This is a preprint version of the paper entitled "AI4COLLAB: An AI-based Threat Information Sharing Platform", which was presented in the context of 2024 IEEE International Conference on Cyber Security and Resilience (CSR). Citation: C. Dalamagkas, D. Asimopoulos, P. Radoglou-Grammatikis, N. MAropoulos, T. Lagkas, V. Argyriou, G. Sargsyan and P. Sarigiannidis, "AI4COLLAB: An AI-based Threat Information Sharing Platform," 2024 IEEE International Conference on Cyber Security and Resilience (CSR), London, United Kingdom, 2024, pp. 783-788, doi: 10.1109/CSR61664.2024.10679429.

AI4COLLAB: An AI-based Threat Information Sharing Platform

Christos Dalamagkas*†, Dimitrios Asimopoulos‡§, Panagiotis Radoglou-Grammatikis¶, Nikolaos Maropoulos¶, Thomas Lagkas*, Vasileios Argyriou‖, Gohar Sargsya** and Panagiotis Sarigiannidis¶
*Department of Computer Science
Democritus University of Thrace, Kavala Campus, 65404, Kavala, GR
Email: tlagkas@cs.duth.gr
†EU Projects Coordination Department
Public Power Corporation S.A., Chalkokondili 22, 10432 Athens, GR
Email: c.dalamagkas@ppcgroup.com
‡Department of Information and Electronic Engineering
International Hellenic University, Sindos Campus 57400 Thessaloniki, GR
Email: dimiasim3@ihu.gr
§MetaMind Innovations
Kila, 50100 Kozani, Greece
¶Department of Electrical and Computer Engineering
University of Western Macedonia, Campus ZEP Kozani, 50100 Kozani, GR
Email: {pradoglou, psarigiannidis}@uowm.gr
‖Department of Networks and Digital Media
Kingston University London, Surrey KT1 2EE, UK
Email: vasileios.argyriou@kingston.ac.uk
**Celesta Advice
Email: g.sargsyan@gmail.com

***Abstract*—In the rapidly evolving field of cybersecurity, Cyber Threat Intelligence (CTI) sharing has become an essential practice to enhance awareness of emerging threats and enable infrastructure owners to defend against cyber incidents more efficiently. AI4COLLAB introduces a comprehensive CTI sharing platform designed to address the various challenges associated with CTI sharing, including data privacy, interoperability, and the speed of information dissemination. AI4COLLAB integrates two major CTI sharing platforms, MISP and OpenCTI, to broaden CTI coverage and leverages advanced AI techniques to automate the detection and anonymization of sensitive information. This ensures compliance with privacy regulations such as the General Data Protection Regulation while maintaining the utility of the shared data. By providing an in-depth analysis of existing CTI sharing solutions and presenting the innovative features of AI4COLLAB, this paper highlights the platform's potential to significantly improve the efficiency, security, and effectiveness of CTI sharing in the cybersecurity community.**

***Index Terms*—Artificial Intelligence, Cybersecurity, Threat Intelligence, Threat Sharing, Anonymisation**

I. Introduction

Cyber threats continuously evolve, harnessing emerging technologies such as Artificial Intelligence (AI) to become more sophisticated. Ransomware attacks no longer focus on critical infrastructure alone but have become intricate, targeting even supply chains. State actors continue to launch cyber espionage and disruptive attacks in pursuit of political, economic, or technological mileage. Another increased attack in the growing environment of the Internet of Things (IoT) is that it extends yet another vulnerability that can drive an attack across inter-connected devices. Phishing and spear-phishing have reached an individual level of social engineering, with the obvious intent of exploiting human psychology to gain unauthorized access to sensitive information [1].

Artificial intelligence and machine learning bring forth solutions to and create challenges for cybersecurity, given their increased adoption by attackers and defenders alike. It is, therefore, a flexible and specialized style of cybersecurity strategy that is instrumental in pushing back against the ever-changing face of these threats.

