Thesis Proposal

Towards a Dataspace for Cyber Threat Intelligence

Navid Rahimidanesh

A thesis presented for the degree of
Master of Science

Department of Computer Science
University of RWTH
Germany
June 9, 2024

Contents

1 Abstract

This is a proposal for a master's thesis in the area of cyber security and data management. Due to the complexity of the ever-changing threat landscape and the interconnected nature of the threats, organizations can benefit from collaborative cyber defense. They can achieve collaboration by sharing cyber security information. There are certain barriers identified in the litereture for this process, such as privacy, trust, and regulatory issues. In this research we are going to investigate the capabilities of a data sharing concept called dataspaces to address these barriers. This document consists of an introduction of the related concepts, a short description of state of the art, a preliminary requirement analysis, a conceptual approach and methods to implement and evaluate the contributions.

2 Introduction

2.1 Motivation

As soon new software get deployed, its vulnerabilities get exposed to the attackers. Due to digitalizations new software is constantly being deployed, therefore organizations face constant challenge of protecting their systems and data from cyberattacks. One thing that can help them, however, is the fact that these threats are interconnected. Same threat actor often uses same techniques to exploit same vulnerabilities to target different organizations. Therefore, different organizations can improve their collective defense by sharing information. They share about threat actors, their motivations, tactics, techniques, and procedures (TTPs), indicators of compromise (IOCs), the systems' vulnerabilities, incident response plans and mitigation strategies. The process of collecting, processing, and sharing this information is called Cyber threat intelligence (CTI) which is more beneficial when it is collaborative (CCTI). Thanks to CCTI, organizations can be one step ahead of the attacker and mitigate the threats before the breach happenning.

However, there are several barriers identified in the literature [1] for CCTI. By analyzing them, we claim that many of them will be resolved by using a platform that fulfills data sovereignty requirements. Data sovereignty is the right of the owner of data to control what happens to their data when it is shared with someone else. CCTI barriers that we claim are related to data sovereignty include for example adversarial usage, privacy law violation risks and safeguarding sensitive information.

A data management and integration scheme that can offer sovereign exchange is Dataspace. There are existing initiatives to standardize a dataspace. Gaia-X and International dataspaces (IDS) are two examples. These projects are getting more popularity, therefore it is important to investigate their potentials and limitations. As far as my knowledge, there is no work that has investigated the potentials of dataspaces in the context of threat information sharing.

2.2 Thesis Goal

Our goal is to encourage more organizations in Europe to adopt collabortive cyber threat intelligence. We do so by tackling the information sharing barriers. We expect that dataspaces can alleviate these barriers significantly. Therefore we investigate its suitablity and limitations. Our methodology consists of four steps: Firstly, we identify the concerns and barriers and how they hinder sharing in different scenarios. Secondly, we try to design a platform that facilites sharing in those scenarios. Thirdly, we implement a protoype based on the results of the last step and lastly we evaluate the effectiveness of the designed platform in solving the problem.

2.3 Outline

The rest of the thesis is structured as follows: Section 3 provides the necessary background information. Section 4 reviews the related work. Section 5 describes the use case and the requirements. Section 6 describes the conceptual approach. Section 7 describes the implementation. Section 8 describes how the evaluation should be. Section 9 describes the timeline and plan.

3 Background

In this section, we will review the relevant concepts and technologies that are necessary to understand the rest of the work.

3.1 Cybersecurity

Cybersecurity, the practice of protecting computer systems, networks, and data, encompasses crucial terms such as confidentiality, ensuring data privacy; integrity, maintaining data accuracy and trustworthiness; availability, assuring data and systems are accessible when needed; authentication, verifying user identities; authorization, granting appropriate access; firewall, filtering network traffic; malware, malicious software; phishing, deceptive attacks; and vulnerability, weaknesses in systems.

Relevant actors in the realm of cybersecurity include hackers who exploit vulnerabilities, security analysts responsible for defending systems, vendors who create security solutions, employees who can pose potential insider threats, government bodies serving as regulators and law enforcement, threat actors with malicious intent, and end users utilizing digital resources in the interconnected digital landscape.

3.2 Cyber Threat Intelligence

Cyber threat intelligence (CTI) is the process of collecting, analyzing, and disseminating information about potential or current cyber threats. CTI relies on collecting data from diverse sources, including security tools, threat feeds, honeypots, forums, social media platforms, and other relevant online and offline sources. The cyber security information produced during this process have three types:

Strategic CTI – Why? Strategic CTI expresses high level insights such as overall threat landscape, the motivations of threat actors, and the business or political impact of the threats. It mainly benefits executive management and other decision making departments by allowing data driven decision making to reduce the risks of cyber attacks.

