of disinformation incidents. DISINFOX adapts disinformation-related information to a CTI-compliant format by incorporating several key elements. First, it utilizes the DISARM framework, which provides a matrix similar to MITRE ATT&CK to characterize the tactics, techniques, and procedures (TTPs) of disinformation incidents. Second, a custom mapping codifies these TTPs along with other relevant information, such as actors and targeted countries, into the STIX2 standard. Finally, the platform integrates with OpenCTI to validate its interoperability, alongside a user-friendly, web-based frontend for visualizing, managing, and analyzing incidents.

DISINFOX employs a modular, containerized architecture comprising four main components: a backend providing a RESTful API independent of other modules, a frontend serving as the ingestion entry point for disinformation incidents, a public API enabling other CTI solutions to extract incidents from the platform, and the DISINFOX OpenCTI connector that validates the interoperability of incidents within a mature CTI tool.

The platform's capabilities were validated through the modeling, storage, sharing, and consumption of over 100 disinformation incidents, demonstrating its technical feasibility. This work highlights the potential of using CTI concepts and tools to systematically combat disinformation threats.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Context

Disinformation has been on the rise in the last decade. However, the idea of influencing adversaries' minds to gain an advantage is not new. From propaganda and deception in World War II [2] to the global disinformation surge during the COVID-19 pandemic [3], the malicious spread of fake information to influence opinions and decisions has been a constant practice throughout history, especially in the geopolitical arena [4].

The popularity and widespread adoption of social networks have altered the dynamics of disinformation. Social networks, especially among younger generations, have become the primary source of information [5]. These platforms, originally intended to connect people and share ideas globally, have become ideal channels for disinformation campaigns to proliferate.

Moreover, attribution has become more challenging. Disinformation campaigns can be executed anonymously, leveraging the inherently decentralized and obfuscated nature of the Internet. Also, the distribution of disinformation has become effortless: malicious content can be spread worldwide at incredible speed, often automatically, drastically reducing costs and amplifying impact for the actors behind these campaigns.

As such, disinformation, much like cyberattacks, constitutes a real hazard in the digital realm and, by extension, can be categorized as a cybersecurity threat that can benefit from the latest advances in this discipline [6, 7, 8, 9].

1.2 Cybersecurity approach to disinformation

At first glance, disinformation campaigns might appear as a rhetorical or sociopolitical problem, targeting a nation's narrative or societal cohesion. However, this perspective is limited. In the modern era, disinformation leverages the opportunities afforded by cyberspace—its reach, speed, automation, anonymity, and scalability—to achieve strategic objectives. This results in unmanageable disinformation attacks with the traditional handling and analysis that cannot be computed or shared in easy ways.

This shift has brought disinformation into the realm of cybersecurity, as it shares key characteristics with other cyber threats. Since 2014, the conflict in Ukraine has demonstrated the use of disinformation as a tool of geopolitical influence [1] [10]. In

[1] Although information operations have broader strategies, disinformation is a big part of them.

many cases, these disinformation campaigns have been executed alongside cyber operations, aimed at destabilizing or degrading Ukraine's infrastructure and morale [11]. These examples have raised awareness among policymakers about the need to treat disinformation as a cybersecurity concern [12, 13].

The discipline of cybersecurity has developed robust analytical frameworks in response to the constant evolution of cyber threats. Focused in these response, we find Cyber Threat Intelligence (CTI), which can be defined as "a discipline focused on understanding the capabilities, intent, motivations, and opportunities of relevant cyber adversaries and their associated Tactics, Techniques and Procedures (TTPs)" [14]. In addition to this, CTI analysis takes advantage of one of the key aspect of cyberattacks: the traces and data left behind in cyberspace. CTI's methodologies and tools, help in the identification and classification of adversarial behaviour. Frameworks like the MITRE ATT&CK [15] provide a standard characterization of the TTPs used in cyber attacks, userful for later correlation and retrspective analysis. Programs such as the Common Vulnerabilities and Exposures (CVE) and Common Vulnerability Scoring System (CVSS) generate common knowledge and scoring about found vulnerabilities, vital to understand and classify this vulnerabilities to prioritize mitigation actions or alert organizations. Platforms such as OpenCTI or MISP help to manage and correlate data about cyber attacks by providing useful views and search operations that improve the retrospective analysis in investigations. Finally, feeds and threat intelligence exchange platforms like Alienvault Open Threat Exchange (OTX) or DigitalSide Threat-Intel Repository serve as a live repository of the latest TTPs, Indicators of Compromise (IoCs) and other pieces of evidence found in attacks. All these platforms can be interconnected in simple ways, as the content that is shared is described with standards like Structured Threat Information eXpression (STIX2) or Trusted Automated Exchange of Intelligence Information (TAXII), which grants interoperable and structured knowledge.

All these CTI solutions have been proven to be useful in countering traditional cyber threats in the cybersecurity realm, however, none of them have been used to take advantage of its powerful investigation and alerting capabilities to try to combat disinformation threats at large scale. In fact, current countermeasures to disinformation campaigns mainly just include manual fact-checking (e.g., Comprobado [2], Newtral [3] or VoxCheck [4]) and per-case reports in natural language which this work considers insufficient. The relying in non standardized ways of describing the disinformation incidents leads to a limited sharing and processing as there are no interoperable or structured

[2] https://comprobado.es/
[3] https://www.newtral.es
[4] https://voxukraine.org/en/voxcheck

1.3 Cyber Threat Intelligence as a tool against disinformation

To improve the handling of the large set of disinformation incidents and evidence, several frameworks and models have been developed to provide structured ways of modeling disinformation [1, 16]. While frameworks like ABCDE (used by European Union (EU) institutions) or SCOTCH provide a fast and abstracted view of a disinformation incident for agile and informed decision-making, others, such as the Disinformation Analysis and Risk Management (DISARM) framework, offer a more detailed and systematic approach by characterizing incidents using TTPs.

The DISARM framework bridges the CTI process of describing an incident with TTPs to the disinformation domain. It achieves this by incorporating concepts from established cybersecurity models, such as the MITRE ATT&CK matrix and the Cyber Kill Chain. By organizing techniques into tactics and aligning them with the phases of a disinformation campaign, DISARM provides a structured matrix that allows analysts to identify and classify specific techniques used in disinformation operations. Furthermore, it facilitates interoperability with CTI platforms by offering a direct mapping of its TTPs to the STIX2 standard and providing an official OpenCTI connector.

As we see, the intersection between disinformation and CTI lies in sharing and understanding adversarial behavior through structured analysis. CTI methodologies, when applied to disinformation, enable the standardization of incident representation and integration with existing cybersecurity tools. This intersection enhances the analytical depth in disinformation studies and provides an opportunity to leverage mature CTI practices to combat disinformation effectively.

1.4 Motivation

Current disinformation databases lack interoperable and structured data capable of easing the processing of disinformation incidents by computers and interconnected systems. These properties are crucial to storing, sharing, managing, and retrospectively analyzing such incidents at a large scale effectively. The treatment of disinformation threats as a cybersecurity issue necessitates adapting disinformation incidents to standardized formats that ensure their computability and shareability with other systems.

In contrast, CTI concepts and tools in the cybersecurity domain provide the standards and methodologies required to improve the current manual and often incomputable management of disinformation threats. Frameworks, matrices, and threat intelligence platforms offer efficient mechanisms to manage cyber incidents, which can be adapted to disinformation threats, as demonstrated by the DISARM framework.

This work addresses these gaps by designing and implementing *DISINFOX* , a CTI platform tailored specifically to disinformation incidents. The platform models disinformation incidents using standardized formats like STIX2, ensuring seamless inter-

operability with existing CTI tools and frameworks. *DISINFOX* provides a scalable architecture that serves both non-technical users through an intuitive web interface and technical users via programmatic APIs for integration with other CTI solutions. Furthermore, the platform is validated by ingesting and managing real-world disinformation incidents, showcasing its capability to bridge the gap between disinformation and cybersecurity practices effectively. To the best of our knowledge, this represents the first technical approach to integrate disinformation incidents into the CTI ecosystem.

2 State of the Art

The growing prevalence of disinformation as a tool for influence and manipulation, particularly in geopolitical contexts, has prompted the development of various platforms and frameworks designed to address this challenge. This section explores related works in the realms of disinformation modeling, threat intelligence platforms, and disinformation databases. The platforms most aligned with the objectives of *DISINFOX* are analyzed, highlighting their main functionalities, strengths, and limitations. Finally, a comparison is presented to evaluate their key characteristics and identify gaps, focusing on the need for an interoperable and structured platform capable of effectively storing and managing disinformation incidents.

2.1 Related works

In the realm of CTI, several platforms and frameworks have been developed to facilitate the sharing and analysis of threat data. The Distributed Security Framework for Reliable Threat Intelligence Sharing [17] emphasizes the importance of a decentralized approach to enhance the reliability and timeliness of shared threat information. Similarly, the Malware Information Sharing Platform (MISP) [18] provides an open-source solution for collecting, storing, and distributing IoCs among organizations, promoting collaborative defense mechanisms. Addressing the need for contextual awareness, **(author?)** [19] propose a context-aware CTI exchange platform, which integrates various data sources to enrich the intelligence gathered, thereby improving the relevance and accuracy of threat assessments. Focusing on the African context, a CTI platform tailored for organizations incorporates data from social media platforms like Twitter, enhancing situational awareness despite not specifically targeting disinformation [20]. Furthermore, a platform designed for correlating CTI from Open Source Intelligence (OSINT) sources demonstrates the effectiveness of aggregating publicly available data to identify potential threats [21]. Leveraging machine learning techniques, the inTIME framework [22] automates the gathering and analysis of web data for CTI, showcasing the potential of artificial intelligence in enhancing cybersecurity measures. Additionally, the TSTEM platform [23] employs cognitive computing to collect CTI from diverse online sources, including social media and websites, facilitating real-time threat detection and analysis. These initiatives underscore the critical role of structured and interoperable CTI platforms in strengthening cybersecurity defenses across various sectors.

Regarding disinformation threats, there are public databases and works that gather disinformation incidents in large quantities. For example, EUvsDisinfo [24], managed by the East Stratcom Task Force, gathers over 18,200 reports on disinformation incidents with summaries and some fixed properties. Similarly, Disinfodex [25], supported by the Harvard Berkman-Klein Center, documents 379 disinformation campaigns on platforms like Google and Facebook, including details about removed resources and policy violations. Additionally, initiatives such as the Media Manipulation Casebook [26] with 36 entries and the 2024 DFRLab's Foreign Interference Attribution Tracker (FIAT) [27] with 86 entries expand on these efforts by coding disinformation campaigns with some variables and visualizing trends. These databases are valuable but none provide a properly structured or interoperable format for sharing, which is essential given the amount data related with disinformation incidents. In this regard, [28] performs an analysis of election-related disinformation campaigns from 2014 to 2024, employing the DISARM framework to model the analyzed incidents and resulting in a rich dataset with 81 campaigns.

However, these disinformation-based repositories are not implementing homogeneous and standardized sharing methodologies for CTI, making it difficult to programmatically consume that intelligence. Considering that threat exchange solutions evidence the utility of community-driven intelligence sharing, a similar approach could be applied to manage disinformation campaigns. Recent initiatives like the Defending Against Deception Common Data Model (DAD-CDM) [29] initiative, launched by OASIS in 2023, have the goal of introducing a common data model for normalizing and sharing information on disinformation campaigns using the STIX standard and using the DISARM framework's advances. Moreover, OpenCTI[30], a popular open-source solution by Filigran for threat exchange and CTI management can serve as a merging point of different CTI feeds and is able to interact with STIX objects inbounds and outbounds. OpenCTI already has a DISARM connector [31] that enables the platform to build reports with the DISARM TTPs and has adapted its solution to better represent disinformation threats in its platform [32]. Nevertheless, the connector is quite basic, as it only ingests the TTPs from DISARM converted to STIX2 objects and its own matrix. This is primarily aimed at generating reports within the OpenCTI platform. However, it is far from enabling the automatic ingestion of disinformation incidents with a database or dataset.

Considering the gaps in the standardization of disinformation data sharing and the lack of comprehensive solutions for managing disinformation as part of the broader CTI ecosystem, this Master Thesis addresses these challenges by proposing a platform that models disinformation incidents using STIX2 objects. This approach allows for seamless interoperability between disinformation data and established CTI systems. By combining the DISARM framework with the flexibility of STIX, this work contributes to the development of a structured, scalable solution that bridges the current gaps in disinformation threat exchange and enhances the capability to address the growing problem of disinformation in a standardized, actionable manner. In order to describe

similar successful threat intelligence platforms, we include Section 2.2

2.2 Threat exchange platforms

In this section, we explore the capabilities and unique features of four prominent CTI platforms: EclecticIQ Threat Intelligence Platform, MISP, OpenCTI, and OTX. These platforms were selected due to their wide adoption and significant impact on the CTI landscape. A comparative table summarizing their key characteristics is provided at the end of this section.