Cyber Threat Intelligence (CTI) involves the proactive collection and assessment of information pertaining to potential and real-time security threats and vulnerabilities. CTI entails the process of gathering data from security research, hacking forums, malware samples, and network monitoring. Such data will include Threat Intelligence indicators of compromise (IoC) such as IP addresses, domain names, hashes of the malicious files, patterns of attack, tactics, techniques, and procedures of the threat actors. Analyzing such information helps to spot patterns and trends, from which the organization and risks may be identified, which can help organizations to predict and prepare accordingly for cyber threats. Sharing CTI is a known best practice in cybersecurity, which raises awareness of new threats and helps owners of the infrastructure

to be better prepared for defense. Affected organizations share details about cybersecurity incidents, including IoCs, artifacts, linked observables, Tactics, Techniques, and Procedures (TTPs), and suggested actions for mitigation, to avoid such incidents happening to others. It is now recognized as a best practice, under European institutions' established direction and legal framework in the NIS2 directive, but also as an obligation for Operators of Essential Services (OES) and critical infrastructure operators [2].

In this paper AI4COLLAB is introduced, a CTI sharing platform that aims to address several challenges resulting from CTI sharing. The CTI sharing community uses multiple open platforms and protocols to store and share CTI incidents and information. Sufficient CTI coverage needs to be achieved through the adoption of different technologies, along with intermediate converters and adapters. A further layer of complexity is added, where CTI data needs to conform to privacy laws (e.g., General Data Protection Regulation (GDPR)) while still being useful. This imposes system architectures and anonymization techniques that protect the identity of the involved organizations and victims but do not destroy useful data. Another challenge is to let CTI sharing be fast, especially when considering OES. That is to say, preprocessing relevant CTI information and redacting sensitive CTI information should require a process that is fast and largely automated [3].

In response to these challenges, AI4COLLAB has been designed to enhance the effectiveness of Cyber Threat Intelligence (CTI) sharing by introducing several key innovations and improvements:

- **Automated Detection of Sensitive Information:** We utilize and compare various AI techniques to automatically detect sensitive information.
- **Anonymization of Sensitive Information:** We automatically anonymize the sensitive information, while ensuring that the CTI data still conforms to the MISP format.
- **Integration of multiple CTI sharing platforms:** Our proposed architecture showcases how two popular CTI platforms, Malware Information Sharing Platform (MISP) and OpenCTI, can be integrated and utilized to widen the range of CTI collection.

The remainder of this paper is organized as follows: Section II presents the related work. Section III delves into the proposed AI4COLLAB platform and Section IV discusses the methods we implemented for detecting sensitive information on the CTI data. Section V provides the evaluation analysis and the experiments done in this work and Section VI concludes the work.