Tactical CTI – How? Tactical CTI is about "how" the threats can cause incidents. Examples are the tactics, techniques, and procedures (TTPs) used by threat actors, vulnerabilities in the organization's security infrastructure, and the strategies that were used to mitigate the impact of the breach. Security teams can achieve more efficiancy by not repeating the work already done leading to more agile cyber incident response.

Technical CTI – What? Technical CTI concerns with the indicators of compromise, meaning concrete technological signs about the attacker or an attack, such as malware signatures or malicious IP addresses. Security teams and system administrators can feed these data to the firewalls and intrusion prevention systems (IPS).

Data Modeling In order to have an effective CTI, suitable data modelling techniques are required. It serves three purposes: (1) to provide a backbone for all relevant information, (2) to specify the data input format for further analysis, (3) to define the desired target for information gathering. [2]

3.3 Threat Intelligence Platforms - TIPs

CTI should be actionable, meaning it the intelligence should lead to cyber security decisions and enforcable. Should it take a lot of effort to process, it will waste security team's resources and provide no benefit. One way of achieving it is automation. Threat intelligence platforms (TIPs) are software tools that are created to address this issue. They can aggregate and correlate the information from different sources, and provide better visualization of the whole data. Furthermore, they can be integrated into other security tools such as firewalls and intrusion prevention systems (IPS) to reduce manual labor and skip the human delay to improve the defensive power.

3.4 Threat Information Sharing

To improve the effectiveness of CTI through collaboration, an approach is to share threat information produced in the CTI process. By sharing information, organizations improve their security posture by working as allies to fight threats together. Zibak et. al. [1] reviewed the literature to find the possible benefits and barriers of threat information sharing, categorized them, and surveyed practitioners of cyber security, to measure their attitude towards these barriers and benefits.

Benefits of Threat Information Sharing The are four categories of benefits identified in [1]. First, the operational benefits include reduce duplicate information handling and support breach detection and response. Second, organisational benefits such as improving overall security posture and situational awareness, combating skills gap, cross-checking different sources, and expanding professional networks. Third, economic benefits which are total cost savings, allowing governmental subsidies, reducing investment uncertainties. Lastly, policy related benefits such as reinforcing the connection with the government agencies.

NIS2 Directive Due to the benefits of threat information sharing, there are some regulations encouraging it. An example regulations that mandates infromation sharing is the directive on measures for a high common level of cybersecurity across the Union, NIS2, that provides legal measures to boost the overall level of cybersecurity in the EU. It ensures EU member states to have a national Computer Security Response Team (CSIRT) that cooperate with each other, and also a culture of information sharing between the public and private sectors in critical sectors. More specifically, organizations that are part of the critical sectors are required to share information about incidents happened to them with the national CSIRT.

Barriers of Threat Information Sharing On the other hand, there are some factors introducing challenge for threat information sharing. Likewise, it could be categorized in four parts [1], operational barriers include lack of standardisation and difficulty of expressing the information, verifying the accuracy and quality, ensuring timeliness, interoperability with automation tools, and protecting private data from being shared. Some barriers fall into the organisational category, such as risk of damaging reputation, forming the necessary trust relationship, market rivalries, and lack of trained staff. Third group of barriers relate to economic issues, free riding effect, effort and cost of the process, and losing customer's trust are examples of such. Last but not least, is policy and regulations such as privacy laws (e.g., GDPR) and other regulations that might differ across countries.

We are going to find out which of these barriers could be alleviated by enforcing "tighter control" of the shared data. Some interesting barriers regarding this is the risk of hackers gaining access to the shared data, i.e., adversarial usage. Another is the risk of violating privacy laws by oversharing, for example, sharing personal data of european citizens outside EU is limited by GDPR. Another is the challenge of safeguarding confidential information, for example, an organization should be cautious to not reveal private information about their infrastructure to their competing organizations. The "tighter control" that was mentioned is similar to the concept of data sovereignty.

3.5 Data Sovereignty

By the increase of the value of data in businesses and data becoming a commodity, protecting data using laws and regulations has become a necessity. Data sovereignty is concept that has arisen in this context. It refers to the right of the owner of the data to have control over their data. By default, if a party is processing data owned by another party, the processing party can technically do anything with the data. Data sovereignty tries to address this issue. One means of achieving data sovereignty is by using usage control. Usage control concerns with introducing and enforcing restrictions on what could (not) happen to the data. It is the generalized version of the traditional access control which only concerns with "who" rather than "how", "where", "why". The data owner defines the usage policies and the usage control mechanism enforces them [3]. A technology that can enable data sovereignty is dataspaces.