EclecticIQ Threat Intelligence Platform

EclecticIQ Threat Intelligence Platform [33] is a CTI platform launched in 2014 designed to support the entire threat intelligence lifecycle, from collection to dissemination. It provides a highly customizable environment tailored to organizational needs, supporting both structured and unstructured data. The platform uses its own format, called EclecticIQ JSON (EIQ JSON), for internal representation, but it also integrates with various data feeds and supports major CTI standards, including STIX2 and TAXII2.

EclecticIQ facilitates intelligence management through an intuitive user interface and advanced visualizations. Its graph-based analysis tools allow users to map relationships between entities, such as threat actors, campaigns, and indicators. The platform's ability to handle large volumes of data, combined with machine learning-enhanced analytics, makes it suitable for enterprises with complex threat landscapes.

The platform provides APIs for automation and integrates seamlessly with SIEMs and other security infrastructure. Although EclecticIQ is closed-source, it offers enterprise-level support and customization options. Its focus on scalability and multi-user collaboration has positioned it as a leading choice for organizations seeking robust CTI capabilities.

Malware Information Sharing Platform (MISP)

Malware Information Sharing Platform (MISP) [18] is an open-source threat intelligence sharing platform widely used in the CTI community. Its primary goal is to enhance collaboration among security practitioners by facilitating the sharing of IoCs and threat intelligence.

MISP supports the STIX2, TAXII and OpenIoC standards and provides an intuitive web-based interface for managing threat data. Its extensibility is one of its strongest attributes, offering numerous plugins and modules for data enrichment, export, and analysis. The platform allows organizations to create and maintain private or public sharing communities, enforcing granular access controls through Traffic Light Protocol (TLP) classifications.

MISP excels in its ability to process large volumes of structured threat data efficiently. However, it focuses more on sharing and less on advanced analytics or visualization compared to other platforms. As a free, open-source solution, MISP is particularly attractive to small and medium-sized organizations with limited budgets.

Open Cyber Threat Intelligence (OpenCTI) Platform

OpenCTI is an open-source CTI platform born in 2018 and developed by Filigran that provides a unified environment for managing and analyzing threat intelligence. It emphasizes interoperability by adhering to CTI standards such as STIX2 and TAXII2 and integrates seamlessly with other tools, including MISP, TheHive or AlienVault OTX.

The platform features a user-friendly web interface with graph-based visualizations for mapping relationships between threat entities. Its ability to store and analyze technical, tactical, and strategic intelligence makes it suitable for a wide range of use cases, from incident response to high-level threat assessments.

OpenCTI supports APIs for automation and offers a modular architecture with connectors such as the MITRE ATT&CK framework, making it highly extensible. Unlike some closed-source solutions, OpenCTI benefits from an active open-source community that continuously contributes enhancements and integrations. Its real-time data processing and focus on collaboration make it a valuable tool for organizations of all sizes.

AlienVault Open Threat Exchange (OTX)

AlienVault Open Threat Exchange (OTX) [34] is a cloud-based CTI platform managed by LevelBlue focused on community-driven threat intelligence sharing. Users can contribute and access a vast repository of threat data, including IoCs, malware samples, and campaign information.

OTX provides compilance with standards such as STIX2, TAXII and OpenIoC and intuitive web interface and an API for automated data ingestion and extraction. It is unique in its use of “pulses”, curated collections of threat data related to specific campaigns or threat actors. This approach simplifies the dissemination of actionable intelligence to security teams.

While OTX is free to use, its advanced features are tied to AlienVault’s commercial offerings, such as its Unified Security Management (USM) platform. Despite this limitation, OTX remains a popular choice for organizations seeking community-driven intelligence without significant financial investment.

Comparison of Platforms

Table 2.1 provides a summary of the key characteristics of the analyzed CTI platforms.

Feature	EclecticIQ	MISP	OpenCTI	OTX	★ *DISINFOX*
Open Source	-	X	X	-	X
Cost	Paid	Free	Free	Free	Free
APIs	X	X	X	X	X
Interoperable standards	STIX2, TAXII2	STIX2, TAXII, OpenIOC	STIX2, TAXII2	STIX2, TAXII, OpenIOC	STIX2
Feed	-	X	-	X	X
Visualization	Advanced	Limited	Advanced	Basic	Basic
Target users	Enterprises	All sizes	All sizes	Community-focus	Community-focus
Disinformation focus	-[1]	-	-	-	X

[1] It provides a new data model for integrating disinformation incidents [32].

Table 2.1: Comparison of CTI Platforms

In all the selected solutions, several shared features reflect their foundational role in the CTI landscape. A critical commonality is the provision of APIs across all platforms, which enhance automation and integration capabilities. These APIs allow for seamless interaction with other systems, enabling users to fetch, analyze, and share threat intelligence programmatically. This automation has become increasingly important in managing the growing volumes of threat data generated in modern security environments.

Cost and accessibility are important differentiators among these platforms. While MISP, OpenCTI, OTX, and *DISINFOX* are free to use, EclecticIQ is a proprietary solution that caters to large enterprises, offering extensive features at a premium price. The open-source nature of MISP, OpenCTI, and *DISINFOX* further enhances their accessibility, allowing smaller organizations and individual researchers to deploy and adapt them without significant financial constraints.

Interoperability is essential for this platforms, facilitated through widely adopted standards such as STIX2 and TAXII. All platforms support structured and standardized data exchange, with MISP and OTX going further by also including compatibility with OpenIOC, which caters to legacy systems. *DISINFOX* , while focusing exclusively on STIX2, aligns its design with a lightweight and specialized approach, emphasizing ease of integration for disinformation-specific use cases.

The inclusion of data feeds sets MISP, OTX, and *DISINFOX* apart from the rest. These platforms offer curated and regularly updated feeds that provide actionable intelligence to their users. This feature proves invaluable for community-driven solutions, as it simplifies the process of accessing ready-made intelligence without requiring significant prior input. In contrast, EclecticIQ and OpenCTI focus more on advanced user-driven data ingestion and customization, reflecting their enterprise-oriented design.

Visualization capabilities are key to analyzing complex threat data effectively. EclecticIQ and OpenCTI provide advanced visualization tools, such as graph-based analysis, to map relationships between threat actors, campaigns, and IoCs. These tools enable analysts to uncover hidden connections and make informed decisions quickly. MISP,

OTX, and *DISINFOX* , on the other hand, prioritize simplicity, offering more basic visualization features. This trade-off reflects their focus on accessibility and community collaboration rather than advanced analytics.

However, the most significant point of divergence lies in the focus areas of the platforms. EclecticIQ, MISP, and OTX primarily cater to traditional cybersecurity use cases, with no explicit mechanisms for addressing disinformation. OpenCTI have its own extension for disinformation. However, it is not easy to use, as it is mixed with all the other cybersecurity concepts. *DISINFOX* , in contrast, is purpose-built to fill this gap. By leveraging the DISARM framework's TTPs and integrating them into a STIX2-compliant format, *DISINFOX* provides a structured and interoperable approach to analyzing and sharing intelligence about disinformation campaigns. This focus on disinformation distinguishes *DISINFOX* from its peers, making it uniquely suited to modern and large scale disinformation threats.

About the target users, EclecticIQ is tailored to large enterprises, providing extensive tools for managing complex threat landscapes and supporting advanced customization for large-scale deployments. MISP and OpenCTI are more focused in the collaboration of single organization or controlled groups of trusted entities. OTX and *DISINFOX* , in contrast, emphasize community-driven collaboration, relying on user-contributed reports and shared intelligence to foster a decentralized and open approach. This design makes them particularly suitable for smaller organizations, independent analysts, and researchers who benefit from the collective insights of a global community without the need for extensive internal infrastructure.

The analysis of these platforms highlights the need for a dedicated solution to address disinformation threats. While existing CTI platforms excel in managing cybersecurity incidents, they fall short in modeling, analyzing, and sharing intelligence about disinformation. *DISINFOX* bridges this gap by providing:

- A specialized framework for disinformation incidents, incorporating TTPs from DISARM to ensure structured and actionable intelligence.
- Interoperability through STIX2, enabling seamless integration with established CTI platforms such as OpenCTI.
- Community-focused accessibility and lightweight deployment, catering to a wide range of users while addressing the specific challenges posed by disinformation campaigns.

DISINFOX represents a novel and essential advancement in the extension of CTI capabilities to the disinformation domain. Its design not only fills a critical gap but also aligns with the principles of modern CTI, enabling organizations to respond rapidly and effective to disinformation threats.

3 Objectives and methodology

The primary objective of this Master Thesis is to **propose *DISINFOX* , an open-source threat intelligence exchange platform** designed to enable the real-time, interoperable exchange of disinformation incidents with client-side CTI consumers. By using CTI standards and methodologies, *DISINFOX* provides a centralized platform for storing, managing, and analyzing disinformation incidents, integrating seamlessly with existing CTI tools to enhance the detection, investigation, and mitigation of this evolving threat. To achieve this, the following sub-objectives have been defined:

1. Define a mapping between the evidence generated in a disinformation incident and a standardized, computable language such as STIX2, establishing the base data model for *DISINFOX* . Use this mapping to model and structure a real disinformation incident as a use case to validate *DISINFOX* .

2. Design and implement the architecture of *DISINFOX* to effectively leverage the defined data model, providing a modular and interoperable client-server framework for the sharing of disinformation incidents.

3. Validate the lifecycle of the stored incidents in *DISINFOX* , from their creation to their integration into other CTI solutions, such as OpenCTI.

This Master Thesis has followed the next methodology:

1. Review of the main state-of-the-art works regarding disinformation modeling, frameworks and databases. This was done mainly with Google and Google Scholar, using Mendeley to store all the relevant references found.

2. Study of CTI standards for threat intelligence sharing, current threat intelligence exchange platforms and their use.

3. Definition and proposal of the project given the identified gaps.

4. Selection of framework to model disinformation incidents and building of new mapping for codifying disinformation incidents with the STIX standard and the selected modeling framework.

5. Design of the DISINFOX's architecture and tech stack.

6. Development and deployment of the full framework: backend, frontend, public API and a connector for a chosen CTI solution.

7. Documentation of the methodology, design, results, and findings through the writing of this thesis.

The different steps of the work were established through weekly meetings with the Thesis Advisors, which also served to validate the previous progress. Additionally, the design and development followed an iterative approach, with each cycle planning the next steps and functions based on the progress made in the previous iterations.

The code of the project was tracked with Git for version control, using GitHub to store the publicly available repository[1].

[1] `https://github.com/CyberDataLab/disinfox`

4 DISINFOX: DISINFOrmation threat eXchange platform

4.1 Mapping of disinformation incidents to CTI objects

For CTI, accurately modeling threats is essential for formal and homogeneous analysis, sharing, and response.

4.1.1 Comparison of disinformation frameworks

Similar to how cyberattacks are deconstructed using cyber kill chains, disinformation attacks require structured modeling to capture their phases and strategies. This enables a common understanding and translation into standardized formats to increase interoperability and automation in combating information threats jointly in both countries and organizations.

A recent article [1] reviews the pros and cons of disinformation-based schemes and taxonomies, having different perspectives and application. Table 4.1 presents a summary of the frameworks considered for modeling disinformation incidents. This section provides a comparative analysis of five prominent frameworks: DISARM, SCOTCH, BEND, ABCDE, and ALERT. These frameworks vary in their focus, design, and applicability, offering diverse approaches to understanding and mitigating disinformation campaigns

Framework description

The Disinformation Analysis and Risk Management (DISARM) framework [35], proposed by the DISARM Foundation, is a comprehensive model inspired by cybersecurity practices. It employs the MITRE ATT&CK model and Cyber Kill Chain analogy, which are widely recognized in the cybersecurity domain. DISARM outlines a four-stage matrix (Plan, Prepare, Execute, and Assess) with specific TTPs, which offers a systematic and structured approach to modeling disinformation disinformation. Additionally, it provides a STIX2 mapping to codify disinformation-related insights effectively with a standardized language for sharing threat intelligence.