II. Related Work

In this section, an overview of CTI sharing solutions is performed by discussing the relevant literature. The discussed solutions aim to improve the CTI sharing procedure in terms of data security, anonymity, performance, and efficiency. The possible data sources and data representations in CTI are thoroughly analysed in [4]. According to this reference, CTI categories include system logs, security and network events, externally sourced observables as well as open-source intelligence (OSINT). For CTI data representation, interoperable and platform-agnostic standards have been identified, including Structured Threat Information Expression (STIX), Cyber Observable eXpression (CybOX) and Common Vulnerability Reporting Framework (CVRF), as well as proprietary formats that work with specific platforms, including the well-known MISP. The CTI formats were examined for their efficiency with respect to different use cases, including email blocklist, spam filters, network intrusion detection and malware analysis. Other known formats include the Incident Object Description Exchange Format (IODEF), the Intrusion Detection Message Exchange Format (IDMEF), and the Open Threat Partner Exchange (OpenTPX). Regarding open-source software solutions in CTI sharing, [5] provides a relevant overview. MISP is a well-known and used CTI sharing solution, that allows organisations to share incidents, providing an open standard, user interface and REST API for creating and publishing MISP events. A similar solution to this is OpenTPX that also uses its own CTI sharing format having as common ability to map with STIX events, enabling interoperability with other platforms. Finally, OpenCTI allows organisations to manage and share knowledge in real-time, adopting the STIX standard and offering Application Programing Interfaces (APIs) and plugins for integration with various tools and platforms, including MISP, TheHive, Trusted Automated Exchange of Intelligence Information (TAXII) services, and AlienVault. Some of the open-source software solutions mentioned above have been improved by research works in terms of performance, data security, anonymity, performance and efficiency, by integrating various technologies on them. For example, some of the existing works employ the blockchain protocol to develop a decentralized and secure CTI sharing infrastructure. In more detail, the authors in [6] propose an Ethereum-based blockchain, the performance of which is evaluated against a Distributed Denial of Service (DDoS) attack. Through blockchain, a CTI report is constructed which takes 55 seconds to reach other nodes, upon the attack is detected. [7] also leverages blockchain to build a blockchain-based MISP, called LUUNU, that is built on top of the Rahasak blockchain. Credibility and transparency are ensured by storing the CTI data in the form of Model Card objects, whilst the anonymity of the participating organisations is ensured by using a self-sovereign identity-enabled mobile wallet, based on smart contracts. It is noteworthy that these works have been performed on top of the MISP platform. Furthermore, AI4COLLAB uses ML models in order to predict and anonymise sensitive data, included in the events, before publishing them to the CTI platforms. Similar works in the context of anonymization have been conducted making the use of anonymization techniques significant in the digital era. In this work, Nikoletos et al. [8] highlight the growing demand for data security while the number of online users increases, bringing forth the issues of safeguarding critical information from misuse. They focuse on the issues of data protection, which comprises of legal, ethical and technical aspects and which urges the use of automated tools in collecting and