3.6 Dataspaces

The term dataspaces term was first coined by Franklin et al. [4] to describe a new abstraction in data management to solve the data integration problem that follows: An organization has interrelated data in diverse origins, encompassing databases, files with various formats, and web services. The task is to query or update the data. Franklin et al. proposed a DataSpace Suppport Platform (DSSP) that helps developers by providing a single query language based on a unified view of the data sources. This implies a pay-as-you-go approach where physically moving and transforming the data is done only by demand.

The same concept applies when several organizations want to integrate their data or exchange it with others. In this context, the term dataspace would refer to the platform consisted of data sources in different organizations to do data exchange defined by a set of standards and protocols to enable interoperability. [5]

Goals of Dataspaces Apart from data sharing and integration, dataspaces can fulfill other requirements. A crucial requirement, that makes dataspaces interesting, is the sovereignty of data. dataspaces can fulfill data sovereignty by keeping the data in the owner's side, and only sharing the metadata publicly. Another requirement is governance of the dataspace. In order to facilitate the cooperation of different participants, a set of policies, rules and protocols should exist. To define them, a governance body is commonly expected to exist [5]. dataspaces should be open, meaning anyone complying with the policies should be able to join, which encourages a fair and non-monopolistic market. This entails an easy access, which means, anyone should be able to connect with a limited effort. dataspaces are usually designed to be decentralized and federated, meaning, there is no entity having direct control over all data exchanges. Different participants can interact with each other directly. This emphasizes the role of interoperability. This is only possible when certain open standards are established. Consequently, dataspaces complying to the same standards can be embedded inside each other enabling cross-data-space exchange [5].

"Dataspaces are defined as: A federated, open infrastructure for sovereign data sharing, based on common policies, rules and standards." [5]

4 Related Work

In this section, we will review some existing solutions related to our problem.

4.1 Existing Threat Intelligence Platforms

By the rise in the amount of CTI available, the need for tools to process them has increased. It lead to emergence of many Threat Intelligence Platforms (TIP). TIPs can fetch CTI from different repositories, process and correlate information, and visualize the results. They can also be used to collaborate and share CTI with other organizations who use the same platform.

Proprietary TIPs There are paid services that provide curated CTI feeds and more complex dashboards. There are many proprietary TIPs available [6]: Anomali ThreatStream [1], ThreatConnect [2], ThreatQ [3], EclecticIQ Platform [4], OpenCTI [5], etc. Some organizations offer open TIPs that allow anyone to publish CTI. However, the source code is not availabe and the platform is managed by the organization. Examples include IBM X-Force Exchange [6] and AlienVault Open Threat Exchange (OTX) [7].

Open Formats and Protocols To create open and interoperable TIPs some standard formats and exchange protocols have evolved. STIX [8], VERIS [9], and the Incident Object Description Exchange Format (IODEF) [10] are the most prominent CTI formats. TAXII [11] is a standard protocol for exchanging CTI that supports both request-response and publish-subscribe model. MISP [12]

[1]https://www.anomali.com/products/threatstream
[2]https://threatconnect.com
[3]https://www.threatq.com/
[4]https://www.eclecticiq.com/
[5]https://filigran.io/solutions/products/opencti-threat-intelligence/
[6]https://exchange.xforce.ibmcloud.com/
[7]https://otx.alienvault.com/
[8]https://oasis-open.github.io/cti-documentation/
[9]https://verisframework.org
[10]https://www.ietf.org/rfc/rfc5070.txt
[11]https://oasis-open.github.io/cti-documentation/
[12]https://www.misp-project.org/

is an open source TIP that is widely used in the industry. It is due to its various features such as efficient IOC database, automatic correlation, flexible data model, different sharing models, and being able to export to and import from other CTI formats. Paice and McKeown [7] encourage the usage of MISP in the UK energy sector after testing different sharing models of MISP in a simulated environment. Pahleven et al. [8] extend the technological capacity of TAXII using Distributed Ledger Technologies (DLT) to enable data non-repudiation and a publish-subscribe middleware to enable real-time sharing. Our contribution can be seen as an extension to standard sharing platforms (e.g. MISP and TAXII) where we add data sovereignty and usage control to it.

4.2 Information Sharing Communities

There are existing communities for threat information sharing. Information Sharing and Analysis Centers (ISACs) are one. These are non-profit organizations that help organizations in a specific sector, usually a critical national infrastructure, e.g. electricity, water, gas, health care, finance, etc., to share CTI with each other.