SCOTCH [36], developed by the Atlantic Council, is a high-level framework for understanding disinformation campaigns, particularly focusing on rapidly assessing influence operations by looking to a more abstract layer and analyzing the source,

Features	★ DISARM	SCOTCH	BEND	ABCDE	ALERT
Proposed by	DISARM Foundation	Atlantic Council	Carnegie Mellon and US Army	Carnegie Endowment for International Peace	QUS Business School, University of Melbourne and IDSA
Disinformation classification	X	X	X	X	X
Use case examples	X	X	X	X	X
Actors analysis	-	X	X	X	X
Countermeasures	X	-	X	X	X
Quantitative analysis	-	-	X	-	-
Supported by	EU, OTAN, ONU	-	-	EU	-
Codification capabilities	STIX2	-	TSV	-	-
Stages	Plan, Prepare, Execute, Assess	-	Framework workflow	-	-
Cyber analogy	MITRE ATT&CK and Cyber Kill Chain	-	-	-	-

Table 4.1: Summary of frameworks analysed (adapted from our recent publication [1])

channel, objective, target, composition and hook. It offers insights into disinformation classification and countermeasures, making it a valuable resource for practitioners who need actionable guidance.

BEND [37], created by Carnegie Mellon University in collaboration with the US Army, provides a structured framework for identifying and responding to disinformation threats. It is notable for including quantitative analysis, disinformation classification, and countermeasures, providing a more technical and measurable approach compared to others. However, it does not include interoperable codification capabilities, which may limit its compatibility with standardized intelligence-sharing formats.

The ABCDE [38] framework, proposed by the Carnegie Endowment for International Peace, takes a more conceptual approach, concentrating on actor analysis and qualitative assessments. While it provides useful insights into the motivations and behaviors of actors involved in disinformation campaigns, it lacks features such as incident stages and codification capabilities, making it less actionable in practice.

Finally, ALERT [39], developed by QUT Business School, the University of Melbourne, and IDSA, offers a broad framework for understanding disinformation campaigns. It presents a taxonomy based on actors, lever, effects and responses, aiming to help security practitioners and policymakers in analyzing disinformation attacks in information systems. However, ALERT is more conceptual than operational, making it better suited for high-level strategic analyses rather than tactical applications.

Comparative analysis

The ability to characterize and model disinformation incidents is the base property of all the frameworks. This capability is particularly useful for organizations aiming

to analyze the diversity of disinformation campaigns. However, all of them have its particularities.

The inclusion of real-world examples helps bridge the gap between theory and application. All the analyzed frameworks—DISARM, SCOTCH, BEND, ABCDE, and ALERT—provide use case examples, making them valuable for practitioners seeking to understand their practical implementation. However, DISARM and SCOTCH excel in demonstrating how their methodologies can be applied to real-world scenarios, offering detailed illustrations that enhance their utility.

Understanding the roles and motivations of actors is a key strength of several frameworks. SCOTCH, BEND, and ALERT emphasize actor analysis, providing tools for identifying and examining the key players involved in disinformation campaigns. However, DISARM does not explicitly offer actor-focused analysis, as it is more centered on technical and procedural aspects. ABCDE, while offering high-level qualitative insights, lacks the practical tools necessary for a detailed examination of actors.

Developing effective countermeasures is a critical aspect of disinformation frameworks. DISARM, BEND, ABCDE and ALERT stand out in this regard by explicitly including countermeasure planning within their models. DISARM, in particular, integrates a mapping between an used techniques and the countermeasures to tackle it, providing a direct and actionable approach. ABCDE and ALERT, also includes countermeasure considerations but they are limited to recommendations for very open scenarios, contrary to the directness offered by DISARM. Conversely, SCOTCH do not explicitly include countermeasures, limiting their operational relevance.

Quantitative analysis is a valuable feature for organizations seeking measurable insights into disinformation campaigns. BEND incorporates quantitative methodologies, enabling users to evaluate the impact and scale of campaigns. However, contrary to some perceptions, DISARM does not explicitly integrate quantitative analysis into its framework, focusing instead on TTPs and technical interoperability. This feature is also absent in SCOTCH, ABCDE, and ALERT, which rely more heavily on qualitative assessments.

Codification capabilities enhance interoperability with existing systems and standards. DISARM is the only framework to adopt STIX2, a widely used standard for threat intelligence sharing, ensuring seamless integration into cybersecurity workflows. BEND supports TSV formatting for use with ORA-PRO software, providing some degree of codification but lacking the standardization advantages of STIX2. SCOTCH, ABCDE, and ALERT do not offer codification features, limiting their ability to integrate into technical ecosystems.

The methodologies and processes defined by the frameworks vary significantly in their structure and detail. DISARM outlines a comprehensive four-stage methodology—Plan, Prepare, Execute, and Assess— with TTPs rooted in cybersecurity practices. BEND adopts a workflow-based approach that focuses on maneuvering narratives and social networks, while SCOTCH, ABCDE and ALERT remain high-level conceptual frameworks, offering general guidance rather than specific methodologies.

Cybersecurity analogies, such as MITRE ATT&CK and the Cyber Kill Chain, provide valuable context for addressing disinformation in technical settings. Among the analyzed frameworks, only DISARM incorporates these analogies, making it uniquely suited for organizations familiar with cybersecurity practices. The other frameworks do not draw on these analogies, adopting broader approaches that may lack the precision needed for technical integration.

After this comparison, DISARM emerges as the most comprehensive model, combining the use of TTPs with codification capabilities and a structured methodology. It is particularly well-suited for organizations with the resources and technical expertise to implement its stages effectively. SCOTCH and ALERT, while less technical, provide valuable tools for actor analysis and classification, making them useful for strategic and conceptual analyses. BEND stands out for its quantitative focus, offering measurable tools for analyzing disinformation threats and their impact. ABCDE, on the other hand, offers a high-level conceptual framework that is valuable for qualitative assessments but lacks actionable features for operational use.

Selection of DISARM as a reference framework

The DISARM framework [35] integrates the concept of TTPs to model the behaviors and actions in disinformation attacks. It merges tools like the MITRE ATT&CK matrix or the Cyber Kill Chain and adapts them to enable a rich description of disinformation incidents by proposing a large set of DISARM in a matrix, detailed in Section 4.1.1. Additionally, the project provides an initial approach[1] to model attack techniques in STIX2, offering a direct mapping of disinformation attack techniques to the `AttackPattern` STIX object type. It also includes an official OpenCTI connector for integrating its TTPs matrix into the platform, enabling visualization and correlation of incidents. The aforementioned capabilities and applications demonstrate that the DISARM framework provides a clear cybersecurity perspective, making it an ideal choice for modeling disinformation incidents within the threat intelligence platform developed in this work.

Moreover, the utility of DISARM has been endorsed by several official EU bodies, including Foreign Information Manipulation and Interference Information Sharing and Analysis Centre (FIMI-ISAC) [40], the European External Action Service (EEAS) [12, 13], European Union Agency for Network and Information Security (ENISA) [9] or Hybrid CoE [41]. It is also employed in disinformation-related reports from Attribution Data Analysis Countermeasures Interoperability (ADAC.IO) [42], the ATHENEA project [43], the European Digital Media Observatory (EDMO) [44] or EU Disinfolab [45], further demonstrating the increasing adoption of this framework.

[1] `https://github.com/DISARMFoundation/DISARM-STIX2`

DISARM TTP Matrix

The core of the DISARM framework is its MITRE ATT&CK-like matrix, which can be visualized online[2]. The matrix permits the descomposition of any incident in phases with associated tactics and techniques. In the following, we formally define the main concepts of the matrix and apply them to a real-world influence operation within the Russia-Ukranian war for a clear comprehension. As this example will also showcase the rest of the paper, we provide some context.

> The *Ukraine Re-sold French Howitzers* (URFH) disinformation incident involved claims that Ukraine had sold CAESAR howitzers—supplied by France as military aid—on the black market. These allegations were propagated by Russian-affiliated media and Telegram channels in July 2022, supported by fabricated evidence and unverifiable reports. The narrative aimed to undermine trust in Western military support for Ukraine and to portray the aid as being misused. Despite lacking credible evidence, the disinformation gained traction within pro-Russian circles, showcasing the manipulation of information to influence public perception during the Russia-Ukraine war [46].

In this sense, to the eyes of the DISARM framework, the operation can be matched to the matrix and its elements which are described next. Table 4.2 illustrates the application of this matrix to the defined use case, supporting the description of the DISARM elements:

1. *Phase*: The most abstract grouping, representing sequential stages of an influence campaign by combining related tactics. There are four phases, including 1) `PLAN` (defining objectives and strategies), 2) `PREPARE` (creating and organizing assets), 3) `EXECUTE` (deploying and amplifying content), and 4) `ASSESS` (evaluating performance and persistence).

   > In the URFH incident, the first three phases of `PLAN`, `PREPARE` and `EXECUTE` can be inferred, but the last one of `ASSESS` is not intuitively interpretable by the analyst.

2. **Tactic**: Specific strategy that can be deployed in a particular Phase to achieve the campaign effects. The `PLAN` phase includes three possible tactics: `Plan Strategy`, `Plan Objectives`, and `Target Audience Analysis`, which outline the strategic groundwork. The `PREPARE` phase encompasses six tactics: `Develop Narratives`, `Develop Content`, `Establish Social Assets`, `Establish Legitimacy`, `Microtarget` and `Select Channels and Affordances`, focusing on operational readiness. The `EXECUTE` phase groups six tactics such as `Conduct Pump Priming`, `Deliver Content`, `Maximize`

[2] `https://disarmframework.herokuapp.com`

`Exposure`, `Drive Online Harms`, `Drive Offline Activity` and `Persist in the Information Environment`, ensuring active dissemination and impact. Lastly, the `ASSESS` phase includes only the tactic of `Assess Effectiveness`, emphasizing evaluation and refinement of the campaign's outcomes. As shown in Table 4.2, DISARM universally tags each tactic with a numerical unambiguous identifier.

> In the URFH use case, Russia would `Plan Objectives` during the `PLAN` phase, `Develop Content` and `Select Channels & Affordances` during the `PREPARE` phase, and `Conduct Pump Priming` and `Deliver Content` during the `EXECUTE` phase.

3. **Technique**: Specific fine-grained action deployed in the real world to complete a tactic. A tactic can have multiple techniques, one may be associated with multiple tactics, and some have sub-techniques for further detail. The DISARM framework covers a wide range of dozens of techniques to interpret any movement of any investigated operation, as mentioned next.

> In the URFH campaign, the actor begins in the `PLAN` phase with the `Plan Objectives` tactic, utilizing `Facilitate State Propaganda` to organize volunteers and disseminate messages favorable to their agenda. Moving to the `PREPARE` phase, the `Develop Content` tactic is employed through `Create Fake Research` and `Demand Insurmountable Proof`, aimed at discrediting opposing narratives and creating doubt about official information. Concurrently, the `Select Channels & Affordances` tactic leverages `Chat Apps`, `Social Networks`, and `Traditional Media` to ensure targeted and broad distribution of the fabricated content. Finally, in the `EXECUTE` phase, the actor applies the `Conduct Pump Priming` tactic using `Use Fake Experts` to lend false credibility to their claims. They further amplify the message through the `Deliver Content` tactic, employing `Cross-Posting`, `One-Way Direct Posting`, and `Attract Traditional Media` to maximize reach across various platforms and audiences.

As a conclusion, the DISARM framework provides a method to characterize and understand a complex influence operation universally.

4.1.2 STIX2 codification of DISARM-modeled disinformation incidents

For the solution to be CTI-compatible, the real-world disinformation incident modeled with DISARM must be translated into STIX2 objects, ensuring computational interoperability between connectors that use this threat intelligence data format. Since

Phase: PLAN		
Tactic	**Technique**	**Rationale**
TA02: Plan Objectives	T0002: Facilitate State Propaganda	*Coordinating volunteers to disseminate messages benefiting Russia.*
Phase: PREPARE		
Tactic	**Technique**	**Rationale**
TA06: Develop Content	T0019.001: Create fake research	*"Experts" claiming that Russia replicated the howitzers*
	T0040: Demand insurmountable proof	*Russian media reframing French' official versions*
TA07: Channels & Affordances	T0043: Chat apps	*Telegram use*
	T0104: Social Networks	*Twitter use*
	T0111: Traditional Media	*News in pro-Russian outlets*
Phase: EXECUTE		
Tactic	**Technique**	**Rationale**
TA08: Conduct Pump Priming	T0045: Use fake experts	*"Experts" claiming that Russia replicated the howitzers*
TA09: Deliver content	T0115.003: One-Way Direct Posting	*Telegram channels to disseminate*
	T0119: Cross-Posting	*Using news sites, Telegram, Twitter and other platforms*
	T0117: Attract Traditional Media	*News reaching mainstream media*

Table 4.2: DISARM phases, tactics and techniques detected in the "Ukraine Re-sold French Howitzers" disinformation campaign by Russian actors in the Russia-Ukraine war.

DISARM already provides its TTPs in STIX2 format, this eliminates the need to create new STIX2 objects for representing the TTPs. Additionally, the decision to use STIX2 aligns with the stack agreed upon between the EU and the United States for sharing disinformation threats [47].

