A Global Analysis of Cyber Threats to the Energy Sector: "Currents of Conflict" from a geopolitical perspective

This is a postprint of a peer-reviewed article, please cite it if using this work:
Gustavo Sanchez, Ghada Elbez, and Veit Hagenmeyer. "A Global Analysis of Cyber Threats to the Energy Sector:"Currents of Conflict" from a geopolitical perspective." atp magazin 67.9 (2025): 56-66.
https://doi.org/10.17560/atp.v67i9.2797

Gustavo Sánchez*
KASTEL Security Research Labs
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
sanchez@kit.edu

Ghada Elbez
KASTEL Security Research Labs
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
ghada.elbez@kit.edu

Veit Hagenmeyer
KASTEL Security Research Labs
Karlsruhe Institute of Technology (KIT)
Karlsruhe, Germany
veit.hagenmeyer@kit.edu

November 20, 2025

Abstract

The escalating frequency and sophistication of cyber threats increased the need for their comprehensive understanding. This paper explores the intersection of geopolitical dynamics, cyber threat intelligence analysis, and advanced detection technologies, with a focus on the energy domain. We leverage generative artificial intelligence to extract and structure information from raw cyber threat descriptions, enabling enhanced analysis. By conducting a geopolitical comparison of threat actor origins and target regions across multiple databases, we provide insights into trends within the general threat landscape. Additionally, we evaluate the effectiveness of cybersecurity tools —with particular emphasis on learning-based techniques— in detecting indicators of compromise for energy-targeted attacks. This analysis yields new insights, providing actionable information to researchers, policy makers, and cybersecurity professionals.

Keywords Artificial Intelligence · Security · Energy · Geopolitics

1 Introduction

In the evolving landscape of cybersecurity, geopolitics plays a defining role in shaping cyber threats, attack motivations, and the strategies employed by both state and non-state actors [7]. The emergence of regional hotspots has led to the formation of cyber alliances, where nations align their cyber defense and offensive capabilities with strategic partners. Countries within alliances such as NATO and BRICS often engage in coordinated cyber operations [9, 10, 14]. Target selection in cyber conflicts is deeply rooted in geopolitical tensions, with adversarial states strategically targeting government agencies, critical infrastructure, and economic assets of rival nations. Thus, scientific literature [8] has identified the role of malware as a pivotal geopolitical tool in the twenty-first century's ever-evolving landscape of international relations and cybersecurity. Cyberattacks in general are often used as instruments of power projection,

*Corresponding author.

deterrence, and asymmetric warfare [13], allowing smaller nations or non-state actors to challenge more technologically advanced adversaries.

However, there is a lack of systematic analysis to verify this intuition across different sources. Gathering insights from different sources is important to avoid potential bias. Beyond direct cyber conflicts, geopolitical factors also shape national cybersecurity practices. Governments have increasingly adopted cyber resilience policies, enforcing stricter regulations on digital infrastructure and cybersecurity cooperation with allies [14]. Meanwhile, adversarial states have developed sophisticated cyber espionage and disruption campaigns, often backed by state-sponsored threat actors [15]. In this regard, Artificial Intelligence (AI) components are now a well-established part of the cyber toolbox, both for defensive and offensive purposes [16]. One of the most vulnerable sectors to geopolitical cyber threats is the energy industry [17, 18, 19]. As nations seek to control global energy markets, cyberattacks against energy infrastructure have become a critical component of geopolitical maneuvering.

The increasing reliance on digital control systems within the energy sector [20] has made it a prime target for cyber-physical attacks, raising concerns over the resilience of national energy security in the face of rising geopolitical tensions. However, existing databases and analytical frameworks often fall short in effectively capturing and analyzing the geopolitical nuances and sectoral focus of cyber threats. This paper aims to address these challenges through three key contributions:

- We leverage Generative AI to extract, structure and interpret information from raw cyber threat descriptions for analytical purposes.
- We conduct a geopolitical analysis with emphasis on the origins and target regions of threat actors and cyber incidents, comparing general trends across databases with those specific to the energy sector.
- We assess the effectiveness of firewalls in detecting Indicators of Compromise (IOCs) for attacks targeting the energy sector, with a focus on AI-based detection.

Additionally, we share the data and code used for this study with the research community[2].

Table 1: Databases used in this work.

Type	Database	Country	Samples	Geographical origins and target regions	Target sectors
Actors	MITRE ATT&CK	USA	163	Often reported in group description text	Often reported in group description text
	ThaiCERT	Thailand	499	Reported explicitly in *JSON* format	Often reported in group description text
	Malpedia	Germany	763	Origin reported explicitly in *JSON* format, target sometimes reported in group description text	Often reported in group description text
Incidents	EuRepoC	Germany	3329	Reported explicitly in tabular form	Reported explicitly in tabular form
	CSIS	USA	580	Reported in incident description text	Reported in incident description text
Reports	AIID	USA	825	Sometimes reported within description	Sometimes reported within description
Malware	Malpedia +VirusTotal API	Germany	3166 families +2400 IOCs	Sometimes reported, e.g., via external URLs	Sometimes reported, e.g., via external URLs

2 Background and Related Work

Threat Actors in the Energy Domain. The energy sector faces a unique blend of cyber threats stemming from a wide range of actors, including state-sponsored groups, cybercriminal organizations, and hacktivists [5]. State-sponsored threat groups are among the most sophisticated adversaries in the energy domain, often leveraging advanced persistent threats (APTs) to conduct espionage, sabotage, or influence operations [21]. These groups operate with the backing of national intelligence agencies, targeting energy infrastructure to gather intelligence on resource distribution, energy policies, and technological advancements, or to disrupt an adversary's energy supply as a form of economic warfare. The energy sector was ranked 19th out of 25 sectors in terms of actual victims of ransomware attacks in 2023 [42]. These groups exploit vulnerabilities in operational technology (OT) and industrial control systems (ICS) to extort payments from energy companies [22], with some even selling stolen data on dark web marketplaces [23]. Hacktivists and ideologically driven actors also pose a significant threat to the energy sector [5], often launching attacks to promote environmental causes, protest against fossil fuel reliance, or disrupt operations of multinational energy corporations. Emerging trends such as AI-driven cyber threats indicate that the energy sector will remain a primary target for cyber adversaries.