EE-ISAC An example ISAC would be European Energy Information Sharing and Analysis Center (EE-ISAC) which has acquired over 30 members from utilities, academia, governmental and non-governmental organizations since its foundation in 2015. Members exchange cyber threat information through plenary meetings, working groups, and a dedicated platform (based on MISP). The information exchange is based on a trust achieved by confidentiality agreements and regular physical meetings with the same members. [9]

4.3 Crowd-Sourced CTI

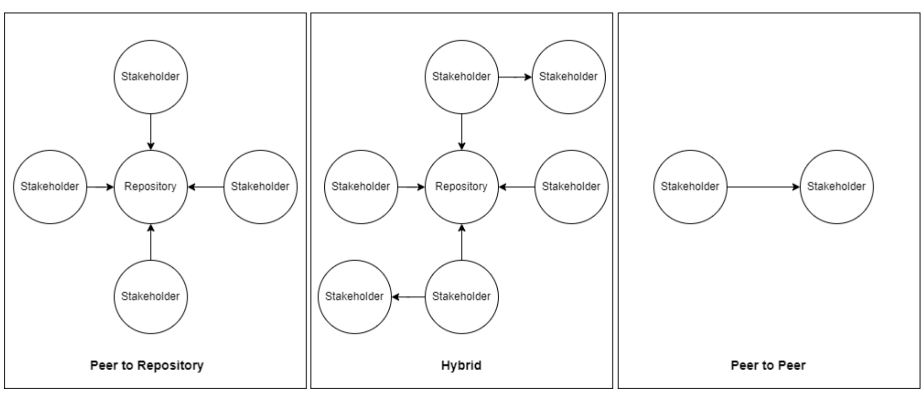

Figure 1: Information Sharing Models

In the context of CTI, information sharing can be either peer-to-peer, peer-to-hub (i.e. Repository), or a combination of the two (Figure 1). Traditional exchange between two organizations is an example of a peer-to-peer sharing. In the peer-to-hub, a hub or repository is used to collect and distribute data. If the repository is openly accessible to anyone, it is called open source CTI [10]. Jesus et al. [10] investigated the state of the art of the open source CTI and found the barriers that have prevented the formation of any widely used open source CTI platform. The barriers mentioned are 1) Legal and regulatory (e.g. GDPR or intellectual property) 2) Interoperability (e.g. different formats) 3) Usefulness and return 4) Market factors (e.g. losing reputation, free-riding) 5) Trust in peers and adversarial usage 6) Confidentiality risks.

After studying these barriers, as well as some technical gaps, they present a confidentiality and privacy analysis of sharing a large sample data set of CTI, to make the claim that it is possible to manage risks of sharing using simple techniques like sanitization. Finally, they propose a set of requirements and a reference architecture for an open source threat intelligence platform.

4.4 Current Dataspace Initiatives

There are several initiatives that are working to create a standard for Dataspaces. The most prominent ones are International Dataspaces (IDS) and Gaia-X. In this section, we will review these two initiatives and compare them.

IDSA The International Dataspaces (IDS) is an initiative with the goal of creating a standard for a distributed software architecture for data exchange with sovereignty. It was launched in 2015 as a Fraunhofer research project funded by the German Federal Ministry for Education and Research [11]. In 2016, the IDS Association (IDSA) was founded as a non-profit organization to continue the research. It resulted in definition of the IDS Reference Architecture Model (IDS RAM). The IDS RAM is the description of IDS components and their interactions without being technology specific [11]. IDS RAM allows anyone to implement the IDS compliant components using any technology. The IDSA also provides a reference implementation of different IDS components called IDS Testbed [13].

IDS RAM IDS RAM defines the following components [12]: Connector, Identity Provider, IDS Broker, Clearing House, IDS Apps, App Store, Vocabulary Provider [13]. Furthermore, it conceptualizes the following roles for the participants: Data Owner, Data Provider, Data Consumer, Data User, App Provider [13]. Also, it defines the standards and procedures to ensure data sovereignty. It uses usage control to enforce the usage policies defined by the data owner. It uses the following components to do so: Usage Control Policy Management Point (UC PMP), Usage Control Policy Decision Point (UC PDP), Usage Control Policy Enforcement Point (UC PEP) [11]. The policies

[13] https://internationaldataspaces.org/offers/reference-testbed/

are defined in a machine-readable format which is an extension of the Open Digital Rights Language (ODRL) [3]. These policies should be mapped to a specific policy language supported by the tool that enforces them. IDSA RAM mentions the following policy enforcement tools: MYDATA [14], LUCON [15], D° [16].