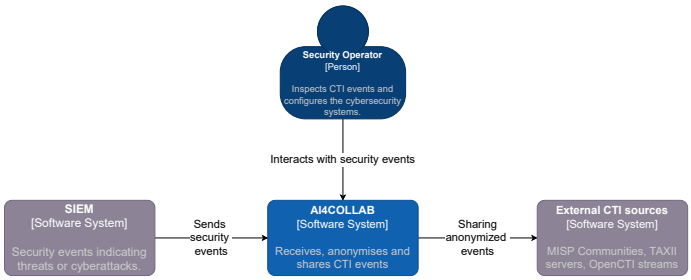

Fig. 1. The System Context diagram of AI4COLLAB

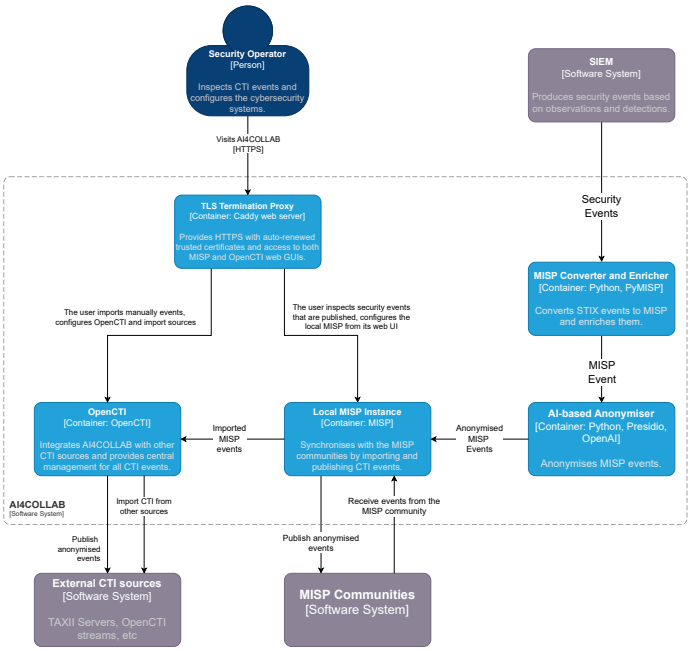

Fig. 2. The Container diagram of AI4COLLAB

anonymizing sensitive data. The work suggests a new process of fully automatic Natural Language Processing (NLP)-based system that will enable both high degree of efficiency and effectiveness, and will be suitable for various data sets across different domains. Moreover, in [9] the authors search the efficacy of text anonymization methods in the context of modern AI capabilities, particularly focusing on the challenge of balancing privacy protection with data utility. It questions the adequacy of current anonymization techniques to mitigate re-identification risks amidst the advancements in AI and big data analytics. Through an experiment with Generative Pre-trained Transformer (GPT) on anonymized texts of notable individuals, the study evaluates the potential for re-identification by AI, leading to a proposal for a novel approach that leverages Large Language Models to enhance text anonymity.

III. Architectural Design

The AI4COLLAB architecture has been defined by following the C4 model methodology [10]. The C4 model is a set of hierarchical abstractions for describing and visualising software architecture. The methodology originally defines four types of diagrams, namely a) System Context diagram, which describes the users and external entities interacting with the system, b) Container diagram, which describes the system as a set of independent services that interact with each other, c) Component diagram, which breaks down each container to describe components as function blocks that perform individual tasks, and d) Code diagram that delves into the implementation of each component to describe its source code operation with the help of Unified Modeling Language (UML) diagrams, entity relationship diagrams, etc.

Given that our custom code implementation is not complicated enough to need Code diagrams, we have considered the System Context, Container and Components diagrams for our work. Hence, in the following subsections, we describe the proposed solution through the aforementioned diagrams.

A. System Context Diagram

Fig. 1 depicts the System Context diagram of AI4COLLAB. This diagram focuses on the interactions of the system with users and external entities. In more detail, AI4COLLAB is used by the CTI analyst or the Security Operation Center operator (both of them will be refereed to as "Security Operator" in this paper), while it receives security events from external security systems and provides anonymized CTI data to CTI communities. The security operator undertakes system configuration, through user interfaces and/or via configuration files, as well as the inspection of CTI data through the available Graphical User Interfaces (GUIs). Incoming security events are provided by external security systems, e.g., Security Orchestration, Automation and Response (SOAR) platforms that provide STIX events describing an identified threat or a cyberattack. Finally, AI4COLLAB interacts with external CTI sources (e.g., TAXII servers, CTI streams, MISP instances), by sharing anonymized CTI data and reports.

B. Container Diagram

The Container diagram of AI4COLLAB is depicted in Fig. 2. This diagram analyzes the AI4COLLAB system into micro services that operate independently as separate system services, but strongly cooperate with each other. While C4 containers share many similarities with Docker container, those two terms are not identical and should not be confused [10]. As depicted in the figure, AI4COLLAB consists of the following containers:

- **MISP Converter and Enricher**: This service is the entry point of AI4COLLAB by asynchronously receiving security events for anonymization. This container undertakes two basic tasks: a) converts incoming events to the MISP format, b) enriches the MISP events with relevant IoCs and MISP taxonomies to represent as much information as possible. Then, the MISP event is provided to the Anonymiser.
- **AI-based Anonymiser**: This container processes MISP events and performs anonymization in two steps: a) detects sensitive information in the content of the event

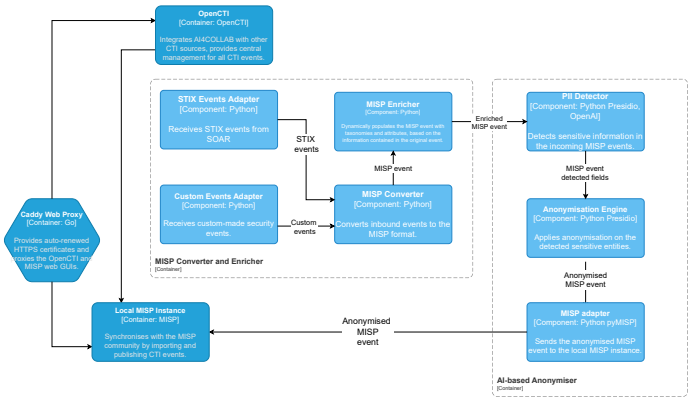

Fig. 3. The Component diagram of AI4COLLAB

by utilising AI algorithms, and b) applies data masking on the identified sensitive information. Finally, the container submits the security event to the local MISP instance.