[2]https://github.com/gus5298/SecurityThreatsGeopolitics

Evaluation of AI-Based Detection. AI-based detection systems offer several advantages, including their ability to analyze vast amounts of data in real time, identify patterns of malicious behavior, and adapt to new attack techniques through continuous learning. These technologies have significantly improved the detection of sophisticated cyber threats, particularly in identifying zero-day vulnerabilities [24], anomalies [25], and multi-stage attacks [26] that traditional signature-based detection methods may miss. However, despite their strengths, AI/ML-based cybersecurity solutions face several limitations and challenges [27]. One major issue is the susceptibility of AI models to adversarial attacks [29], where threat actors manipulate input data, e.g., to evade detection [28] by an intrusion detection system (IDS). Additionally, the reliance on historical data for training AI models can lead to biases [30], making them less effective against novel attack techniques. The interpretability of AI-driven decisions remains another challenge [31]. While AI enhances threat detection, human analysts are still essential for contextualizing threats, investigating alerts, and making strategic response decisions. The ongoing development of explainable AI (XAI) [32] and hybrid AI-human collaboration models aims to address these challenges by improving the transparency and reliability of AI-driven cybersecurity solutions. However, to fully realize the benefits of AI in cybersecurity, more effort is required to mitigate its limitations and enhance its adaptability.

Related Work. Recent research has shed light on the multifaceted challenges of detecting and mitigating cyber threats, with some looking into topics related to geopolitics:

Yuan *et al.* [3] offer a comprehensive threefold contribution. Their measurement study assesses the effectiveness of existing Phishing Website Detection (PWD) systems across diverse regions, revealing limitations in current approaches. Building on this, they propose enhancements to adapt PWD techniques for both Western and Chinese websites, and underscore the urgency of the issue by releasing all associated tools and datasets to stimulate real-world solution development. Skopik *et al.* [4] take a different angle by presenting a tool that categorizes news items through advanced machine learning algorithms. By extracting and indexing key entities (such as company names, products, CVEs, and attacker groups) their tool groups related news into coherent "stories". This approach not only aids in the rapid identification of emerging trends but also automates report summarization and leverages a collaborative ranking system to prioritize critical information. Turning to smart grid security, Sande-Ríos *et al.* [5] provide a detailed analysis of the adversaries targeting these critical infrastructures, making reference to geopolitics. They build an adversarial model that considers the attack surface, adversary motivations, goals, and capabilities. Regarding analysis of IOCs, Van Liebergen *et al.* [6] analyze the VirusTotal file feed over one year, reviewing 328 million reports for 235 million samples. Their study reveals that despite a feed volume 17 times lower than that of antivirus telemetry, VirusTotal detects eight times more malware.

While these contributions improve our perception of cyber threat detection and response, several research gaps remain. First, comparing diverse data sources across geopolitical contexts—particularly within the smart grid environment—requires further exploration. Second, enhancing comprehension of AI's contribution and performance metrics is necessary. Lastly, despite advancements in automating categorization and summarization, developing systems capable of integrating heterogeneous data sources into cohesive, actionable insights remains a critical requirement. In this paper, we address these gaps.

3 Parsing Methodology

3.1 Databases

We identify relevant databases from varied origins and in different formats. We introduce these sources in this subsection; an overview can be found in Table 1.

(1) MITRE ATT&CK [34]. A globally recognized framework created by the Mitre Corporation (United States) that documents known adversary tactics, techniques and procedures. Their ATT&CK Groups database focuses on cataloging APT groups and cybercriminal organizations, detailing their methods, tools, and motivations. This resource is widely used for threat intelligence and security strategy development. They also provide advanced information on malware families. **(2) ThaiCERT APT Groups [35].** A database maintained by Thailand's national cybersecurity agency. It includes detailed profiles of APT groups, their campaigns, tools, and tactics, with information on their impact. This datasets overlaps with Malpedia [36], so it is not investigated further. **(3) Malpedia [36].** An open platform hosted by Fraunhofer FKIE (Germany) dedicated to providing detailed information on malware families and their associated threat groups. It offers comprehensive descriptions of threat actors, their campaigns, and the malware they deploy, facilitating cross-referencing between malware and the groups behind them. **(4) EuRepoC [37].** The European Repository of Cyber Incidents (EuRepoC) is an independent research consortium established to enhance understanding of the cyber threat landscape within the European Union and globally. Its primary mission is to promote data-driven discussions and policy making in cybersecurity while raising awareness of cyber threats. EuRepoC provides an analytical framework to assess and compare the lifecycle of cyber incidents, with a focus on technical, political, and legal dimensions. **(5)**

CSIS Signifcant Cyber Events List [38]. Based in the United States, the Center for Strategic and International Studies (CSIS) think tank maintains a Significant Cyber Events database, tracking major cybersecurity events worldwide. **(6) AIID reports [39].** The Artificial Intelligence Incident Database (AIID) is a centralized repository that documents real-world incidents involving artificial intelligence systems. It includes cases of AI-related vulnerabilities, misuse, and failures in various domains.
Note. The Chinese CERT [43] releases geographical information in monthly textual summaries, but only distinguishes between "national" and "cross-border" incidents, limiting granularity and therefore is not considered in this study.

3.2 Data Processing with Generative AI

Parsing unstructured threat intelligence data requires a robust and scalable approach to extract meaningful insights. In this paper, we utilize a Generative AI Model (gemini-1.5-flash-latest) to structure raw cyber incidents text data into structured fields for further analysis, as exemplified in Figure 1. The parsing strategy is designed to minimize Large Language Model (LLM) token consumption.

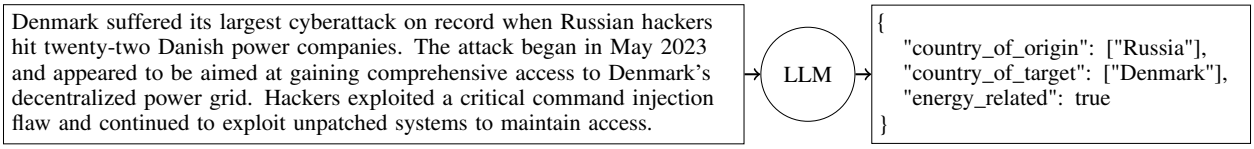


Figure 1: Example of a cyberattack incident description and extracted fields.