Gaia-X Gaia-X is an initiative, launched in 2019, that aims to foster creation of an infrastructure that allows for free and easy exchange of data and services between organizations and evade the vendor lock-in imposed by current proprietary cloud and service providers. To do so, regulations and technical specifications that are based on European values, applicable to any existing cloud and edge technology stack are going to be defined. The goal is to bring transparency, controllability, portability and interoperability across data and services. By facilitating data collection and sharing between organizations, a vibrant data ecosystem across Europe and beyond could evolve. Gaia-X Association deliverables include federation services, common policy rules and an architecture of standards. Federation services can be utilized by the ecosystem participants to achieve a global interoperability, compliance and effortless set up. This includes, "Identity and Trust", "Federated Catalog" and "Data Exchange services". [14]

Comparison of IDS and Gaia-X In comparison to IDSA, Gaia-X is less mature and still in the development phase whereas IDSA is used in the industry. It focuses on cloud infrastructures and businesses operating within EU, in contrast to IDSA which is more on the technicalities of the sovereign data exchange. Finally, Gaia-X can use IDSA in the data exchange layer [15].

Other Data Spaces Other data management researches based on dataspaces include "Trusted Integrated Knowledge Dataspace For Sensitive Healthcare Data Sharing" (TIKD) [16], which fulfills the following requirements: secure collaborative knowledge graph database of potentially personal data with fine grained access control and privacy-aware data interlinking. Another example is Real-time Linked Dataspace (RLD) [17] which is designed for the Smart Environments, supporting a pay-as-you-go data integration management system for real-time heterogeneous data sources that provides unified query interface based on linked data technologies. However, these data spaces lack the data sovereignty and usage control features that are necessary for our use case.

5 Use Case and Requirements

In this section, we will describe the use case and the requirements of the proposed platform. A platform for sharing threat information could be useful for

[14] https://www.dataspaces.fraunhofer.de/en/software/usage-control/mydata.html
[15] https://www.dataspaces.fraunhofer.de/en/software/usage-control/lucon.html
[16] https://www.dataspaces.fraunhofer.de/en/software/usage-control/d.html

any set of organizations that work together to achieve a common goal who fundamentally use IT systems in their operations. A good platform facilitate creation and operation of sharing communities. Example communities are sector specific ISACs such as ISACs for critical infrastructures, including, energy, water, finance, transportation, and healthcare. In order find concrete requirements, we need to narrow our focus to a specific use case. In this work, we will focus on the energy sector, and more specifically the cyber security of the smart grids.

5.1 Cyber Security in the Energy Sector

In order to improve the efficiency of energy distribution grids, efforts are done to make them smart, i.e. smart grids. Smart grids use different information technology (IT) components to collect and process data. However, these components are susceptible to cyber threats. They are an interesting target for attackers, specially advanced state-sponsored attackers, due to the level of damage that is achievable by a successful attack in the energy sector.

A threat actor for smart grids can be an advanced persistent threat (APT) supported by an enemy government, or a gang of experienced cyber criminals intended to disrupt the energy supply by attacking different actors in the supply chain. By doing so, they can reach their goal of causing a blackout, exfiltrating sensitive information, or gaining financial benefits (e.g. ransomware).

The important threats that smart grids face are listed by Wallis et al. [9]: Data injection attacks on state estimation [18], distributed denial of service (DDoS) and denial of service (DoS) attacks [19], targeted attacks, coordinated attacks, hybrid attacks, and advanced persistent threats [20]. Moreover, in recent years, ransomware campaigns have emerged as a significant risk to the sector [21].

5.2 Threat Information Sharing in Energy Sector

Due to the criticality of the cyber security in the energy sector, organizations in the energy supply chain try to improve their security posture. By using similar technologies, these organizations share the same vulnerabilities. Therefore, they attract the same attackers. These attackers use specific set of tactics, techniques, and procedures (TTP) to attack their victims over and over again. As a result, the victims can prepare themselves for these threats by knowing the TTPs that were used against other victims.

Therefore, energy sector organizations can benefit from sharing threat information with each other. That is where they can use our proposed platform. Therefore, the participants using our platform are different organizations active in the energy supply chain. Here, we assume that participants are the security team of the aforementioned organizations, or a managed security service if the organization does not have its own security team. That is because the security team is the entity responsible for handling the CTI, and we do not consider other types of security information (e.g. logs) in this work.

It is good to note that organizations in the energy sector are subject to the NIS2 directive, therefore they are obliged by the goverment to share cyber security information with other participants in the supply chain and the government.

5.3 Requirements

A comprehensive list of requirements requires more research of the existing processes to find the gaps. However, we can foresee that a useful platform should fulfil the requirements structured in the following 3 usage scenarios.