- **Local MISP Instance**: This container reflects the local installation of the MISP platform, which allows the end user to participate in MISP communities and populate their threat inventory. After the anonymized MISP events are pushed to the local MISP instance, the security operator can access the web-based GUI of the instance, inspect the events, edit or correct them if necessary, and approve their dissemination within the connected MISP communities.
- **OpenCTI**: This container is a complete installation of the OpenCTI platform. It periodically imports the MISP events from the local MISP instance through a dedicated plugin, by generating corresponding STIX reports. Moreover, OpenCTI provides a web-based GUI for the inspection of the available CTI information. Finally, the imported events can be disseminated, while also additional CTI intelligence can be retrieved, through TAXII servers and CTI streams.
- **TLS Termination Proxy**: This container is based on the Caddy[1] server and acts as a reverse web proxy and TLS termination proxy, aiming to provide HTTPS access to the web interfaces of MISP and OpenCTI. If access through the Internet is needed, this container can also automatically generate trusted HTTPS certificates and automatically renew them through the Let's Encrypt[2] service.

C. Component Model

The containers can be further analysed into components, as depicted in Fig. 3. The component analysis reveals implementation details for the described containers as well as important function blocks that comprise each service. Since MISP and OpenCTI are already existing open-source tools, the component analysis focuses on two services that have been implemented by us: a) the MISP Converter & Enricher, b) the AI-based Anonymizer.

1) MISP Converter & Enricher components: The MISP Converter & Enricher container consists of the following components:

- **Event adapters**: Depending on the type of the incoming event, dedicated components are used to receive each type of event. AI4COLLAB processes STIX events as well as custom-made events coming from custom tools. Hence, there are two event adapters that are working in parallel. Each event adapter utilizes Apache Kafka [11] and subscribes to a distinct topic to receive new events asynchronously. As long as a new event arrives, the event adapter delivers the events to the MISP converter component.
- **MISP Converter**: This component converts the incoming events to the MISP format. In case of STIX events, the component utilizes the MISP-STIX[3] library. For custom events, though, custom code is used to match the event fields one-by-one. Finally, the component validates the MISP event, by constructing a pyMISP object that represents the MISP event. After successful validation, the event is passed to the MISP Enricher.
- **MISP Enricher**: This component enhances the MISP event with MISP taxonomies and attributes. The component uses the misp-modules[4] service of MISP in order to retrieve additional information for the IoCs contained in the MISP event.

2) AI-based Anonymizer components: The AI-based Anonymizer container consists of the following components:

- **PII Detector**: This component aims to detect sensitive information in the enriched MISP event. The detector can use multiple alternative detection methods, however, only one is activated at runtime. The employed detection methods are described in Section IV. Dedicated Python packages are used for each detection method, in particular the OpenAI[5] library is used to interact with ChatGPT, and the Presidio Analyzer[6] for applying the Presidio AI models. As a result, the component provides the original MISP event and the detected Private Identifiable Information (PII) entities.
- **Anonymisation Engine**: Given the identified sensitive content of the event, this component applies data masking on the sensitive information. As a result, this component outputs the anonymized MISP event, in which the sensitive data is replaced by wildcard characters (e.g., # or *).
- **MISP adapter**: This component utilises the pyMISP[7] library in order to interact with the local MISP instance and submit the anonymized MISP event. First, the component creates and saves the event on the local MISP