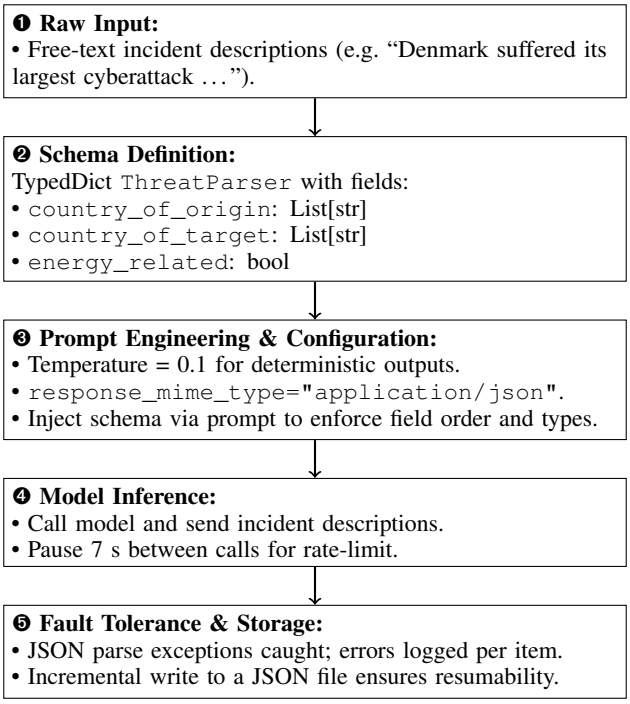


Figure 2: Flow of the generative-AI parsing pipeline, from raw description to structured JSON.

As represented in Figure 2, the parsing workflow begins with loading the input dataset containing unstructured threat intelligence data (❶). The next step is codifying our target JSON format in a Python *TypedDict*[3], that we name *ThreatParser* (❷). Embedding this schema directly in the system prompt ensures that Gemini outputs precisely the expected keys (and in the correct order), avoiding downstream validation errors. We empirically tuned the model's temperature to 0.1—balancing variability and determinism—and enforced the target JSON format (❸).

To respect API rate limits and manage token costs, a fixed 7-seconds delay is inserted between calls (❹). All responses undergo a 'try/except' JSON parse: failures (e.g., reaching the API quota limit) are recorded with an error flag (❺).

[3] A TypedDict type represents dictionary objects with a specific set of string keys, and with specific value types for each valid key.

Partial results are written immediately to a JSON file so that parsing can resume after any interruption. The structured output captures key features, including the attack's origin, target, and domain (i.e., whether it is energy-related or not). Finally, all parsed data is stored in JSON format, preserving structured insights for downstream research and visualization.

3.3 Evaluation

To assess the performance of our generative-AI parser in classifying descriptions as energy-related, we created a stratified evaluation set of 200 threat descriptions (100 energy, 100 non-energy) drawn at random from the EuRepoC dataset. Each entry was labeled based on the `receiver category` attribute, and then fed to the same pipeline. The evaluation results yielded an overall accuracy of 84.0%. The following confusion matrix (rows = true class; columns = predicted) provides further details: $\begin{pmatrix} 91 & 9 \\ 23 & 77 \end{pmatrix}$. We observe that 91 true non-energy incidents are correctly classified, 9 non-energy are mislabeled as energy, 23 energy are mislabeled as non-energy, and 77 true energy are correctly identified. From this matrix we compute additional class-specific metrics for the energy class:

$$\text{Precision}_{\text{energy}} = \frac{77}{77+9} \times 100\% \approx 89.5\%, \quad \text{Recall}_{\text{energy}} = \frac{77}{77+23} \times 100\% = 77.0\%,$$
$$F_{1,\text{energy}} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \approx 82.7\%.$$

Upon inspection, the sentences that are not correctly labeled are (1) those that lack explicit details[4] about the energy targets, and (2) those labeled as "Government", e.g., in the cases where the target is the US Department of Energy. These results demonstrate that our schema-enforced, low-temperature prompting approach achieves strong overall accuracy, high precision on the energy class, and acceptable recall given the challenging, unstructured nature of threat descriptions.

For comparison, the spaCy [44] rule-based baseline (using a `PhraseMatcher` with 16 domain-specific terms[5] such as "energy", "power grid", etc.) reached an overall accuracy of 81%. Its confusion matrix $\begin{pmatrix} 95 & 5 \\ 34 & 66 \end{pmatrix}$, corresponds to a precision of 93% and a recall of 66% for the energy class, giving an F_1 score of 77%.

In summary, these results demonstrate that our schema-enforced, low-temperature Gen-AI pipeline not only outperforms a rule-based baseline—boosting recall by + 11 percentage points (77% vs. 66%) on energy-related cases—but also maintains comparably high precision, underscoring the practical value of generative AI for accurately extracting nuanced domain signals from unstructured threat descriptions.

Table 2: Comparison of classification metrics for energy-domain detection.

Metric	Our Gen-AI Parser	spaCy Baseline
Accuracy	84.0%	81.0%
Precision (energy)	89.5%	93.0%
Recall (energy)	77.0%	66.0%
F_1 (energy)	82.7%	77.0%

4 Geopolitical Big Data Analysis Results

The reporting of geographical origins and target regions varies significantly across different cyber threat databases, reflecting disparities in structure, detail, and accessibility. Table 1 summarizes the geographical reporting practices of major cyber threat intelligence resources. Some databases provide structured, machine-readable formats that enable streamlined analysis. For instance, ThaiCERT and Malpedia report geographical origins explicitly in JSON format, facilitating automation and consistency. However, target regions in Malpedia are often embedded within descriptive text, requiring additional processing via the proposed generative AI pipeline. EuRepoC explicitly documents both origins and targets in tabular form, offering a standardized approach allowing comparative analysis. In contrast, databases such as MITRE ATT&CK Groups and CSIS rely on unstructured descriptive text to document geographical information. While rich in qualitative details, these formats necessitate our processing technique to structure the data for analysis. Not all

[4] An example of this is: "Likely Iranian State-sponsored hackers (Crowd strike) have conducted a series of destructive attacks on Saudi Arabia over the last two weeks, erasing data and wreaking havoc in the computerbanks of the agency running the country's airports and hitting five additional targets".

[5] Requires manually gathering keywords via expert knowledge, while our Gen-AI approach only requires to specify the domain itself, e.g.: energy.

databases include geographical reporting. Notably, AIID omits geographical origins and targets entirely, reducing their utility for geopolitical studies. This omission prevents understanding the spatial distribution of cyber threats. Using our generative AI pipeline, we are still able to extract whether the AI vulnerabilities are related to the energy domain.

Takeaway 1. There is heterogeneity in reporting practices among cyber threat databases. Structured reporting (e.g. JSON), enhances consistency and facilitates cross-database comparisons. The reliance on descriptive text or omission of geographical information hinders the ability to conduct comprehensive geopolitical analyses. Our generative AI pipeline allows for better analysis by parsing relevant unstructured information.