STIX2 (Structured Threat Information eXpression) [48] is a standardized data model

designed to facilitate the exchange of threat intelligence information, traditionally related to cyberattacks. It organizes data into a bundle of interconnected objects, each representing predefined aspects of an incident, such as observed behaviors, threat actors, tools or techniques. This structured approach ensures consistent and agreed representation and enables seamless integration between systems.

However, although there are standardized ways of transforming cybersecurity knowledge to STIX2 objects, there are no guidelines for representing disinformation incidents yet. Therefore, we have conceptualized a way to abstract the nature of disinformation incidents to fit them in the already available STIX2 objects, providing an equivalency between a disinformation incident and a cybersecurity incident. This is also powerful, as it supports the integration and correlation in the same domain and common language of information and cyber threats, which is important for today's context.

Disinformation entities through STIX Domain Objects (SDOs)

STIX Domain Objects (SDOs)		
Property	**STIX2 object**	**Rationale**
Incident	`IntrusionSet`	Group of actions done by some entity
Actor	`ThreatActor`	Author of the incident
Technique	`AttackPattern`	DISARM technique launched
Country	`Location`	Geographic point of the targeted region

Table 4.3: Mapping between disinformation properties (nodes) and STIX2 object types

Firstly, the STIX Domain Objects (SDO) define specific concepts usually found in the CTI ecosystem [48]. As shown in Table 4.3, we map the details related to a disinformation incident to particular standardized STIX objects as follows:

- *Incident*: The core element of a disinformation incident, characterizing it through key properties such as the `name`, `description`, or `first seen date`. It is mapped to a `IntrusionSet` STIX object, traditionally used to represent a group of cybersecurity activities and resources with shared objectives, aligning well with the strategical nature of disinformation incidents. The `IntrusionSet` serves as the central object characterizing an incident, linking all related entities.

  Listing 4.1 presents a simplified STIX2 representation of the URFH *Incident*. Note that the `type` field specifies the SDO type of the element (`IntrusionSet`). The `name` field contains the title of the identified incident, while the `description` field holds the text content of the associated report. Additionally, the `first_-`

`seen` field represents the date when the incident was first identified. Fields, created,

- *Incident*: The core element of a disinformation incident, characterized by key properties such as `name`, `description`, and `first_seen`. It is mapped to an `IntrusionSet` STIX object, traditionally used to represent a group of cybersecurity activities and resources with shared objectives. This aligns well with the strategic and coordinated nature of disinformation incidents. The `IntrusionSet` serves as the central object characterizing the incident, linking all related entities.

  Listing 4.1 provides a simplified STIX2 representation of the URFH *Incident*. The fields `id`, `type`, `created`, `modified` and `spec_version` represent the STIX metadata that define and identify the object itself. The remaining fields, such as `name`, `description`, `labels`, and `first_seen`, form the payload of the object, containing the core details about the disinformation incident.

Listing 4.1: `IntrusionSet` SDO representation of the URFH *Incident*

```
{
  "id": "intrusion-set--76271730-...",
  "type": "intrusion-set",
  "created": "2024-12-25T23:35:11.86288Z",
  "modified": "2024-12-25T23:35:11.86288Z",
  "spec_version": "2.1",
  "name": "Ukraine re-sold French howitzers for profit",
  "description": "Claims that Ukraine had sold CAESAR howitzers...",
  "labels": [ "incident", "disinformation"],
}
```

- *Actor*: The entity, whether an organization, group, or individual, is believed to be responsible for orchestrating the *Incident*. It is mapped to a `ThreatActor` STIX object, which is actually designed to represent the malicious cyberattacker.

  Listing 4.2 contains the STIX2 representation of the *Actor* responsible for the URFH incident. In this representation, the key field is the `name`, which stores the name of the actor attributed in the source report: Russia. Additionally, the `threat_actor_types` field categorizes the actor as a `nation-state`, indicating its classification within the threat intelligence ecosystem.

Listing 4.2: `ThreatActor` SDO related to the URFH *Incident*

```
{
  "id": "threat-actor--7ebead2d-...",
  "type": "threat-actor",
  "created": "2024-12-25T23:27:53.696031Z",
  "modified": "2024-12-25T23:27:53.696031Z",
  "spec_version": "2.1",
  "name": "Russia",
  "labels": ["threat-actor"],
  "threat_actor_types": ["nation-state"]
}
```

- *Technique*: The specific DISARM technique used in the disinformation incident

that supported the *Actor* actions to achieve its goals. As Section 4.1 mentions, the DISARM foundation already translated this information to the `AttackPattern` STIX object for encapsulating the malicious techniques.

Listing 4.3 presents the STIX2-formatted representation associated with the *Facilitate State Propaganda* DISARM technique employed in the URFH *Incident*. Notice how the `name` and `description` fields correspond to the official name and description of the technique [3], respectively. The `kill_chain_phases` field specifies the overarching tactic in the DISARM matrix: `plan-objectives`, which is utilized by OpenCTI to display the techniques with color-coded visualizations.

In this case, note that the `created` and `modified` timestamps differ more than a year from those of the other listed SDO. This discrepancy arises because these objects were originally created by DISARM in its repository some time ago and the codification process in the platform uses these original SDO instead of creating new ones.

Listing 4.3: Simplified `AttackPattern` SDO related with the URFH *Incident*

```
{
  "id": "attack-pattern--70717452-...",
  "type": "attack-pattern",
  "created": "2023-09-14T20:38:04.999444Z",
  "modified": "2023-09-14T20:38:04.999444Z",
  "created_by_ref": "identity--f1a0f560-...",
  "name": "Facilitate State Propaganda",
  "description": "Organise citizens around pro-state messaging...",
  "external_references": [
    {
      "external_id": "T0002",
      "source_name": "mitre-attack",
      "url": "https://github.com/DISARMFoundation/DISARMframeworks/blob/main/generated_pag
      ↪  es/techniques/T0002.md"
    }
  ],
  "kill_chain_phases": [
    {
      "kill_chain_name": "mitre-attack",
      "phase_name": "plan-objectives"
    }
  ],
  ...
}
```

- *Country*: The world location to which the disinformation attack was targeted to. They are mapped to `Location` STIX objects as they represent a geographic point.

  Listing 4.4 presents the STIX2-formatted representation of one of the targeted countries identified in the URFH incident: France. In this `Location` SDO, two

[3] `https://github.com/DISARMFoundation/DISARMframeworks/blob/main/generated_pages/techniques/T0002.md`

fields are significant: the `country` field, which stores the value `France`, and the `name` field, which redundantly stores the same value for clarity and identification.

Listing 4.4: `Location` SDO related to the URFH *Incident*

```
{
  "id": "location--be5032fd-0b5c-5170-beb7-c7b499afa4bd",
  "created": "2024-12-25T23:27:52.703244Z",
  "modified": "2024-12-25T23:27:52.703244Z",
  "spec_version": "2.1",
  "country": "France",
  "name": "France",
  "type": "location"
}
```

In this context, a disinformation incident can be described using the aforementioned objects. It is important to note that STIX entities are independent of their relationships. This separation is leveraged to flexibly connect entities and expand knowledge, enabling adaptable and extensible modeling through multiple incidents.

Disinformation relations through STIX Relationship Objects (SROs)

STIX Relationship Objects (SROs)		
Relationship	**STIX2 object**	**Rationale**
Incident → *Technique*	`uses`	A *Technique* is used in an *Incident*
Incident → *Actor*	`attributed-to`	An *Incident* is attributed to some *Actor*
Incident → *Country*	`targets`	An *Incident* targeted to some *Country*

Table 4.4: Mapping between disinformation relationships (edges) and STIX2 object types

The STIX Relationship Objects (SRO) link the SDO and describe the generated CTI [48]. As shown in Table 4.4, we define three types of standard STIX relationships that relate two pieces of information (SDO) through their unique identification (`id`):

- *Incident* $\overset{uses}{\rightarrow}$ *Technique*: Represents the relationship between a *DISARM Technique* and the *Incident* in which it was employed. Typically, an *Incident* involves multiple *Techniques*, resulting in many such relationships.

  Listing 4.5 shows the STIX2 representation of the URFH disinformation technique. The `relationship_type` field is set to `uses`, aligning with our definition. The `source_ref` field references the `id` of the `IntrusionSet` representing the URFH *Incident*, while the `target_ref` field points to the `id` of the `AttackPattern` representing the particular DISARM *Technique*.

Listing 4.5: Simplified `uses` SRO related to the URFH *Incident*

```
{
  "id": "relationship--1dce08d4-3650-4f78-8d55-1a08055ffbf3",
  "relationship_type": "uses",
  "source_ref":"intrusion-set--76271730-6e05-51f0-bf4c-6a7c7b53d9b0",
  "target_ref":"attack-pattern--70717452-f7e3-4ce8-956f-39a4d34c5cfb",
  "type": "relationship",
  ...
}
```

- *Incident* $\overset{attributed\ to}{\rightarrow}$ *Actor*: Represents the relationship between an *Actor* and the *Incident* attributed to it.

  Listing 4.6 shows the STIX2 representation of the URFH attribution. The `relationship_type` field is set to `attributed-to`. The `source_ref` field references the `id` of the `IntrusionSet` representing the URFH *Incident*, and the `target_ref` field points to the `id` of the `ThreatActor` representing the URFH*Actor*.

Listing 4.6: Simplified `attributed-to` SRO related to the URFH *Incident*

```
{
  "id": "relationship--dd7da138-6850-4b6b-ae0f-8f20c2502882",
  "relationship_type": "attributed-to",
  "source_ref":"intrusion-set--76271730-6e05-51f0-bf4c-6a7c7b53d9b0",
  "target_ref":"threat-actor--7ebead2d-9a79-505f-8998-026100724eab",
  "type": "relationship",
  ...
}
```

- *Incident* $\overset{targets}{\rightarrow}$ *Country*: Represents the relationship between a *Country* and the *Incident* that targeted it.

  Listing 4.7 shows the STIX2 representation of the URFH target. The `relationship_type` field is set to `targets`, indicating the targeting relationship. The `source_ref` field refers to the `id` of the `IntrusionSet` representing the URFH *Incident*, and the `target_ref` field points to the `id` of `Location` object representing the *Country* targeted in URFH incident.

Listing 4.7: Simplified `targets` SRO related to the URFH *Incident*

```
{
  "id": "relationship--c476d1ee-1c33-4989-a51c-3dd4ef64dcf5",
  "relationship_type": "targets",
  "source_ref":"intrusion-set--76271730-6e05-51f0-bf4c-6a7c7b53d9b0",
  "target_ref":"location--be5032fd-0b5c-5170-beb7-c7b499afa4bd",
  "type": "relationship"
  ...
}
```

To sum up, these STIX2 SDO and SRO objects constitute standard representations of DISARM-modeled incidents. In order to be exchanged between CTI peers, they are encapsulated in a STIX2 Bundle, a container used to package and share multiple

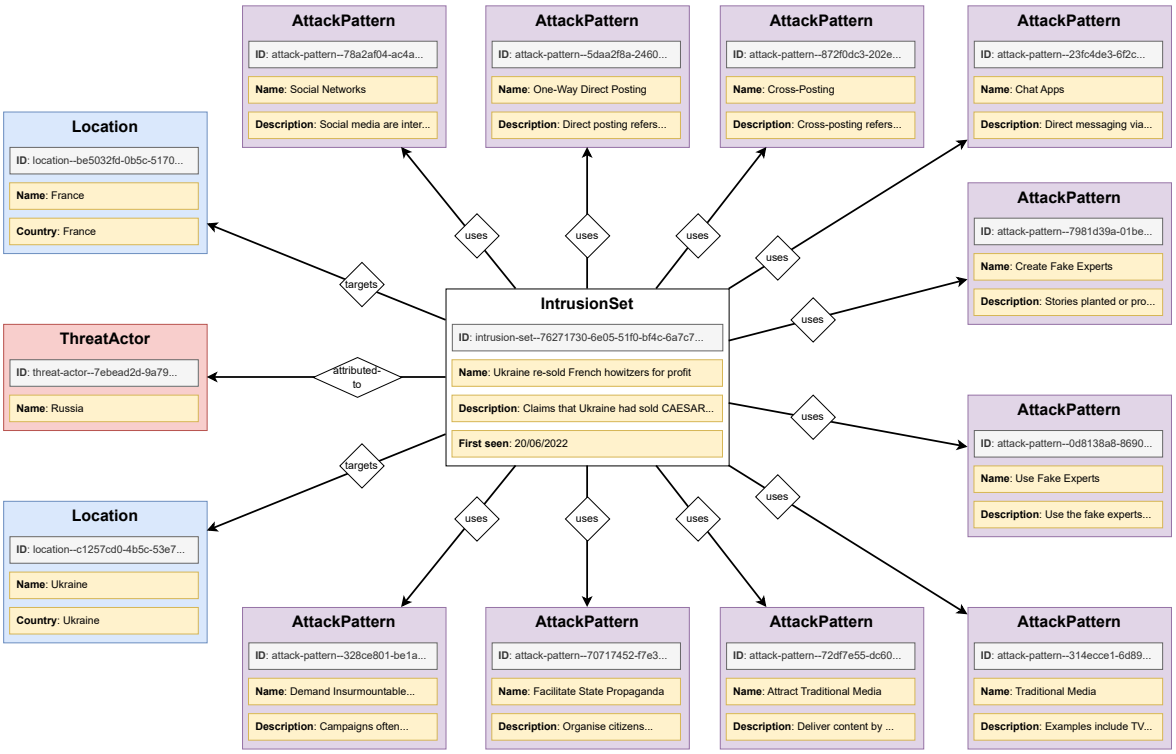

Figure 4.1: Graph representation of the STIX Bundle representing the modeled URFH disinformation incident

STIX objects [48]. Visually, the STIX2 Bundle can be seen as a graph in Figure 4.1. The corresponding simplified, machine-readable STIX2 Bundle object is shown in Listing 4.8, and the full version is available in the project repository[4].