[1]https://caddyserver.com/
[2]https://letsencrypt.org/
[3]https://pypi.org/project/misp-stix/
[4]https://github.com/MISP/misp-modules
[5]https://pypi.org/project/openai/
[6]https://pypi.org/project/presidio-analyzer/
[7]https://github.com/MISP/PyMISP

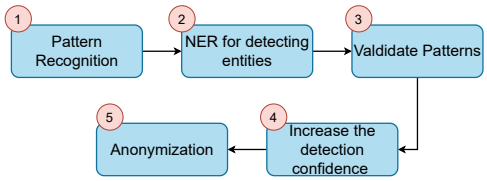

Fig. 4. The methodology of Presidio Model

instance and then it publishes the event to the connected MISP communities. If configured by the user during deployment, the MISP adapter may not publish the event, allowing the security operator to inspect the event first in the MISP GUI.

IV. PII Detection Methods

This section further analyzes the PII Detector component, in terms of PII detection methods that are leveraged to detect sensitive information. We have implemented three detection methods, by employing a) the Presidio model [12], b) GPT-2 and c) ChatGPT [13].

In the domain of data anonymization, Microsoft Presidio emerges as a robust, purpose-built tool that leverages advanced machine learning techniques to detect and anonymize sensitive information in text. Presidio operates by first identifying a wide range of personal data types, such as names, addresses, social security numbers, and credit card information, using a combination of predefined and customizable detectors. These detectors are grounded in pattern recognition, checksum validation, and contextual analysis, ensuring a high degree of accuracy in identifying sensitive data. Once identified, Presidio employs a series of anonymization strategies, including substitution, redaction, and generalization, to effectively obscure the identified information as shown in Fig. 4.

The data detection and anonymization process in the Presidio Model involves five steps. First, Regex utilizes pattern recognition techniques to identify specific sequences of characters within data. Next, Named Entity Recognition leverages natural language processing and machine learning to detect and classify named entities in text. Following this, a Checksum is used to validate data integrity by checking for patterns and ensuring the data has not been altered. Context Words are then analyzed to enhance detection accuracy by understanding the surrounding context. Finally, Anonymization employs various techniques to obscure personal or sensitive information, thereby protecting privacy.

Large language models (LLMs) such as the GPT have introduced a new approach to data privacy. These models have extensive knowledge of various linguistic environments, idiomatic expressions and semantic details since they were trained on a variety of datasets. LLMs especially those in GPT family are very good at parsing and changing texts without losing originality and meaning which makes them be often used for anonymization. The process involves using an annotated dataset containing sensitive information to fine-tune the GPT model so that it is capable of accurately identifying

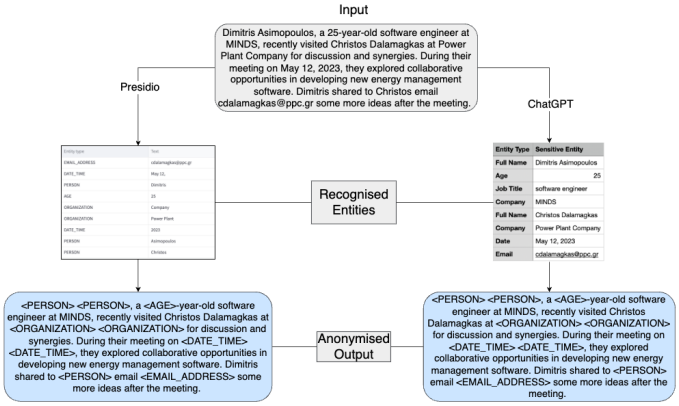

Fig. 5. Example of Anonymisation Process

TABLE I
Evaluation Metrics - CoNLL-2003 dataset

Model	F1-score	Precision	Recall
Presidio	0.85	0.88	0.83
GPT-2	0.71	0.70	0.79

and replacing specific categories of sensitive data. Unlike other methods like substitution or masking, GPT models can generate replacements that are contextually suitable for sensitive information while maintaining cohesiveness and legibility of the text. An example of the anonymization process using Presidio and ChatGPT is shown in Fig. 5.