Scenario 1: Peer to Peer Sharing An organization (org 1) after having heard of the platform finds it useful. To join the platform, org 1 passes the necessary certification process. It trusts another organization (org 2) because they also passed the same certification process. They are able to connect to each other because of the interoperability of the different components. They agree to exchange information. To do so, first they negotiate on the scope of exchange and terms and conditions. This can include usage control policies. For example, they agree to only process data in servers located within the EU (to comply with GDPR). Afterwards, they proceed with the actual exchange.

Scenario 2: Share within community A utility provider (Acme) has found a new malware in their network. They share the related IOC to their trusted community. They set the usage policy to only allow the data to be read by the participants that have a certain minimum level on a certain trust metric. Furthermore, Acme does not want to share with competing companies. Therefore, it blacklists identified participants that compete in the same market.

Scenario 3: Paid Service and Rating A security organization specialized in selling CTI feeds, wants to use the platform to sell its services. It has a set of CTI feeds that it wants to share with its customers. It wants to charge its customers based on the number of IOCs they receive. Customers are also invited to rate the quality of the feeds they receive. This ratings are used by the customers to choose between different feeds.

6 Conceptual Approach

To support the use case scenarios mentioned, we are going to design and implement a sharing platform based on dataspaces for cyber threat intelligence. A high level diagram is depicted in figure 2. We used a color coding in the diagram: green components are the main contributions of this work. Yellow, are the components that we use or adjust to have a functional prototype. And white components are mentioned in the diagram to show the context, but we are not going to implement them.

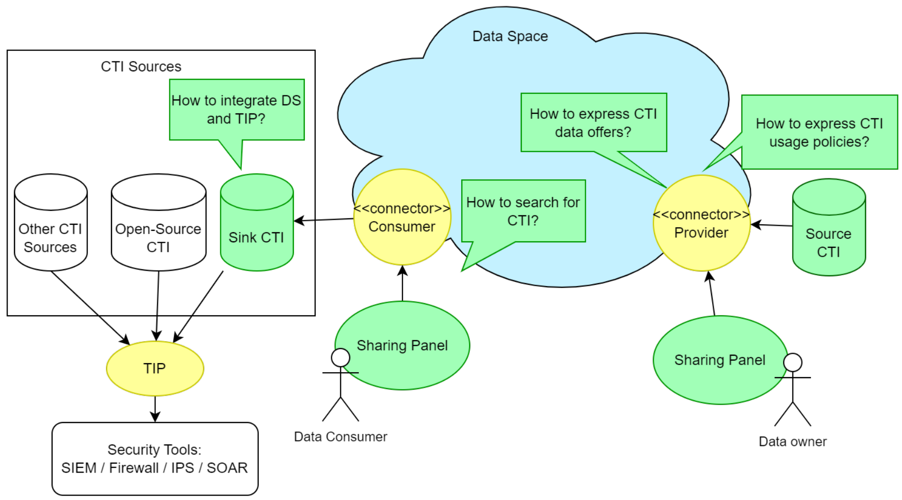

Figure 2: System Architecture Diagram

6.1 Dataspace Connector

We have to set up a functional dataspace to create the backbone of the platform to support sovereign data exchange. It has the ability to handle publishing data offers, secure data exchange, contract negotiation, and enforcing usage policies. Therefore, a dataspace connector component is necessary.

6.2 Data Modelling

We need to create data models to be used in the platform. The data models should cover the identified requirements. We expect that we need data models in three different areas:

Usage Policies: We will specify the formats and data models to express the usage policies based on our requirements, such as prevent access outside EU, or allow only certain fields to be read.

CTI: Here we focus on the threat intelligence data that is going to be exchanged. How it will be modeled, using which standard formats, or whether define a new abstract format to support many of existing ones.

Data Discovery and Negotiation: Here we should answer the following questions: How one should find published data offers? How to express the search query? After selecting the data, how the negotiation of the contract takes place?

6.3 Integration with TIP

The sharing platform is only useful when it is integrated to existing tool and platforms for threat intelligence. We should find how data can be extracted from existing threat databases and how the received data could be integrated in the TIP in the data consumer side.

6.4 Sharing Panel

To allow the user of the platform to trigger the necessary functions we need a user interface. The panel is a component that calls the corresponding endpoints in the dataspace connector. It should support the different usage scenarios of the platform, such as searching for offers, negotiation of contracts, specifying usage policies, and uploading the data to the dataspace.