From the provided plots in Figure 3, each subfigure shows the top 5 of either threat origins or targets, with bars grouped by energy vs. non-energy related. The three relevant data sources to answer this question—CSIS, Malpedia, and EuRepoC—each provide a slightly different perspective on which countries appear most frequently. In general, the same few countries dominate both energy and general threats (i.e., Russia and China as origins, and the USA as target). However, the bar heights (counts) can differ substantially. For instance, a country might be the top origin overall but drop to a lower position (or even disappear) in the energy-related subset, indicating that some origins are especially active in the general domain but less so in energy. The energy-related bars show a higher concentration in fewer countries—often, one or two countries account for most of the energy-focused incidents—suggesting that certain threat actors specialize in energy infrastructure attacks. By contrast, the general category includes many different targets (government agencies, corporations, etc.). This can make the energy domain appear narrower but more heavily hit by a smaller set of threat actors. The USA appears as a top general target across datasets, while in energy incidents, the Middle East as a region takes the first place by a small difference (according to Malpedia).

In Figure 4, we group each incident's origin (and, analogously, its target) by major geopolitical alliances—namely NATO, BRICS, or "Other"—using up-to-date membership rosters as of January 2025. Specifically, our NATO list includes the 32 member states from Albania through the United States (adding Sweden, which joined in July 2023), while the BRICS bloc encompasses not only the original five (Brazil, Russia, India, China, South Africa) but also the ten countries formally participating in the 2024–25 BRICS expansion (Egypt, Ethiopia, Indonesia, Iran, and the United Arab Emirates), providing a relevant, up-to-date overview. Any country not found in either list is classified as "Other".

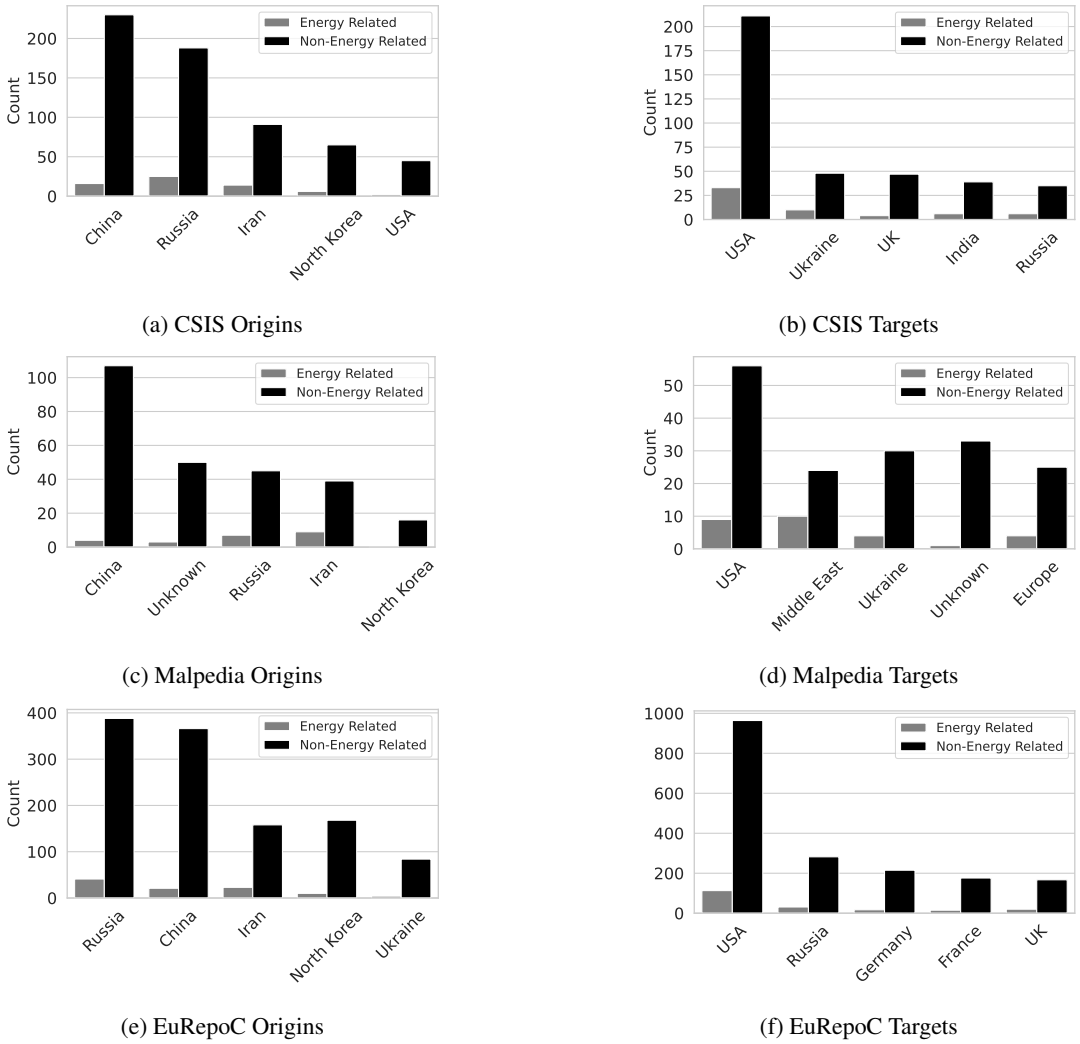
Concretely, we take each record's country-of-origin field (which may be a list), standardize the country names, and then map each to its alliance via a lookup. We then group by alliance membership; every origin country contributes separately to the aggregate counts, so that each pair (incident, country) becomes its own row. Finally, we assign an alliance label, and plot counts for "Energy Related" versus "Non-Energy Related" in matching bar charts. This clustering by alliance reveals, for example, that BRICS countries dominate non-energy threat origins across all datasets, whereas energy-focused incidents show a comparatively higher concentration in "Other" or NATO members depending on the source. In the Malpedia dataset (Figures 4c and 4d), we observe the biggest contrast between alliances.

Takeaway 2. General threats tend to involve a broader set of countries both in terms of origins and targets, whereas energy-related attacks are often more narrowly concentrated. The same few countries dominate the overall ranking—e.g., Russia, China, or the USA—but for energy-focused incidents, certain actors seem to specialize.

From the plots in Figure 5, it is clear that cyber incidents in conflict regions are neither uniformly distributed over time nor concentrated in a single geographic theater. Instead, they form distinct clusters that align with specific escalations in each ongoing conflict. In the Russia–Ukraine timeline, a noticeable uptick appears around 2014 (following the annexation of Crimea) and accelerates substantially from 2022 onward. Israel–Palestine incidents show intermittent bursts—fewer overall than Russia–Ukraine, but still recurring during major flare-ups in the region. Meanwhile, the China–Taiwan timeline is comparatively sparser but displays a steady trickle of events spanning multiple years, suggesting a slow-burn pattern of cyber activity. Each conflict region thus exhibits its own rhythm of threat incidents, typically intensifying around key military escalations.