V. Evaluation Analysis

For the evaluation analysis of AI4COLLAB, we compare the results of the three PII detection methods (Presidio, GPT-2, ChatGPT) in terms of F1 score, precision and recall. In particular, we compare and evaluate GPT-2 and Presidio on the detection of sensitive data on the ConLL-2003 [14] open source dataset. Next, we compare and evaluate Presidio and ChatGPT on the recognition of sensitive entities into different MISP events created by the MISP Converter.

Table I provides the overall metrics of the evaluation made in the models using the CoNLL-2003 dataset. The results are in the scale 0 to 1.

Presidio has proven to be very effective in the anonymization area. With a precision of 0.83, the majority of its predictions were valid and accurate. Its recall was also 0.88, demonstrating how well the model identified and captured a sizeable number of relevant entities from the dataset. With an accuracy and recall balance, the F1 score of 0.85 indicates a well-rounded performance. This strong result demonstrates Presidio's accuracy and breadth of coverage for data anonymization tasks and validates its competence as an anonymization tool.

Table I offers a concise summary of the performance metrics for the GPT-2 model in an anonymization task. The model demonstrates a Precision of 0.70, meaning that 70% of its identifications are accurate. It achieves a Recall of

TABLE II
EVALUATION METRICS FOR MISP EVENTS

Model	F1-score	Precision	Recall
Presidio	0.75	0.90	0.82
ChatGPT	0.89	0.80	0.84

0.79, indicating that it correctly identifies 79% of all relevant instances. The F1-score, which balances Precision and Recall, is 0.71, suggesting a good equilibrium between these metrics. Overall, these figures indicate that the GPT-2 model performs effectively in anonymizing data, though there is potential for improvement, particularly in enhancing precision without significantly compromising recall.

Besides the evaluation that was executed on the dataset, the models where also tested to a majority of MISP events created. For this evaluation process, Microsoft Presidio and ChatGPT were used. The results are presented in Table II.

The classification reports for Presidio and ChatGPT reveal distinct performance characteristics in entity recognition tasks. Presidio demonstrated a Precision of 0.75, indicating that 75% of its detected entities were correct, while its Recall was 0.90, suggesting it successfully identified 90% of all relevant instances. This yielded an F1-score of 0.82, reflecting a solid balance but highlighting the model's tendency towards false positives.

Conversely, ChatGPT achieved a higher Precision of 0.89, meaning nearly 89% of its detected entities were accurate, and a Recall of 0.80, indicating it detected 80% of all relevant instances. Its F1-score stood at 0.84, showing a well-rounded performance but also pointing to some missed entities. These results suggest that while Presidio excels in detecting almost all relevant entities, it does so at the cost of a higher false positive rate. In contrast, ChatGPT maintains a better precision with fewer false positives but misses a few relevant entities compared to Presidio.

VI. CONCLUSIONS

In this work, the AI4COLLAB platform was presented, a CTI solution that focuses on integrating multiple CTI protocols and standards, as well as AI, for improving threat intelligence sharing amongst multiple stakeholders and operators. The entire architecture, based on the C4 model approach, was presented in detail, revealing the internal implementation details of the proposed platform. The usage of AI was demonstrated on detecting and anonymizing sensitive information, realized as a way to foster voluntary collaboration amongst the stakeholders as well as to strengthen trust between the participants.

Future work on AI4COLLAB will focus on enhancing the anonymization techniques, by investigating privacy-preserving techniques (e.g., k-anonymity and differential privacy) to ensure compliance with evolving privacy laws while maximizing data utility. Moreover, we plan to investigate more LLM-based detection models, including GPT-4 and Google Gemini.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 and Horizon Europe research and innovation programme under grant agreement No 101070450 (AI4CYBER). Disclaimer: Funded by the European Union. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the European Commission can be held responsible for them.