7 Realization / Implementation

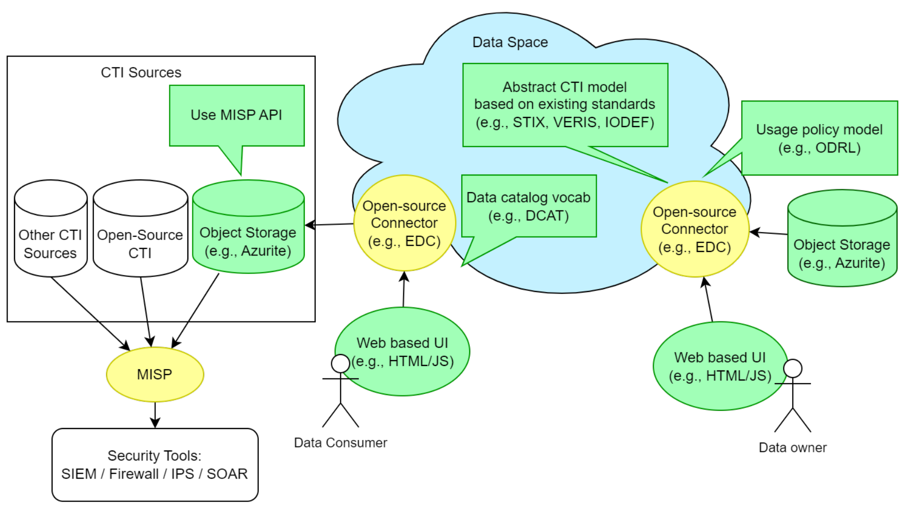

Figure 3: System Architecture Diagram with Technologies

To implement the components mentioned in the last section we are going to use different technologies and tools. An overview is illustrated in figure 3.

Dataspace Connector There are different dataspace specifications, but IDS RAM is one that most likely suit our use case. There are open-source implementations of IDS components such as IDS Connector. An open-source implementation is Eclipse Dataspace Connector (EDC) [17]. However, if during the analysis or implementation milestone, we realize that it does not suit our use case, we might switch to simpler solutions such as Solid [18].

Data Modelling For expressing the data models we try to stick to the standard formats as much as possible. The result of this task is the vocabularies and ontologies. For usage policies, there are different formats, the most prominent one is ODRL [19]. Also, for the CTI data model we analyze existing standards such as STIX, VERIS, IODEF. Finally, for the data catalogs a common format is DCAT [20].

Integration with TIP We will an open-source threat intelligence platform, MISP [21], and use its HTTP API to insert the data.

Sharing Panel The panel is a python script with different arguments for different functionalities. In a real-world scenario it is a graphical user interface which is out of scope of this thesis.

8 Evaluation

The evaluation is consisted of two approaches, analytical and experimental. In the analytical approach, we will go through the identified requirements and validate the designed architecture against them, to measure the extent to which the requirements are fulfilled. Moreover, will dive deeper in the security and privacy requirements and perform a comparative analysis against our baseline, MISP. We will also take a look at the threats in the energy sector and see how our platform could mitigate them.

The experimental approach is based on simulations using the created prototype. By executing different use case scenarios on the prototype we can validate qualitively the effectiveness of the platform in action and see how the requirements are satisfied. We can also measure quantitively some performance metrics such as throughput, delay and resource consumption. Similar to the analytical approach, we dive deep in the security and privacy requirements and simulate some attack vectors and see how the platform can help defend against them.

9 Project Plan

The work is planned to take 6 month (26 weeks), and consists of 5 Milestones:

[17] https://github.com/eclipse-edc/Connector
[18] https://solidproject.org
[19] https://www.w3.org/TR/odrl-model/
[20] https://www.w3.org/TR/vocab-dcat-3/
[21] https://www.misp-project.org/

Milestone 1 (1 Dec - 21 Dec: 3 weeks): Foundation

- Related work: Review the related works
- Define use case scenarios: Find concrete scenarios and requirements

Milestone 2 (21 Dec - 12 Jan: 3 weeks): Analysis and Modelling

- Information Modelling: Define required data models
- System Design: Specify the components and the data flows

Milestone 3 (12 Jan - 9 Mar: 7 weeks): Prototype Implementation

- Setup Dataspace: Implement the components for the dataspace
- Integration with TIP
- Sharing Panel

Milestone 4 (9 Mar - 7 Apr: 4 weeks): Evaluation

Milestone 5 (7 Apr - 29 May: 7 weeks): Writing and buffer

References

[1] A. Zibak and A. Simpson, "Cyber Threat Information Sharing: Perceived Benefits and Barriers," in *Proceedings of the 14th International Conference on Availability, Reliability and Security*, ARES '19, (New York, NY, USA), pp. 1–9, Association for Computing Machinery, Aug. 2019.

[2] M. Husák, L. Sadlek, S. Špaček, M. Laštovička, M. Javorník, and J. Komárková, "CRUSOE: A toolset for cyber situational awareness and decision support in incident handling," *Computers & Security*, vol. 115, p. 102609, Apr. 2022.