Takeaway 3. Heightened conflicts correlate with increased frequency of attacks. Russia's persistent cyber campaigns since 2014 reflect ongoing regional disputes. In contrast, the Israel–Palestine conflict shows sharp, periodic spikes in activity that mirror its cyclical hostilities, while China–Taiwan experiences a steady, lower-intensity flow of incidents, suggesting an enduring contest of influence rather than a single, large-scale campaign.

We perform a series of experiments to classify and analyze incidents from the AIID, which contains almost 800 records. The goal is to identify and categorize AI incidents related to specific topics, with a particular focus on the energy domain. Figure 6a shows that the top alleged harmed parties in the AIID dataset include a broad range of groups—such as general-public, democracy, or Tesla drivers—while energy- or power-sector victims do not prominently appear among the top ten. It is important to note that, in the case of Tesla cars, AI vulnerabilities lie in the autonomous driving



(a) CSIS Origins

(b) CSIS Targets

(c) Malpedia Origins

(d) Malpedia Targets

(e) EuRepoC Origins

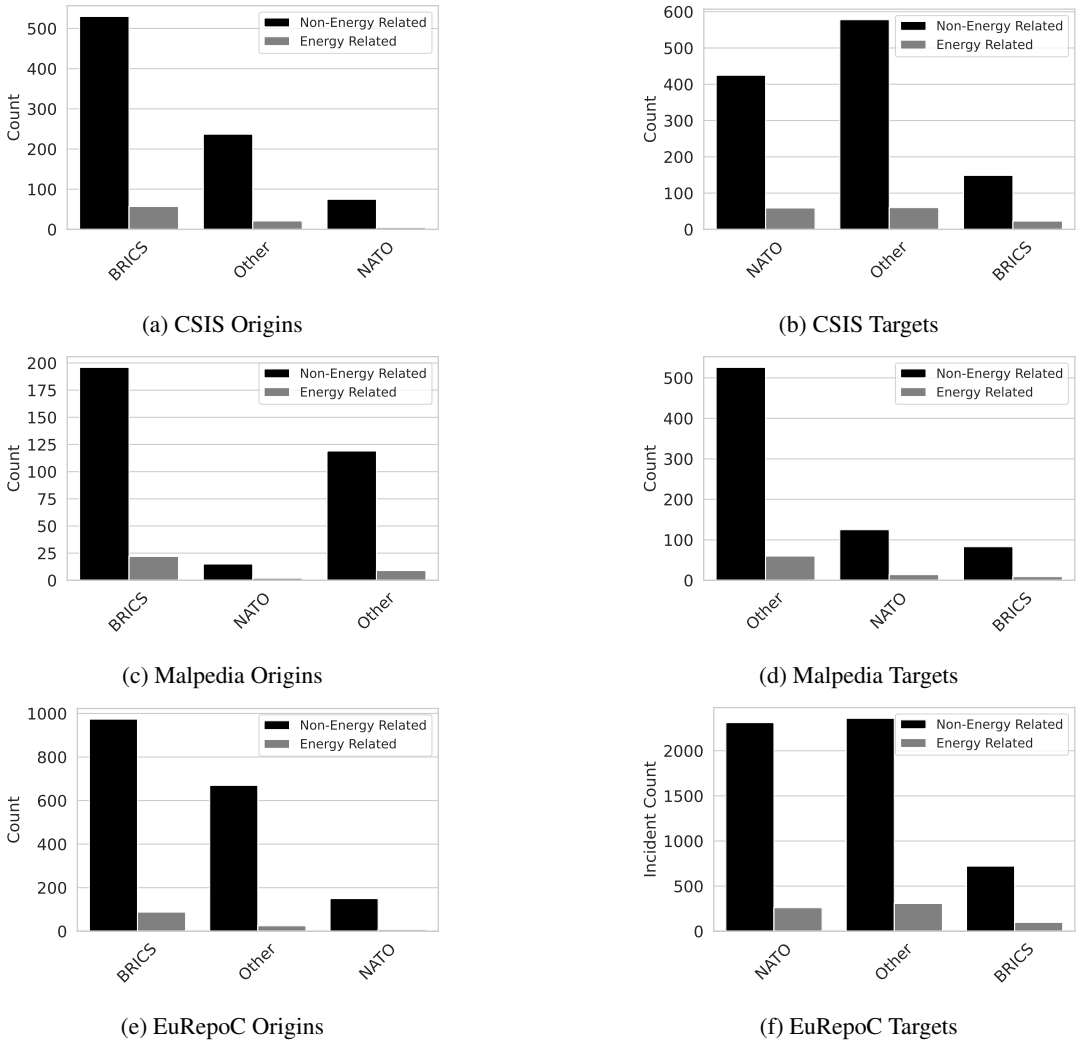
(f) EuRepoC Targets

Figure 3: Top 5 threat origins and targets per dataset.

or manufacturing automation processes, and not in the electric vehicle infrastructure. Relative to other domains, explicit AI-based incidents in the energy sector are less common or, at least, less frequently reported. Meanwhile, Figure 6b, which plots AI threat incidents over time and highlights energy-related events with red triangles, reveals that such incidents do exist but remain comparatively sparse. To shed light into this, we describe the specific AI and energy related incidents in Table 3; in some of these incidents, the energy relation is relatively loose: the generative AI model considers e.g., smart home devices as part of the smart grid, and the climate crisis a subclass of the energy domain. As AI technologies continue to integrate into critical infrastructure, including power grids and smart devices, the potential for energy-targeted attacks will likely grow.

Table 3: Most energy-related Cyber Incidents in AIID.

Date	Alleged Developer	Alleged Harmed	Details
2016-10-08	Tesla	Tesla	Poor Performance of Tesla Factory AI Robots (assembling lithium batteries)
2014-01-21	Nest Labs	Fire victims	Smoke + CO Alarm could inadvertently silence genuine alarms (smart home)
2019-03-01	Scammers	UK Energy Firm's CEO	Fraudsters Used AI to Mimic Voice
2020-04-14	Belgian action group	Belgian government	Deepfake of Belgian Prime Minister Urging Climate Crisis Action
2019-02-01	YouTube	YouTube users	Recommendation Algorithm Allegedly Promoted Climate Misinformation Content
2024-10-14	Portland Water Bureau	City of Portland	Portland Water Bureau Algorithm Reportedly Allocates Utility Bill Discount to High-Wealth Consumer



(a) CSIS Origins

(b) CSIS Targets

(c) Malpedia Origins

(d) Malpedia Targets

(e) EuRepoC Origins

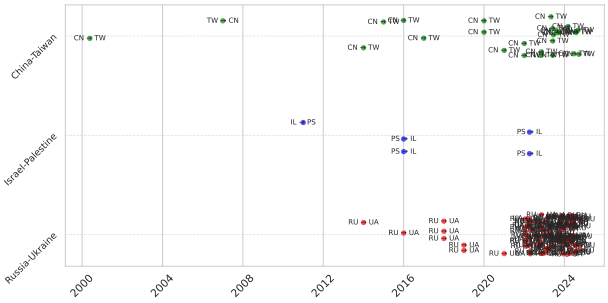
(f) EuRepoC Targets

Figure 4: Top 5 threat origins and targets per dataset, clustered by alliances.