[4] `https://github.com/CyberDataLab/disinfox/blob/main/backend/data/urfh_incident.json`

Listing 4.8: Simplified STIX2 bundle of uploaded disinformation incident


```
{
  "id": "bundle--3351770d-0656-4b3b-862f-6e81742669a3",
  "type": "bundle"
  "objects": [
  {
      "description": "Claims that Ukraine had sold CAESAR howitzers...",
      "first_seen": "2022-06-20T00:00:00Z",
      "id": "intrusion-set--76271730-6e05-51f0-bf4c-6a7c7b53d9b0",
      "name": "Ukraine re-sold French howitzers for profit",
      "type": "intrusion-set",
      ...
    },
    {
      "id": "threat-actor--7ebead2d-9a79-505f-8998-026100724eab",
      "name": "Russia",
      "type": "threat-actor",
      ...
    },
    {
      "country": "France",
      "id": "location--be5032fd-0b5c-5170-beb7-c7b499afa4bd",
      "name": "France",
      "type": "location"
      ...
    },
    {
      "created_by_ref": "identity--f1a0f560-2d9e-4c5d-bf47-7e96e805de82",
      "description": "Organise citizens around pro-state messaging. Coordinate paid or volunteer
      ↪  groups to push state propaganda.",
      "external_references": [
        {
          "external_id": "T0002",
          "source_name": "mitre-attack",
          "url": "https://github.com/DISARMFoundation/DISARMframeworks/blob/main/generated_pages/
          ↪  techniques/T0002.md"
        }
      ],
      "id": "attack-pattern--70717452-f7e3-4ce8-956f-39a4d34c5cfb",
      "name": "Facilitate State Propaganda",
      "type": "attack-pattern",
    },
    {
      "id": "relationship--1dce08d4-3650-4f78-8d55-1a08055ffbf3",
      "relationship_type": "uses",
      "source_ref": "intrusion-set--76271730-6e05-51f0-bf4c-6a7c7b53d9b0",
      "target_ref": "attack-pattern--70717452-f7e3-4ce8-956f-39a4d34c5cfb",
      "type": "relationship",
      ...
    },
    {
      "id": "relationship--c476d1ee-1c33-4989-a51c-3dd4ef64dcf5",
      "relationship_type": "targets",
      "source_ref": "intrusion-set--76271730-6e05-51f0-bf4c-6a7c7b53d9b0",
      "target_ref": "location--be5032fd-0b5c-5170-beb7-c7b499afa4bd",
      "type": "relationship"
      ...
    }
    ...
  ],
}
```

4.2 Design and implementation

The *DISINFOX* framework provides comprehensive, end-to-end support for modeling and sharing disinformation incidents. It encompasses the entire process, from uploading incidents in computational language to a centralized server, to the consumption of intelligence by client-side applications.

4.2.1 Design of the DISINFOX framework

The *DISINFOX* framework is inspired by well-established deployment models of traditional cybersecurity OTX schemes [49]. It is designed to handle real-world disinformation incidents originating from diverse sources, such as individual initiatives, news sites, or government reports. Figure 4.2 illustrates the technological stack, showcasing the process from uploading incidents to the platform to their integration within a CTI system. The *DISINFOX* framework features two main components:

- ***DISINFOX* platform**: The DISINFOX platform serves as the centralized repository for standardized, disinformation-based knowledge, providing a persistent source of intelligence and a user-friendly management. It ingests disinformation incidents using a two-phase pipeline:
  1. *DISARM Modeling*: Applied to represent the techniques used in each incident, as described in Section 4.1. Out of the DISARM framework, complementary details such as actor names, affected countries and other contextual data is also ingested.
  2. *STIX2 Representation*: The extended model of the disinformation incident is transformed into STIX2 format, generating SDO and SRO and inserting them into the database, as detailed in Section 4.1.2. This transformation ensures a standarized and machine-readable representation of the incident.

- ***DISINFOX* clients**: They are responsible for consuming and operationalizing disinformation-related knowledge. This paper introduces a custom *DISINFOX* OpenCTI Connector integrated with the OpenCTI platform. The *DISINFOX* OpenCTI Connector retrieves STIX2-encoded incidents from the *DISINFOX* platform and imports them into OpenCTI, enabling visualization and correlation with other CTI objects. Nevertheless, the *DISINFOX* client could be other CTI consumers by implementing the corresponding HTTP API based on STIX2 and DISARM standards.

In the following section, we describe the architecture deploying the DISINFOX framework.

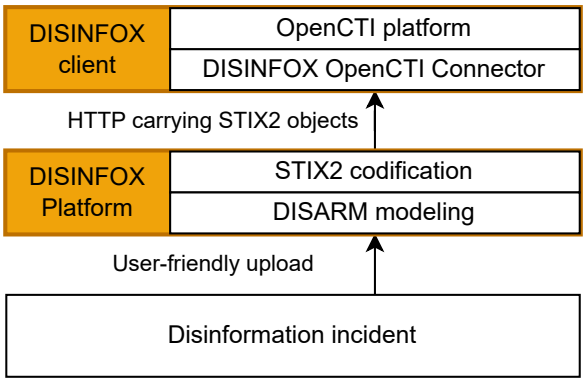

Figure 4.2: Technical stack of the DISINFOX framework

4.2.2 Implementation of the DISINFOX architecture

The framework has been designed through a service-oriented architecture to maximize interoperability while maintaining scalability and modularity. The publicly available implementation[5] relies on Docker containers for each service, ensuring ease of deployment and seamless communication (Figure 4.3).

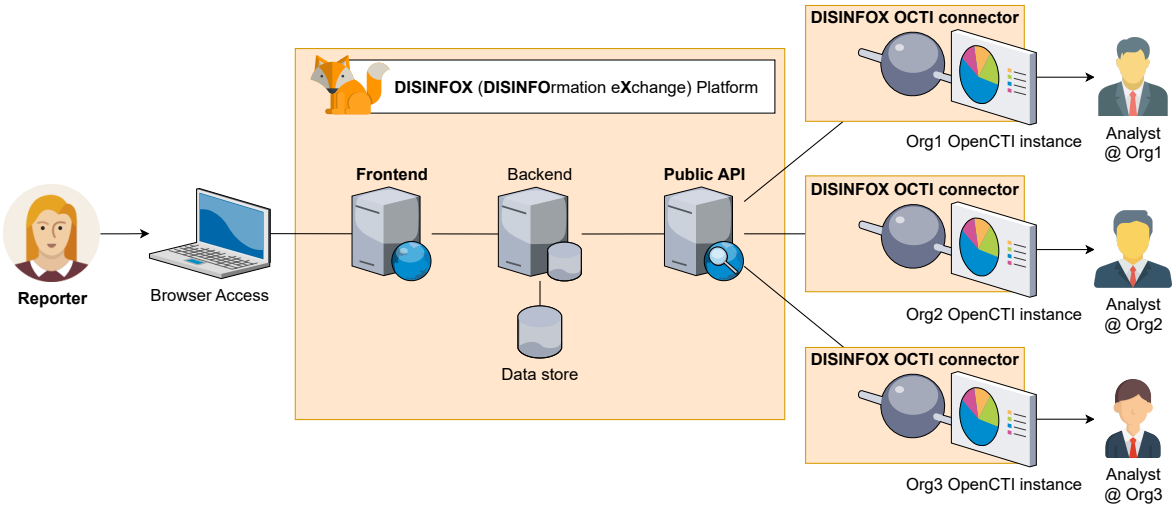

Figure 4.3: DISINFOX architecture

- **Frontend:** A web-based interface designed for non-technical users enables them to share and view disinformation incidents easily. Built with Python 3 and Flask, it uses Jinja2 templates to render responsive and visually unified HTML

[5] https://github.com/CyberDataLab/disinfox

pages using Bootstrap 5.3[6]. Also, Stixview[7] was integrated to generate interactive STIX2 graphs, providing enhanced visualization of incidents. The frontend interacts with the backend to upload user-submitted incidents, display platform data, and manage user accounts.

- **Backend REST API:** This component manages STIX2 objects and user data within the platform while interfacing securely with the data store. Developed with Python 3 and Flask, it provides a REST interface for handling STIX2 objects, enabling easy integration with future components and functionalities. Decoupling the backend from the frontend ensures the system remains agnostic to frontend technologies. The backend primarily sends STIX-formatted bundles to the frontend while ingesting and validating incidents submitted in the frontend. Using the STIX2 library, the backend transforms submitted data into well-formatted STIX2 objects and inserts them directly into the MongoDB collection. Additionally, this backend validates the public API requests and serves to it STIX2 objects for external CTI platforms.

- **Data Store:** A MongoDB database was selected for its native capability to store STIX2 objects. Various database types were evaluated, with SQL-based DBMSs discarded due to the extensive transformation required for STIX2 objects. Document-oriented DBMSs were preferred for their compatibility with JSON (the format used by STIX), offering flexibility and simplicity in handling the data. While graph databases could meet the requirements, their complexity and steep learning curve rendered them less suitable. Among document-oriented DBMSs, MongoDB was chosen for being open-source, providing robust Python library support, offering an official Docker image, and ranking as the most popular document database[8].

  Although *DISINFOX* is designed to function without preloaded data, allowing incidents to be added dynamically, the open-source code provide a dataset of 118 incidents from a variety of sources. This dataset includes incidents from [28], the DISARM repository [9], and several new incidents introduced in this work.

- **Public REST API:** This API, also built with Flask and Python 3 exposes endpoints for programmatic access to *DISINFOX* 's incident repository managed by the backend, allowing CTI connectors and other software to retrieve data. Users must authenticate requests by including an API key, which is generated in the Profile section of the frontend interface.

[6] https://getbootstrap.com/docs/5.3/getting-started/introduction/
[7] https://github.com/traut/stixview
[8] https://db-engines.com/en/ranking/document+store
[9] https://github.com/DISARMFoundation/DISARMframeworks/blob/main/DISARM_MASTER_DATA/DISARM_DATA_MASTER.xlsx

- **DISINFOX OpenCTI Connector:** Our publicly available Python 3 connector[10] for the OpenCTI platform serves as a proof of concept for demonstrating *DISINFOX* 's interoperability. This connector retrieves new content from *DISINFOX* and integrates it seamlessly into OpenCTI. Thanks to using STIX2 natively, no extra steps for the ingestion to OpenCTI are needed.

  OpenCTI was chosen as the platform to build the connector and validate the interoperability of the platform due to several key factors. First, it is part of the technology stack for disinformation sharing agreed upon by the EU and the United States [47]. Second, OpenCTI demonstrates a commitment to adapt its platform to better support disinformation management [32]. Third, it is the most popular open-source platform capable of ingesting STIX2. Lastly, OpenCTI offers a comprehensive guide for building connectors and has strong Python library support through the *ctipy* library.

While DISINFOX relies on all these modules for full functionality, only the frontend and the public REST API directly interact with external users, serving as the primary entry points to the platform.

4.3 Lifecycle of disinformation incidents and validation

The following subsections detail how an disinformation incident is managed and shared within *DISINFOX* . To illustrate the process, the URFH use case related to the Ukraine war is referenced throughout. Figure 4.4 outlines the main steps in the lifecycle, from incident upload to ingestion by other CTI platforms. These steps were performed to generate 118 disinformation incidents from the ingestion of *DISINFOX* 's default dataset[11].