REFERENCES

[1] P. R. Grammatikis, P. Sarigiannidis, E. Iturbe, E. Rios, A. Sarigiannidis, O. Nikolis, D. Ioannidis, V. Machamint, M. Tzifas, A. Giannakoulias, M. Angelopoulos, A. Papadopoulos, and F. Ramos, "Secure and private smart grid: The SPEAR architecture," in *2020 6th IEEE Conference on Network Softwarization (NetSoft)*. IEEE, Jun. 2020.

[2] C. Singh, "The european approach to cybersecurity in 2023: A review of the changes brought in by the network and information security 2 (nis2) directive 2022/2555," *International Company and Commercial Law Review*, no. 5, pp. 251–261, 2023.

[3] G. Sakellariou, P. Fouliras, I. Mavridis, and P. Sarigiannidis, "A reference model for cyber threat intelligence (cti) systems," *Electronics*, vol. 11, no. 9, 2022. [Online]. Available: https://www.mdpi.com/2079-9292/11/9/1401

[4] A. Ramsdale, S. Shiaeles, and N. Kolokotronis, "A comparative analysis of cyber-threat intelligence sources, formats and languages," *Electronics*, vol. 9, no. 5, p. 824, 2020.

[5] T. Chantzios, P. Koloveas, S. Skiadopoulos, N. Kolokotronis, C. Tryfonopoulos, V.-G. Bilali, and D. Kavallieros, "The quest for the appropriate cyber-threat intelligence sharing platform." in *DATA*, 2019, pp. 369–376.

[6] D. Mendez Mena and B. Yang, "Decentralized actionable cyber threat intelligence for networks and the internet of things," *IoT*, vol. 2, no. 1, pp. 1–16, 2020.

[7] E. Bandara, S. Shetty, R. Mukkamala, A. Rahaman, and X. Liang, "Luunu—blockchain, misp, model cards and federated learning enabled cyber threat intelligence sharing platform," in *2022 Annual Modeling and Simulation Conference (ANNSIM)*. IEEE, 2022, pp. 235–245.

[8] S. Nikoletos, S. Vlachos, E. Zaragkas, C. Vassilakis, C. Tryfonopoulos, and P. Raftopoulou, "Rog§: A pipeline for automated sensitive data identification and anonymisation," in *2023 IEEE International Conference on Cyber Security and Resilience (CSR)*, 2023, pp. 484–489.

[9] C. Patsakis and N. Lykousas, "Man vs the machine in the struggle for effective text anonymisation in the age of large language models," *Scientific Reports*, vol. 13, 09 2023.

[10] A. Vázquez-Ingelmo, A. García-Holgado, and F. J. García-Peñalvo, "C4 model in a software engineering subject to ease the comprehension of uml and the software," in *2020 IEEE Global Engineering Education Conference (EDUCON)*, 2020, pp. 919–924.

[11] K. M. M. Thein, "Apache kafka: Next generation distributed messaging system," *International Journal of Scientific Engineering and Technology Research*, vol. 3, no. 47, pp. 9478–9483, 2014.

[12] D. P. Kotevski, R. I. Smee, M. Field, Y. N. Nemes, K. Broadley, and C. M. Vajdic, "Evaluation of an automated presidio anonymisation model for unstructured radiation oncology electronic medical records in an australian setting," *International Journal of Medical Informatics*, vol. 168, p. 104880, 2022.

[13] I. Ullah, N. Hassan, S. S. Gill, B. Suleiman, T. A. Ahanger, Z. Shah, J. Qadir, and S. S. Kanhere, "Privacy preserving large language models: Chatgpt case study based vision and framework," *arXiv preprint arXiv:2310.12523*, 2023.

[14] E. F. Tjong Kim Sang and F. De Meulder, "Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition," in *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, 2003, pp. 142–147. [Online]. Available: https://www.aclweb.org/anthology/W03-0419