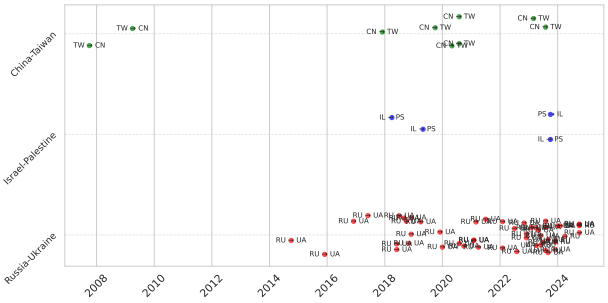
Takeaway 4. These observations imply that AI-related vulnerabilities in the energy domain exist but they appear overshadowed by a wider array of AI threats affecting broader consumer or societal areas.

We analyze the performance of various antivirus engines, with an emphasis on SentinelOne and Acronis, which explicitly report the usage of static machine learning models. We filter out cyber threat groups related to the energy domain. Then, we extract the information about tools that they leverage (e.g., malware families). From the tool names, we fetch IOCs (i.e., file hashes) via Malpedia. These hashes are then submitted to the VirusTotal platform via API to gather detection results from various antivirus engines, both AI-based and traditional. The VirusTotal results were stored locally to avoid redundant queries, preserving API quota for future experiments.

Our analysis shows that only a fraction of IOCs (12.85%) are explicitly tied to energy-focused threats, and that Static ML tools (Acronis and SentinelOne) detect only around 46.8% of malicious indicators, whereas other engines (the *Others* category) detect roughly 88.4%—on par with the overall average (88.4%). This indicates that, for the malware samples used to target the energy domain, traditional solutions are outperforming the ML-based ones. Figure 7 presents a side-by-side comparison of each antivirus engine's detection rate, with Acronis and SentinelOne (the static ML engines) emphasized in red. The top-performing engine in this dataset achieves a higher detection rate than either of these ML-based solutions.



(a) EuRepoC Dataset

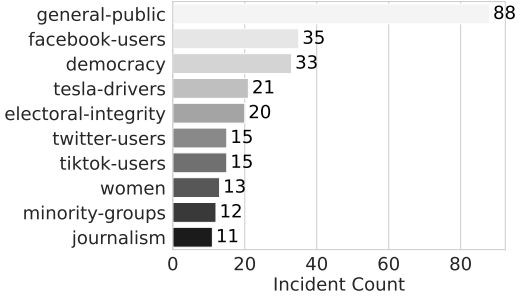


(b) CSIS Dataset

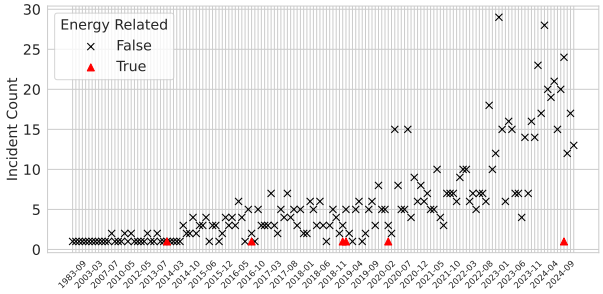
	Date	Initiator	Target	Details
CSIS	2015-12	Russia	Ukraine	DoS and SCADA attacks
	2016-12	Russia	Ukraine	Ukrenergo shut down
	2017-06	Russia	Ukraine	Ransomware
	2018-06	Russia	Ukraine	Backdoors
	2020-01	Russia	Ukraine	Infiltration
	2020-05	China	Taiwan	Malware attacks
	2020-10	Russia	Ukraine	Officers indicted (’16 attacks)
	2022-08	Russia	Ukraine	State energy website hack
	2022-11	Russia	Ukraine	Energy and logistics sector hack
	2022-12	Russia	Ukraine	Sandworm APT
	2023-12	Ukraine	Russia	Encryption/deletion of data
EuRepoC	2024-09-19	China	Taiwan	Spear-phishing
	2024-04-19	Russia	Ukraine	Sandworm APT
	2023-06-15	China	Taiwan	CVE-2023-2868
	2023-01-29	Ukraine	Russia	Gazprom hack-and-leak
	2023-01-31	Russia	Ukraine	Sandworm APT
	2022-08-17	Russia	Ukraine	Energoatom website hack
	2022-02-28	Russia	Ukraine	Sandworm APT
	2022-04-12	Russia	Ukraine	Sandworm APT
	2022-04-12	Russia	Ukraine	Sandworm APT

(c) Energy-Related Cyber Incidents in Current Conflicts

Figure 5: Comparison of conflict chronologies (top row) and energy-related cyber incidents (bottom row).



(a) Top 10 Alleged Harmed Parties in the AIID dataset.



(b) AI Threat Incidents Over Time in the AIID dataset.

Figure 6: Analysis of AI Threat Incidents in the AIID dataset.

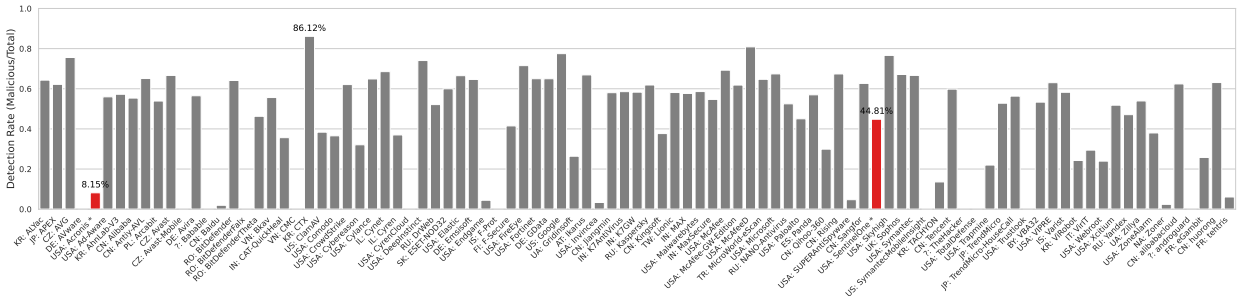


Figure 7: Detection rate by antivirus engine. Static ML engines (Acronis and SentinelOne) are highlighted in red.