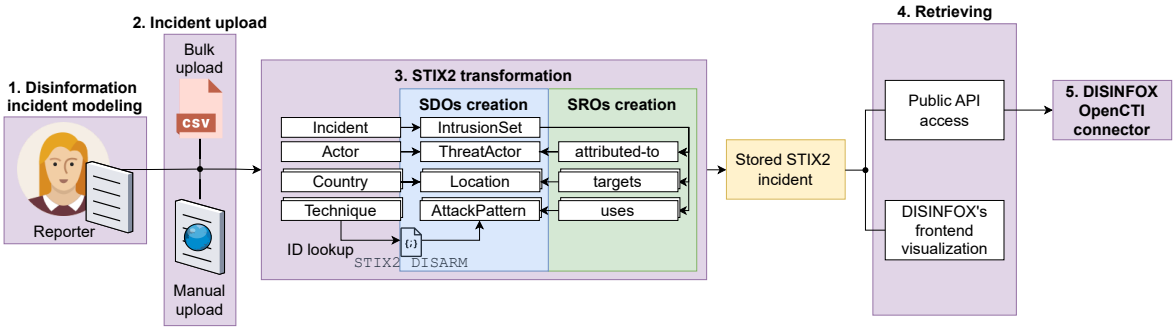

Figure 4.4: Incident lifecycle

[10] https://github.com/CyberDataLab/opencti-connector-disinfox
[11] https://github.com/CyberDataLab/disinfox/blob/main/backend/data/merged_Foulde_DSRM_additions.csv

4.3.1 Incident modeling

The Reporter uploading the incident to the platform must first model it using DISARM TTPs, as described in Section 4.1.

4.3.2 Incident upload

When a Reporter user accesses *DISINFOX* 's frontend, they can upload incidents using one of two methods:

- **Manual individual upload:** This is the simplest method for uploading a single identified incident. As illustrated in Figure 4.5, the user needs to fill out a form with the following fields: incident name, description, date, target countries, threat actors, and identified DISARM techniques.
- **Bulk upload:** This method is ideal for importing a large set of disinformation incidents. The user can upload either a CSV file[12] or a JSON file containing a STIX2 bundle with the incidents they wish to import. When using this method, the platform performs an intermediate transformation to format each individual incident, simplifying the creation of STIX2 objects.

The interactive form provides a user-friendly way for the Reporter to upload all the necessary information about a disinformation incident. The incident modelled in the previous section can be used as an example of how to fill out the form. Figure 4.5 illustrates the form fields filled with the required information for the incident. The title is entered as *Ukraine re-sold French howitzers for profit*, while the description contains a summary of the source report. The date field is filled with *June 20, 2022*, the date of the first evidence of disinformation, which corresponds to the first related post. The target countries, *Ukraine* and *France*, are selected as they were both targets of the false claims. The threat actor is identified as *Russia*, as noted in the source report. Finally, the DISARM techniques are listed according to those identified in Table 4.2.

Once incidents are uploaded using either method, the platform performs validation checks on the submitted data and transforms the incidents into individual STIX2 objects.

4.3.3 Automated STIX2 transformation

The process of creating STIX2 objects from incident data is guided by the mapping established in Section 4.1.2 and follows these steps:

1. The form data is used to create individual SDO, temporarily stored using Python's *stix2* library. First, an `IntrusionSet` SDO is created using the title, description, and date provided in the form, which populate the `name`, `description`,

[12]The CSV file must follow a specific template based on the one used in this working paper [28].

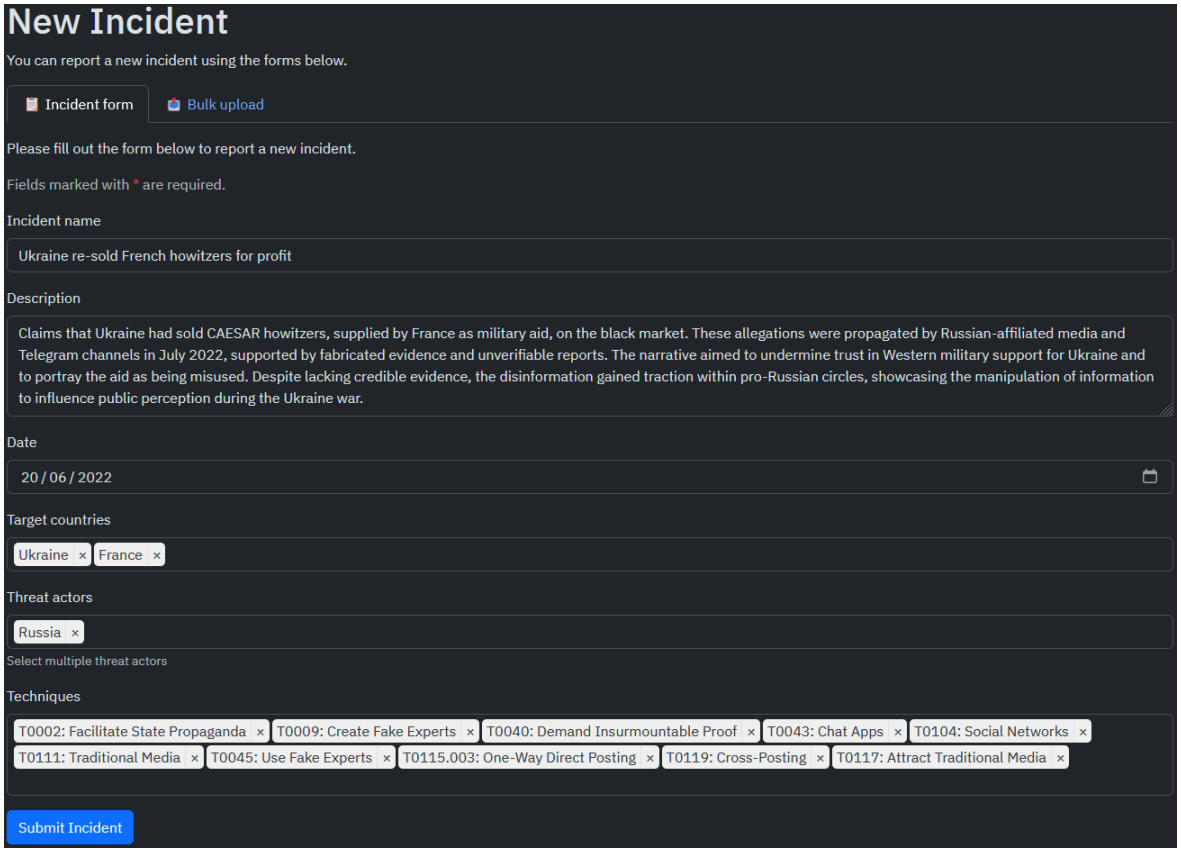

Figure 4.5: Manual individual upload form.

and `first_seen` properties of the object, respectively. Next, a `ThreatActor` SDO is generated using the threat actor names specified in the form. Finally, a `Location` SDO is created using the country names provided in the form.

Using the previously modeled incident example, the first three objects in Listing 4.8 demonstrate how these properties are populated and aligned with the uploaded data.

2. All DISARM techniques are represented as `AttackPattern` SDO, pre-built and stored in the `DISARM.json` file in STIX2 format. The DISARM techniques selected in the form are iterated through and matched against their corresponding entries in the JSON file. For each matching technique, the JSON object is converted into a Python `stix2` object and temporarily stored in a list.

   The fifth object in Listing 4.8 illustrates how a DISARM technique identified in the incident is represented in STIX2 format. Note that the `created` date in the `AttackPattern` SDO reflects the last update of `DISARM.json`, not the upload date of the incident in *DISINFOX* .

3. SRO are generated to link the previously created SDO. These SRO establish relationships between the `IntrusionSet` and the associated `ThreatActor`, `Location`, and `AttackPattern` SDO.

   The final object in Listing 4.8 illustrates a `targets` relationship SRO. Note how the `source_ref` and `target_ref` properties link the `IntrusionSet` and `Location` SDO, respectively.

4. All generated SDO and SRO are inserted into the platform's database.

The disinformation threat landscape is constructed from the STIX2 objects stored in the database, forming a structured and interoperable dataset for further analysis and sharing.

4.3.4 Retrieving stored incidents

Once the platform has ingested incidents, they can be retrieved in several ways.

- **For non-technical and casual users**, the most effective way of checking incidents is by looking at the frontend's listing. This listing shows the name, short description and date of the stored incidents. Once the user has found an interesting incident, he can view its details to get more information from it (Figure 4.6). All incident's information is shown graphically and intuitively: name, full description, date, target countries with a map, actors, used techniques, a graph showing the STIX2 relationships of this incident and the raw STIX2 bundle that represents this incident. Additionally, users can generate a PDF or Word report with all the detailed information about the incident to export it to other media and can select the incident as a favorite, so it can be easily found in its Profile.

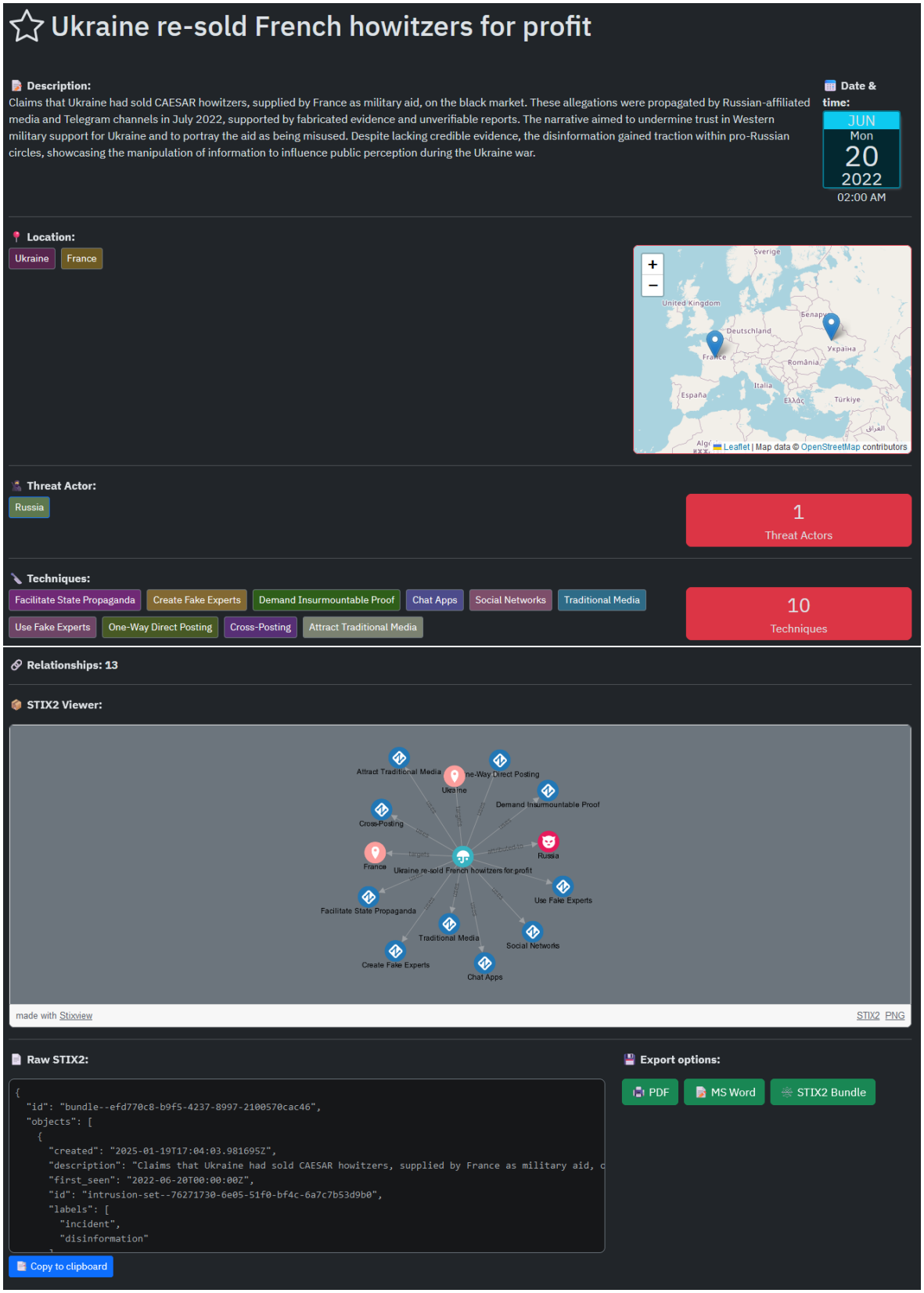

Figure 4.6: Visualization of a disinformation incident at the DISINFOX frontend web page

- **Technical users and specialized CTI developers** have the option to use the Public API to query the platform. Access to the API requires presenting an API key in the HTTP `Authorization` header, ensuring proper access monitoring and security. To obtain an API key, developers must register on the platform and navigate to the API Key section in their Profile Once the API key is obtained, the Public API can be queried, as demonstrated in the messages between the connector and the Public API shown in Figure 4.7.