[3] A. Eitel, C. Jung, R. Brandstädter, A. Hosseinzadeh, S. Bader, C. Kühnle, P. Birnstill, G. Brost, Gall, F. Bruckner, N. Weißenberg, and B. Korth, "Usage Control in the International Data Spaces," tech. rep., Zenodo, Mar. 2021. Version Number: 3.0.

[4] M. Franklin, A. Halevy, and D. Maier, "From databases to dataspaces: a new abstraction for information management," *ACM SIGMOD Record*, vol. 34, pp. 27–33, Dec. 2005.

[5] Reiberg, A. a. Niebel, and Crispin, "What is a Data Space?," 2022.

[6] T. D. Wagner, K. Mahbub, E. Palomar, and A. E. Abdallah, "Cyber threat intelligence sharing: Survey and research directions," *Computers & Security*, vol. 87, p. 101589, Nov. 2019.

[7] A. Paice and S. McKeown, "Practical Cyber Threat Intelligence in the UK Energy Sector," Mar. 2023. Publisher: Springer.

[8] M. Pahlevan, A. Voulkidis, and T.-H. Velivassaki, "Secure exchange of cyber threat intelligence using TAXII and distributed ledger technologies - application for electrical power and energy system," in *Proceedings of the 16th International Conference on Availability, Reliability and Security*, ARES '21, (New York, NY, USA), pp. 1–8, Association for Computing Machinery, Aug. 2021.

[9] T. Wallis and R. Leszczyna, "EE-ISAC—Practical Cybersecurity Solution for the Energy Sector," *Energies*, vol. 15, p. 2170, Mar. 2022.

[10] V. Jesus, B. Bains, and V. Chang, "Sharing Is Caring: Hurdles and Prospects of Open, Crowd-Sourced Cyber Threat Intelligence," *IEEE Transactions on Engineering Management*, pp. 1–20, 2023.

[11] B. Otto, "The Evolution of Data Spaces," in *Designing Data Spaces : The Ecosystem Approach to Competitive Advantage* (B. Otto, M. ten Hompel, and S. Wrobel, eds.), pp. 3–15, Cham: Springer International Publishing, 2022.

[12] B. Otto, S. Steinbuss, A. Teuscher, and S. Lohmann, "IDS Reference Architecture Model," tech. rep., Zenodo, Apr. 2019. Version Number: Version 3.0.

[13] H. Pettenpohl, M. Spiekermann, and J. R. Both, "International Data Spaces in a Nutshell," in *Designing Data Spaces : The Ecosystem Approach to Competitive Advantage* (B. Otto, M. ten Hompel, and S. Wrobel, eds.), pp. 29–40, Cham: Springer International Publishing, 2022.

[14] H. Tardieu, "Role of Gaia-X in the European Data Space Ecosystem," in *Designing Data Spaces* (B. Otto, M. Ten Hompel, and S. Wrobel, eds.), pp. 41–59, Cham: Springer International Publishing, 2022.

[15] P. D. B. Otto, "GAIA-X and IDS," tech. rep., Zenodo, Jan. 2021. Version Number: 1.0.

[16] J. Hernandez, L. McKenna, and R. Brennan, "TIKD: A Trusted Integrated Knowledge Dataspace For Sensitive Healthcare Data Sharing," in *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, (Madrid, Spain), pp. 1855–1860, IEEE, July 2021.

[17] E. Curry, W. Derguech, S. Hasan, C. Kouroupetroglou, and U. ul Hassan, "A Real-time Linked Dataspace for the Internet of Things: Enabling "Pay-As-You-Go" Data Management in Smart Environments," *Future Generation Computer Systems*, vol. 90, pp. 405–422, Jan. 2019.

[18] R. Deng, P. Zhuang, and H. Liang, "False Data Injection Attacks Against State Estimation in Power Distribution Systems," *IEEE Transactions on Smart Grid*, vol. 10, pp. 2871–2881, May 2019.

[19] Q. Wang, W. Tai, Y. Tang, H. Zhu, M. Zhang, and D. Zhou, "Coordinated Defense of Distributed Denial of Service Attacks against the Multi-Area Load Frequency Control Services," *Energies*, vol. 12, p. 2493, Jan. 2019. Number: 13 Publisher: Multidisciplinary Digital Publishing Institute.

[20] R. Leszczyna, *Cybersecurity in the electricity sector: managing critical infrastructure*. Cham: Springer, 2019.

[21] M. Keshavarzi and H. R. Ghaffary, "I2CE3: A dedicated and separated attack chain for ransomware offenses as the most infamous cyber extortion," *Computer Science Review*, vol. 36, p. 100233, May 2020.