Takeaway 5. Although ML-based cybersecurity solutions are frequently touted as being especially agile at catching new or sophisticated attacks, the data here suggests they may lag behind more traditional or hybrid approaches when confronting malware associated with the energy domain.

In this work, we analyze the use of AI—specifically, static machine learning—as a defensive method for detecting IOCs, and generative AI to support natural language tasks. However, the malicious use of AI to enhance cyber operations is an emerging threat (unless done ethically, e.g., for web application penetration testing [41]) with significant geopolitical implications and the potential to impact energy systems. Recent reports [40] have highlighted three cases of AI-related cyber operations involving OpenAI's ChatGPT: **(1) SweetSpecter**: This suspected China-based adversary uses OpenAI's services for reconnaissance, vulnerability research, scripting support, and evasion of anomaly detection, as well as for development. **(2) CyberAv3ngers**: This group, suspected of being affiliated with the Iranian Islamic Revolutionary Guard Corps (IRGC), employs GPT models to conduct research on programmable logic controllers. CyberAv3ngers is known for disruptive attacks on industrial control systems and programmable logic controllers (PLCs) in sectors such as water, manufacturing, and energy. Their targeted infrastructure is typically associated with Israel, the United States, or Ireland. **(3) STORM-0817**: An Iranian threat actor identified as STORM-0817 is developing malware and tools to scrape social media. This actor was found to use OpenAI's models to debug malware, receive coding assistance for building a basic Instagram scraper, and translate LinkedIn profiles into Persian, aimed at identifying potential targets. These three cases —detected based on credible tips— illustrate that while AI offers robust defensive capabilities, adversaries are increasingly harnessing these technologies to advance their offensive operations—raising new challenges for global cybersecurity, especially in critical sectors like energy.

Recommendations. Building on our findings, we propose:

1) **Extension to Other Automation Domains.** To leverage the same generative-AI pipeline to ingest and normalize data from multiple automation sectors. This is applicable to actors, incidents, reports and malware descriptions. While our primary focus is the energy sector, the generative-AI pipeline and subsequent geopolitical clustering extend readily to other automation environments—such as process industry, automotive manufacturing, or water treatment. This can be easily achieved by swapping the energy-specific keyword for a domain-appropriate one, e.g., using *automotive-related* instead of *energy-related* in the schema definition during prompt engineering (see Figure 2). This translates into a simple code update, demonstrating that our method can provide actionable risk insights across diverse automation sectors with minimal reconfiguration. The possibility of summarizing results for different domains into an "Automation Threat Dashboard" for real-time comparative risk assessment is especially interesting.
2) **Dynamic Adversarial Augmentation.** Generating key synthetic examples of adversarially-crafted threat descriptions and including them in the prompts to harden the parser against emerging jargon or evasion tactics.

5 Conclusion

This paper describes different threat intelligence databases, and introduces a way to parse their unstructured data using generative AI for further analysis. Our findings indicate that cyber threats targeting the energy sector follow distinct patterns compared to general threats. Specialized threat actors frequently emerge as the top origins for energy-focused incidents, coinciding with geopolitical flashpoints. AI-based detection tools—specifically static ML solutions—do not currently outperform traditional antivirus engines for detecting malware used against the energy domain. We observe that AI components are both attack vector and attack surface. These insights, obtained using our generative AI

parsing methodology, underscore the need for geopolitical awareness, focused threat intelligence, and improved model prompting tailored to avoid the impact of bias and adversarial perturbations.

References

[1] Kaya, Y., Chen, Y., Saha, S., Pierazzi, F., Cavallaro, L., Wagner, D., and Dumitras, T. (2024). "Demystifying Behavior-Based Malware Detection at Endpoints." *arXiv:2405.06124*.

[2] Saeed, M. H., Ali, S., Paudel, P., Blackburn, J., and Stringhini, G. (2024). "Unraveling the Web of Disinformation: Exploring the Larger Context of State-Sponsored Influence Campaigns on Twitter." In *RAID'24*.

[3] Yuan, Y., Apruzzese, G., and Conti, M. (2025). "Beyond the West: Revealing and Bridging the Gap Between Western and Chinese Phishing Website Detection." *Computers & Security*.

[4] Skopik, F., Akhras, B., Woisetschläger, E., Andresel, M., Wurzenberger, M., and Landauer, M. (2024). "On the Application of Natural Language Processing for Advanced OSINT Analysis in Cyber Defence." In *ARES'24*.

[5] Sande-Ríos, J., Canal-Sánchez, J., Manzano-Hernández, C., and Pastrana, S. (2024). "Threat Analysis and Adversarial Model for Smart Grids." In *IEEE EuroS&PW*.

[6] van Liebergen, K., Caballero, J., Kotzias, P., and Gates, C. (2023). "A Deep Dive into the VirusTotal File Feed." In *DIMVA'23*.

[7] Nocetti, J. (2018). "The Geopolitics of Cyberconflict." *Politique étrangère*, Institut français des relations internationales.

[8] Leventopoulos, S., Gritzalis, D., and Stergiopoulos, G. (2024). "Malware as a Geopolitical Tool." In *Malware: Handbook of Prevention and Detection*. Springer.

[9] Smeets, M. (2019). "NATO Members' Organizational Path Towards Conducting Offensive Cyber Operations: A Framework for Analysis." In *CyCon'19*.

[10] Belli, L. (2021). "Cybersecurity Policymaking in the BRICS Countries: From Addressing National Priorities to Seeking International Cooperation." *The African Journal of Information and Communication*.

[11] Radu, N. C.-C. (2024). "The People's Republic of China and the Threats to NATO's Cyber Security." In *Romanian Military Thinking International Conference*. DOI: 10.55535/RMT.2024.4.17.

[12] Izycki, E., Van Niekerk, B., and Ramluckan, T. (2023). "Cyber Diplomacy: NATO/EU Engaging with the Global South." In *CyCon'23*.

[13] Todd, G. H. (2009). "Armed Attack in Cyberspace: Deterring Asymmetric Warfare with an Asymmetric Definition." *AFL Rev.*

[14] Fracalossi de Moraes, R. (2020). "Whither Security Cooperation in the BRICS? Between the Protection of Norms and Domestic Politics Dynamics." *Global Policy*. Wiley Online Library.

[15] Hunter, L. Y., Albert, C. D., and Garrett, E. (2021). "Factors That Motivate State-Sponsored Cyberattacks." *The Cyber Defense Review*. JSTOR.