  The request to the `/incidents` endpoint should include the `newer_than` parameter, which takes an ISO 8601 datetime string with microsecond precision. This parameter specifies the point in time from which the last edited STIX2 objects will be retrieved, making it particularly useful for reducing traffic and retrieval times by fetching only new or updated information from the platform. If all the objects needs to be retrieved, the epoch datetime can be used.

  This retrieval method allows developers to easily integrate incident data into their own applications in a RESTful manner. Extending this functionality to support the ingestion of new incidents through the API is a goal for future development.

This two methods are essential to provide an useful way of retrieving incidents for two different use cases.

4.3.5 Ingesting incidents from the DISINFOX OpenCTI connector

As it has been stated, the Public API eases the work of incident retrieval for applications that want to use *DISINFOX* 's incidents, especially to connect it to other CTI solutions.

To demonstrate this, the proof-of-concept *DISINFOX* connector for OpenCTI 6.4.2[13] was developed. Although the DISINFOX connector can be used standalone with an OpenCTI installation, it is recommended to first install the DISARM connector[14]. The DISARM connector not only inserts all AttackPattern SDO from DISARM into OpenCTI, but also provides the DISARM matrix and other additional objects that enhance the utility of the *DISINFOX* connector. This allows the DISINFOX incidents shared with OpenCTI to be analyzed using the matrix, complementing all the other visualization options available in OpenCTI

That ensure its utility a proof-of-concept OpenCTI connector has been built to prove the interoperability of *DISINFOX* 's incidents. As Figure 4.7 shows, this connector works in a very simple way thanks to using STIX2 natively:

1. The OpenCTI platform registers the connector and performs the first run of *DISINFOX* 's connector.

[13] `https://github.com/CyberDataLab/opencti-connector-disinfox`

[14] `https://github.com/OpenCTI-Platform/connectors/tree/master/external-import/disarm-framework`

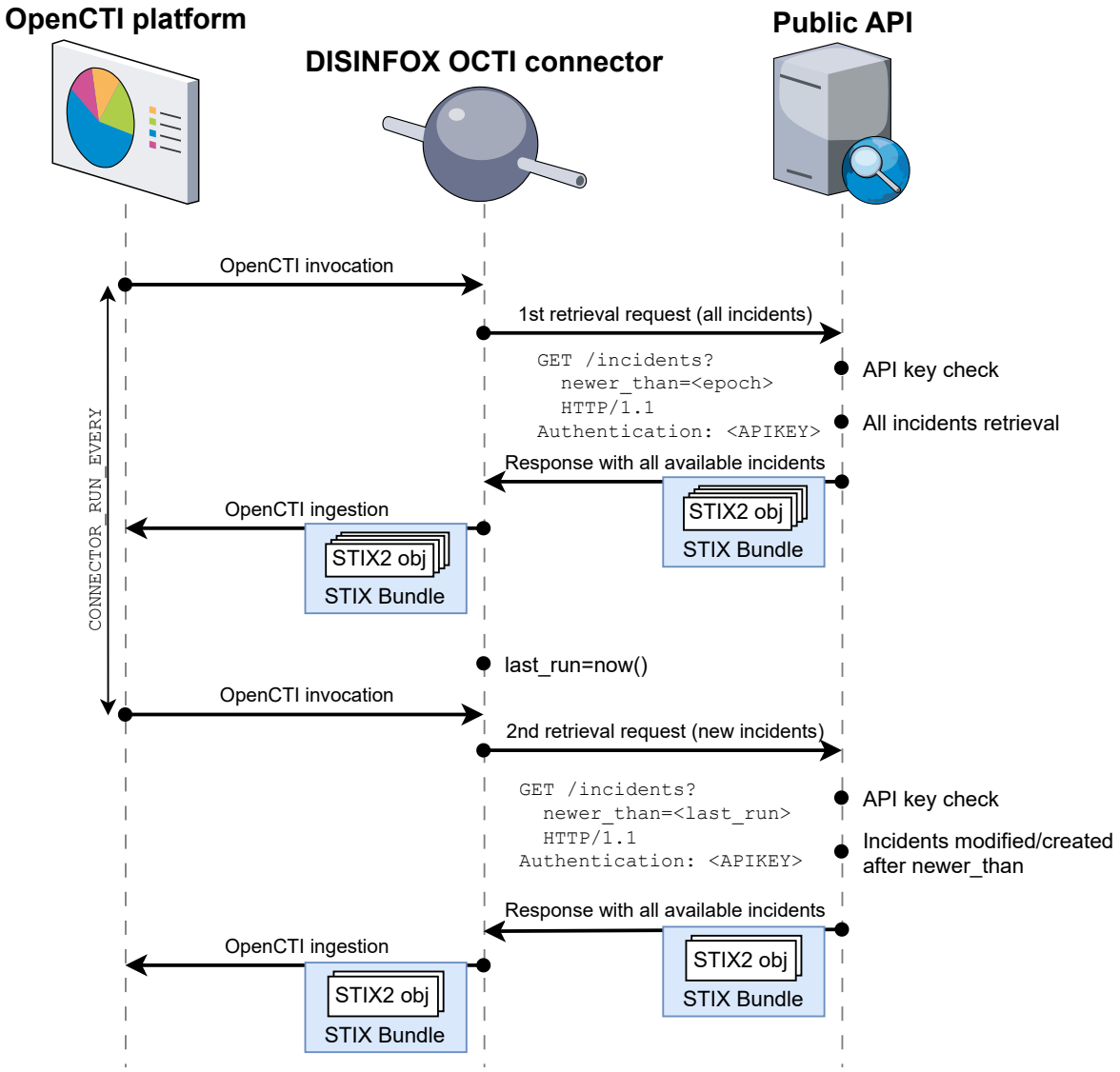

Figure 4.7: DISINFOX's proof-of-concept OpenCTI connector messages

2. *DISINFOX* connector sends a request to *DISINFOX* Public API with the `newer_than` parameter set with the epoch timestamp, as this is the first run and all the incidents need to be retrieved. It also includes an Authorization header with the `API key` that the used have been included in the `.env` file, previously obtained through its *DISINFOX* 's profile.

3. *DISINFOX* Public API check the API key in the request headers, if it is valid, it start retrieving all the incidents from the backend and send them back to the *DISINFOX* connector as a response. The body of this response will contain all the STIX2 objects representing all the incidents uploaded to the platform.

4. *DISINFOX* connector inserts the STIX2 objects from the API response to OpenCTI without any extra transformation.

5. The last operations are repeated just changing the `newer_than value`, that now will be set to the last time that the connector was set. The next call to the connector will be done depending on the time set in the `CONNECTOR_RUN_EVERY` parameter set in the installation of the connector to OpenCTI.

Now, all SDO and SRO are stored in OpenCTI. A listing of all the ingested disinformation incidents can be easily seen in the *Threats > Intrusion Set* section.

The presented use case can be used as an example to see the analysis that can be done in the OpenCTI platform. Apart from the *Overview* section that shows a summary of the properties (name, description, first seen date, etc.), the *Knowledge* tab of the Ukranian incident offers much more interesting data. The left picture of Figure 4.8 shows the *Diamond* graph that summarizes the relationships of the intrusion set in 4 dimensions: *Adversary*, where we find Russia as the threat actor; *Capabilities*, where attack patterns (DISARM techniques) such as *Use Fake Experts* or *One-Way Direct Posting* can be directly found; *Victimology*, where France and Ukraine are set as the targets of this intrusion set; and *Infraestructure*, which is unused. If the *VIEW ALL* button in the *Capabilities* frame or the *Attack patterns* button in the right bar is selected, OpenCTI displays the view in the right picture of Figure 4.8. This is the matrix view, which shows the used attack patterns in the matrix model that is selected, in this case, the DISARM matrix, which has been installed thanks to the DISARM connector. Notice how all the used attack patterns in the Ukranian incident are painted in red are placed under their corresponding tactic in the DISARM matrix.

These are just an example of the possibilities of using OpenCTI to manage disinformation incidents but other actions such as Cyber Kill Chain analysis or correlation with other incidents by taking into account its common DISARM techniques or target locations can be achieved. Overall, disinformation analysts can embed this connector to its workflow to monitor, correlate and asses disinformation incidents with a potentially shared view with other cybersecurity incidents, providing a rich picture of the current picture of the threat landscape.

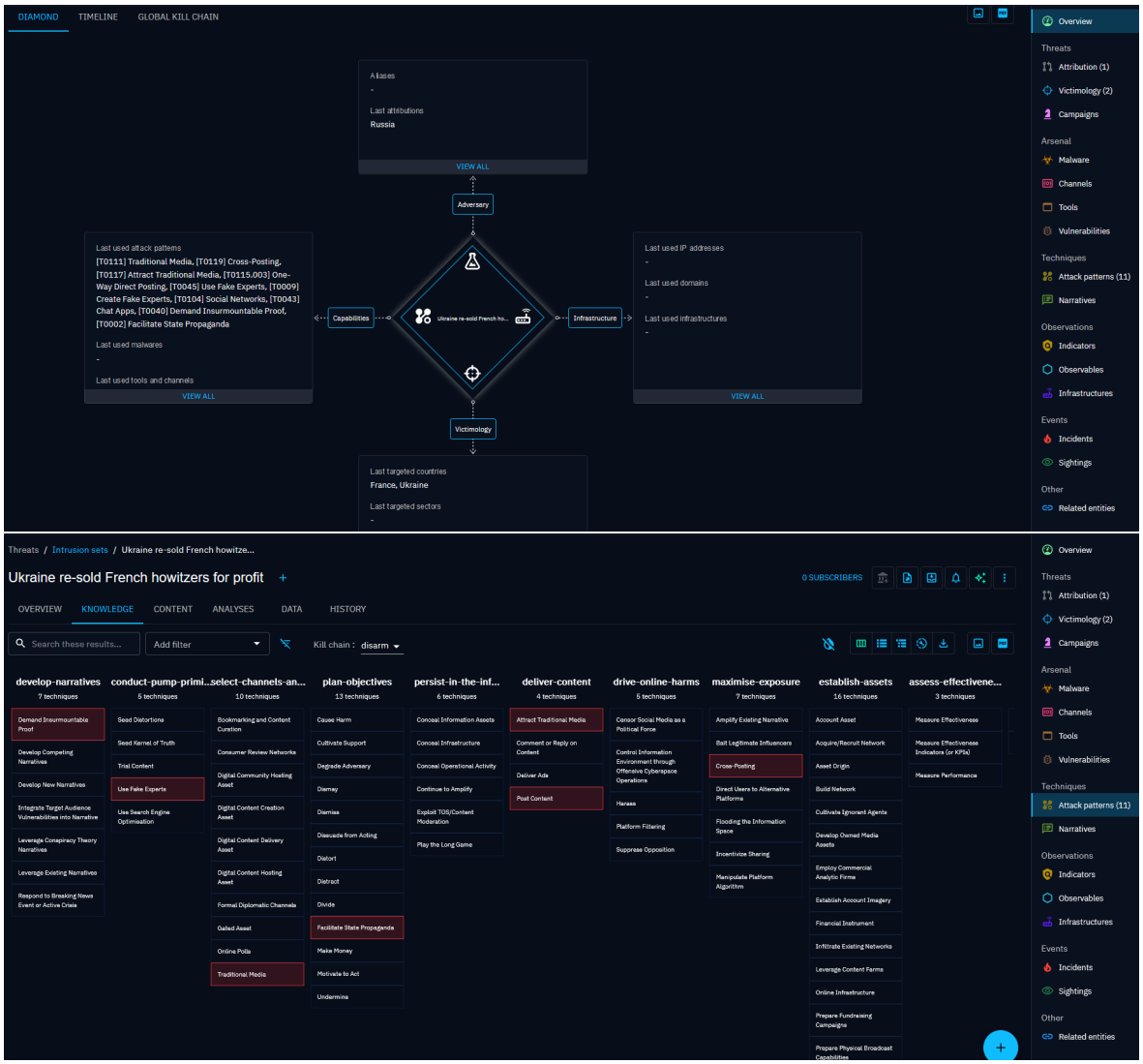

Figure 4.8: OpenCTI *Knowledge* tab in the page of the modeled intrusion set

5 Conclusion and Future Work

This Master Thesis introduced *DISINFOX* , a threat intelligence exchange platform tailored for disinformation incidents. By implementing CTI methods, *DISINFOX* provides a structured and interoperable approach to managing disinformation threats, a domain traditionally reliant on unstructured natural language.

The platform was successfully deployed and validated, bridging the gap between disinformation analysis and CTI standards. The use of the DISARM framework and a custom STIX2 mapping enabled the seamless representation of disinformation incidents, ensuring compatibility with established CTI platforms like OpenCTI. The integration demonstrated the practicality of modeling and managing disinformation using tools originally developed for cybersecurity.