[16] Truong, T. C., Diep, Q. B., and Zelinka, I. (2020). "Artificial Intelligence in the Cyber Domain: Offense and Defense." *Symmetry*. MDPI.

[17] Maliarchuk, T., Danyk, Y., and Briggs, C. (2019). "Hybrid Warfare and Cyber Effects in Energy Infrastructure." *Connections*. JSTOR.

[18] Pollard, M. (2024). "A Case Study of Russian Cyber-Attacks on the Ukrainian Power Grid: Implications and Best Practices for the United States." *Pepperdine Policy Review*.

[19] Whitehead, D. E., Owens, K., Gammel, D., and Smith, J. (2017). "Ukraine Cyber-Induced Power Outage: Analysis and Practical Mitigation Strategies." In *CPRE'17*. IEEE.

[20] Baidya, S., Potdar, V., Ray, P. P., and Nandi, C. (2021). "Reviewing the Opportunities, Challenges, and Future Directions for the Digitalization of Energy." *Energy Research & Social Science*. Elsevier.

[21] Lu, Q., Peng, Z., Wu, L., Ni, M., Luo, J., and others. (2024). "Detecting the Cyber-Physical-Social Cooperated APTs in High-DER-Penetrated Smart Grids: Threats, Current Work and Challenges." *Computer Networks*. Elsevier.

[22] Nguyen, L.-H., Nguyen, V.-L., Hwang, R.-H., Kuo, J.-J., Chen, Y.-W., Huang, C.-C., and Pan, P.-I. (2024). "Towards Secured Smart Grid 2.0: Exploring Security Threats, Protection Models, and Challenges." *IEEE Communications Surveys & Tutorials*.

[23] Pantelis, G., Petrou, P., Karagiorgou, S., and Alexandrou, D. (2021). "On Strengthening SMEs and MES Threat Intelligence and Awareness by Identifying Data Breaches, Stolen Credentials and Illegal Activities on the Dark Web." In *ARES'21*.

[24] Ahmad, R., Alsmadi, I., Alhamdani, W., and Tawalbeh, L. (2023). "Zero-day Attack Detection: A Systematic Literature Review." *Artificial Intelligence Review*. Springer.

[25] Siniosoglou, I., Radoglou-Grammatikis, P., Efstathopoulos, G., Fouliras, P., and Sarigiannidis, P. (2021). "A Unified Deep Learning Anomaly Detection and Classification Approach for Smart Grid Environments." *IEEE Transactions on Network and Service Management*. IEEE.

[26] Jia, Y., Gu, Z., Du, L., Long, Y., Wang, Y., Li, J., and Zhang, Y. (2023). "Artificial Intelligence Enabled Cyber Security Defense for Smart Cities: A Novel Attack Detection Framework Based on the MDATA Model." *Knowledge-Based Systems*. Elsevier.

[27] Arp, D., Quiring, E., Pendlebury, F., Warnecke, A., Pierazzi, F., Wressnegger, C., Cavallaro, L., and Rieck, K. (2022). "Dos and Don'ts of Machine Learning in Computer Security." In *USENIX Sec'22*.

[28] Mumrez, A., Sánchez, G., Elbez, G., and Hagenmeyer, V. (2023). "On Evasion of Machine Learning-Based Intrusion Detection in Smart Grids." In *SmartGridComm'23*. IEEE.

[29] Sánchez, G., Elbez, G., and Hagenmeyer, V. (2024). "Attacking Learning-Based Models in Smart Grids: Current Challenges and New Frontiers." In *e-Energy'24*.

[30] Andresini, G., Pendlebury, F., Pierazzi, F., Loglisci, C., Appice, A., and Cavallaro, L. (2021). "Insomnia: Towards Concept-Drift Robustness in Network Intrusion Detection." In *AISec@CCS'21*.

[31] Neupane, S., Ables, J., Anderson, W., Mittal, S., Rahimi, S., Banicescu, I., and Seale, M. (2022). "Explainable Intrusion Detection Systems (x-IDs): A Survey of Current Methods, Challenges, and Opportunities." *IEEE Access*.

[32] Moustafa, N., Koroniotis, N., Keshk, M., Zomaya, A. Y., and Tari, Z. (2023). "Explainable Intrusion Detection for Cyber Defences in the Internet of Things: Opportunities and Solutions." *IEEE Communications Surveys & Tutorials*. IEEE.

[33] Guato Burgos, M. F., Morato, J., and Vizcaino Imacaña, F. P. (2024). "A Review of Smart Grid Anomaly Detection Approaches Pertaining to Artificial Intelligence." *Applied Sciences*. MDPI.

[34] MITRE. (n.d.). *MITRE ATT&CK Groups*. From `https://attack.mitre.org/groups/` (Accessed: Feb 2025).

[35] ETDA. (n.d.). *APT Groups*. From `https://apt.etda.or.th/cgi-bin/aptgroups.cgi` (Accessed: Feb 2025).

[36] Fraunhofer FKIE. (n.d.). *Malpedia Actors*. From `https://malpedia.caad.fkie.fraunhofer.de/actors` (Accessed: Feb 2025).

[37] EUREPOC. (n.d.). *EUREPOC Table View*. From `https://eurepoc.eu/table-view/` (Accessed: Feb 2025).

[38] Center for Strategic and International Studies. (n.d.). *Significant Cyber Incidents*. From `https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents` (Accessed: Feb 2025).

[39] Incident Database. (n.d.). *Incident Summaries*. From `https://incidentdatabase.ai/summaries/incidents/` (Accessed: Feb 2025).

[40] OpenAI Threat Intelligence. (2024, October). *Influence and Cyber Operations: An Update*. From `https://cdn.openai.com/threat-intelligence-reports/influence-and-cyber-operations-an-update_October-2024.pdf` (Accessed: Feb 2025).

[41] Sánchez, G., Olayinka, O., and Pasikhani, A. (2024). *Web Application Penetration Testing with Artificial Intelligence: A Systematic Review.* In *NCA'24*. IEEE.

[42] PwC Germany. (n.d.). *Under the Lens: The Energy Sector*. From `https://www.pwc.de/de/energiewirtschaft/under-the-lens-the-energy-sector.pdf` (Accessed: Feb 2025).

[43] CERT China. (n.d.). *CERT China*. From `https://www.cert.org.cn/publish/english/55/index.html` (Accessed: Feb 2025).

[44] Explosion AI, *Rule-based Matching — PhraseMatcher*, `https://spacy.io/usage/rule-based-matching#phrasematcher`, accessed 25 July 2025.