The implementation delivered key results and contributions:

- A study of the most relevant disinformation modeling frameworks for the construction of the platform. The DISARM framework was selected after a comparative analysis, with its limitations addressed by complementing it with new STIX2 objects that were able to characterize and structure a more comprehensive picture of real disinformation incidents.

- The successful development and implementation of *DISINFOX* 's architecture within a modular, dockerized environment, consisting of a database, backend, frontend, and public API. This architecture supports real-time data exchange via a RESTful API, enabling both manual data input through a user-friendly interface and automated retrieval by external CTI solutions. The integration of the proposed framework, consisting of the *DISINFOX* platform and its clients, was validated by the successful storage, management, and visualization of a dataset of over 118 real disinformation incidents. Interoperability with mature CTI tools was also reached through a proof-of-concept connector for OpenCTI, allowing incidents to be analyzed, visualized, and correlated with this existing CTI platform.

- The lifecycle of stored disinformation incidents in *DISINFOX* was verified by analyzing the process from ingestion in the frontend to visualization in OpenCTI, validating a full adherence to the *EU-US Trade and Technology Council*'s technology stack for addressing FIMI [47].

All these contributions highlight the improve collaboration and automation in the fight against disinformation, while also demonstrating the potential of CTI method-

ologies in a broader context. Despite these achievements, several limitations need to be addressed:

- The dataset consisted of only 118 ingested incidents, which limits the depth and quality of the correlations and insights that can be generated from the platform.
- The manual application of DISARM TTP remains a bottleneck, increasing the time required to model and upload new incidents, which can hinder scalability.
- The current STIX2 mapping is minimal, and expanding it could enhance the richness of the incident data, providing a more detailed representation within DISINFOX and OpenCTI.
- The public API does not yet integrate with standards like TAXII, limiting the interoperability and automated sharing of incidents with other CTI platforms.

Future work will focus on addressing these limitations. Automation of the modeling process using Large Language Models (LLM) will simplify the identification of TTP and accelerate data ingestion. Expanding the dataset with diverse incidents will provide deeper insights into the disinformation threat landscape. Integration with emerging standards like the DAD-CDM initiative by OASIS will enhance the representation and interoperability of incident data. Finally, implementing TAXII for data transport will ensure seamless exchange with a broader range of CTI tools.

Bibliography

[1] G. C. L. de Molina, F. S. González, P. Nespoli, J. Pastor-Galindo, and J. A. Ruipérez-Valiente, "Analyzing frameworks to modeldisinformation attacks in online social networks," in *9th National Conference on Cybersecurity Research (JNIC 2024)*, pp. 92–99, 2024.

[2] W. Rector, "Deceivingly decisive: U.s. army military deception and counterintelligence," 2021.

[3] A. Mourad, A. Srour, H. Harmanani, C. Jenainati, and M. Arafeh, "Critical impact of social networks infodemic on defeating coronavirus covid-19 pandemic: Twitter-based study and research directions," *IEEE Transactions on Network and Service Management*, vol. 17, pp. 2145–2155, 2020.

[4] W. Hutchsinson, "Information warfare and deception," *Informing Science: The International Journal of an Emerging Transdiscipline*, vol. 9, pp. 213–223, 1 2006.

[5] S. Fotopoulos, "Traditional media versus new media: Between trust and use," *European View*, vol. 22, pp. 277–286, 2023.

[6] K. Baraniuk and P. Marszałek, "The potential of cyber threat intelligence analytical frameworks in research on information operations and influence operations," *Przegląd Bezpieczeństwa Wewnętrznego*, vol. 16, pp. 279–320, 12 2024.

[7] World Economic Forum, "Global risks report." `https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf`, 2024.

[8] D. de Seguridad Nacional (DSN), "TRABAJOS DEL FORO CONTRA LAS CAMPAÑAS DE DESINFORMACIÓN. INICIATIVAS 2024." `https://cpage.mpr.gob.es`, 2014.

[9] European Union Agency for Network and Information Security, "Enisa threat landscape 2024," 2014.

[10] B. van Niekerk, "The evolution of information warfare in ukraine: 2014 to 2022," *Journal of Information Warfare*, vol. 22, pp. 10–31, 02 2023.

[11] O. Kravchenko, V. Veklych, M. Krykhivskyi, and T. Madryha, "Cybersecurity in the face of information warfare and cyberattacks," *Multidisciplinary Science Journal*, vol. 6, p. 2024ss0219, 1 2024.

[12] European External Action Service's (EEAS) Stratcom, "1st EEAS report on foreign information manipulation and interference threats," 2 2023.

[13] European External Action Service's (EEAS) Stratcom, "2nd EEAS report on foreign information manipulation and interference threats," 1 2024.

[14] M. DeBolt, "Cyber threat intelligence capability maturity model version 1.0," 2024.

[15] MITRE Corporation, "MITRE ATT&CK," 2013.

[16] A. B. López, J. Pastor-Galindo, and J. A. Ruipérez-Valiente, "Frameworks, modeling and simulations of misinformation and disinformation: A systematic literature review," 2024.

[17] D. Preuveneers, W. Joosen, J. Bernal Bernabe, and A. Skarmeta, "Distributed security framework for reliable threat intelligence sharing," *Security and Communication Networks*, vol. 2020, no. 1, p. 8833765, 2020.

[18] C. Wagner, A. Dulaunoy, G. Wagener, and A. Iklody, "Misp: The design and implementation of a collaborative threat intelligence sharing platform," in *Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security*, WISCS '16, (New York, NY, USA), p. 49–56, Association for Computing Machinery, 2016.

[19] M. Motlhabi, P. Pantsi, B. Mangoale, R. Netshiya, and S. Chishiri, "Context-aware cyber threat intelligence exchange platform," *International Conference on Cyber Warfare and Security*, vol. 17, pp. 201–210, 03 2022.

[20] M. Mutemwa, J. Mtsweni, and N. Mkhonto, "Developing a cyber threat intelligence sharing platform for south african organisations," in *2017 Conference on Information Communication Technology and Society (ICTAS)*, pp. 1–6, March 2017.

[21] G. González-Granadillo, M. Faiella, I. Medeiros, R. Azevedo, and S. González-Zarzosa, "Etip: An enriched threat intelligence platform for improving osint correlation, analysis, visualization and sharing capabilities," *Journal of Information Security and Applications*, vol. 58, p. 102715, 2021.

[22] P. Koloveas, T. Chantzios, S. Alevizopoulou, S. Skiadopoulos , and C. Tryfonopoulos , "intime: A machine learning-based framework for gathering and leveraging web data to cyber-threat intelligence," *Electronics*, vol. 10, no. 7, 2021.

[23] P. Balasubramanian, S. Nazari, D. K. Kholgh, A. Mahmoodi, J. Seby, and P. Kostakos, "Tstem: A cognitive platform for collecting cyber threat intelligence in the wild," 2024.

[24] "EUvsDisinfo | Detecting, analysing, and raising awareness about disinformation - EUvsDisinfo — euvsdisinfo.eu." `https://euvsdisinfo.eu/`. [Accessed 20-01-2025].

[25] G. Harman, R. Tarrant, A. Tolbert, N. Ungerleider, and C. Wolf, "Disinfodex." `https://disinfodex.org`, 2020.

[26] "Media Manipulation Casebook — mediamanipulation.org." `https://mediamanipulation.org/`. [Accessed 20-01-2025].

[27] Digital Forensic Research Lab (DFRLab), "Interference 2024 — interference2024.org." `https://interference2024.org/`. [Accessed 20-01-2025].

[28] M. Fulde-Hardy, "Working paper presenting a dataset, a methodology, and a codebook to guide future applications of structured frameworks enabling threat assessment." 2024.

[29] DAD-CDM, "Home - DAD-CDM Open Project — dad-cdm.org." `https://dad-cdm.org/`. [Accessed 20-01-2025].

[30] "OpenCTI | Filigran — filigran.io." `https://filigran.io/solutions/opencti/`. [Accessed 20-01-2025].

[31] "connectors/external-import/disarm-framework at master · OpenCTI-Platform/connectors — github.com." `https://github.com/OpenCTI-Platform/connectors/tree/master/external-import/disarm-framework`. [Accessed 20-01-2025].

[32] S. Hassine, "How OpenCTI helps to fight disinformation and foreign interferences | Filigran Blog — filigran.io." `https://filigran.io/how-opencti-helps-to-fight-disinformation-and-foreign-interferences/`. [Accessed 20-01-2025].

[33] "Threat Intelligence Platform — eclecticiq.com." `https://www.eclecticiq.com/threat-intelligence-platform`. [Accessed 20-01-2025].

[34] "LevelBlue - Open Threat Exchange — otx.alienvault.com." `https://otx.alienvault.com/`. [Accessed 20-01-2025].

[35] S. Terp and P. Breuer, "Disarm: a framework for analysis of disinformation campaigns," in *2022 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, pp. 1–8, 2022.

[36] S. Blazek, "Scotch: A framework for rapidly assessing influence operations," *Atlantic Council*, 2021.

[37] J. T. Blane, *Social-Cyber Maneuvers for Analyzing Online Influence Operations*. PhD thesis, United States Military Academy, 2023.

[38] J. Pamment, *The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework*. Carnegie Endowment for International Peace., 2020.

[39] K. C. Desouza, A. Ahmad, H. Naseer, and M. Sharma, "Weaponizing information systems for political disruption: The actor, lever, effects, and response taxonomy (alert)," *Computers & Security*, vol. 88, 2020.

[40] Foreign Information Manipulation and Interference - Information Sharing and Analysis Centre (FIMI-ISAC), "Fimi-isac collective findings i: Elections." https://www.enisa.europa.eu/sites/default/files/publications/Foreign%20Information%20Manipulation%20and%20Interference%20%28FIMI%29%20and%20Cybersecurity%20-%20Threat%20Landscape.pdf, 10 2024.

[41] H. Newman, "Foreign information manipulation and interference defence standards: Test for rapid adoption of the common language and framework 'disarm'," tech. rep., The European Centre of Excellence for Countering Hybrid Threats (Hybrid CoE), 2022.

[42] V. Smith, S. Campbell, and A. Maunder, "A comprehensive review of disarm framework and its compatibility with related frameworks used to model foreign information manipulation and interference." https://adacio.eu/a-comprehensive-review-of-disarm, 2025.

[43] ATHENEA Project, "Policy brief conclusions and recommendations from the athena project on foreign information manipulation and interference." https://project-athena.eu/wp-content/uploads/2024/12/Policy-Brief-ATHENA.pdf, 2024.

[44] E. Panizio, "Disinformation narratives during the 2023 elections in europe." https://edmo.eu/publications/disinformation-narratives-during-the-2023-elections-in-europe/, 2024.

[45] N. Hénin, "Fimi: Towards a european redefinition of foreign interference." https://www.disinfo.eu/publications/fimi-towards-a-european-redefinition-of-foreign-interference/, 2023.

[46] R. Osadchuk and Digital Forensic Research Lab, "How russia promoted the claim that ukraine re-sold french howitzers for profit." https://medium.com/dfrlab/how-russia-promoted-the-claim-that-ukraine-re-sold-french-howitzers-for-profit-fd51f71a9362, July 2022. [Accessed 25-12-2024].

[47] European External Action Service (EEAS), "Ttc ministerial foreign information manipulation and interference in third countries." https:

//www.eeas.europa.eu/eeas/trade-and-technology-council-fourth-ministerial---annex-foreign-information-manipulation-and_en, 2023.

[48] OASIS, "Stix version 2.1." https://docs.oasis-open.org/cti/stix/v2.1/stix-v2.1.html, June 2021.

[49] W. Tounsi and H. Rais, "A survey on technical threat intelligence in the age of sophisticated cyber attacks," *Computers & Security*, vol. 72, pp. 212–233, 2018.

List of Acronyms and Abbreviations

ADAC.IO	Attribution Data Analysis Countermeasures Interoperability
CTI	Cyber Threat Intelligence
CVE	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Scoring System
DAD-CDM	Defending Against Deception Common Data Model
DISARM	Disinformation Analysis and Risk Management
EDMO	European Digital Media Observatory
EEAS	European External Action Service
ENISA	European Union Agency for Network and Information Security
EU	European Union
FIMI-ISAC	Foreign Information Manipulation and Interference Information Sharing and Analysis Centre
IoCs	Indicators of Compromise
LLM	Large Language Models
OSINT	Open Source Intelligence
OTX	Open Threat Exchange
SDO	STIX Domain Object
SRO	STIX Relationship Object
STIX2	Structured Threat Information eXpression
TAXII	Trusted Automated Exchange of Intelligence Information
TTPs	Tactics, Techniques and Procedures