Journal of Electrical Engineering and Computer (JEECOM)
Vol. 8, No. 1 (2026), DOI: 10.33650/jeecom.v8i1.17163
p-ISSN: 2715-0410; e-ISSN: 2715-6427

Deep Learning Driven Ransomware Detection: A Bibliometric Analysis of Research Trends, Knowledge Structure, and Future Directions

Guntoro Guntoro*[1], Lisnawita Lisnawita[2], Loneli Costaner[3], Wenni Syafitri[4]
[1,2,3]Department of Informatic Engineering, Universitas Lancang Kuning, Pekanbaru, Indonesia
[4] Department of Computer Science, Universitas Lancang Kuning, Pekanbaru, Indonesia

Article Info

Article history:

Received April 03, 2026
Revised April 28, 2026
Accepted April 30, 2026

Keywords:

Ransomware detection
Deep learning
Machine learning
Behavioral analysis
Bibliometric analysis

ABSTRACT

Ransomware has become a major cybersecurity threat because it encrypts data, disrupts services, and evolves rapidly beyond conventional signature-based detection. This study presents a bibliometric analysis of deep learning-driven ransomware detection research, with emphasis on behavioral and dynamic analysis, API-call sequences, system calls, and future research directions. A total of 481 records were retrieved from the Web of Science Core Collection; after screening one retracted publication and two editorial materials, 478 eligible records remained. The eligible corpus was analyzed using Biblioshiny and Bibliometrix. Results show rapid growth from 2022 to 2025, with IEEE Access, Computers & Security, Sensors, Scientific Reports, and International Journal of Information Security among the prominent sources. Keyword and thematic analyses indicate a shift from static detection toward behavior-aware, sequence-based, explainable, and real-time ransomware detection. The findings highlight the need for robust datasets, cross-dataset validation, low-latency inference, explainable deep learning, and integrated detection systems for practical cybersecurity deployment.

Corresponding Author:
Guntoro Guntoro
Universitas Lancang Kuning, Jl. Yos Sudarso No. KM. 8, Umban Sari, Pekanbaru, 28266, Indonesia
Email: guntoro@unilak.ac.id; Orchid ID : https://orcid.org/0000-0002-3084-7582

1. INTRODUCTION

Ransomware is a high-impact malware threat because it compromises data availability by encrypting files, disabling services, and disrupting operations. Its detection problem is not confined to recognizing a known binary; it also requires identifying early execution activities before encryption becomes irreversible. This has increased interest in detection systems that can operate across endpoints, enterprise networks, cloud infrastructures, and Internet of Things environments while remaining robust to obfuscation, polymorphism, and previously unseen variants [1], [2], [21].

Static and dynamic or behavioral analysis are detection paradigms and data-acquisition strategies, whereas machine learning (ML) and deep learning (DL) are computational methods applied to the resulting data. Static analysis examines artifacts such as Portable Executable headers, file structure, and entropy. Dynamic or behavioral analysis records runtime evidence, including API and system calls, file operations, process activity, memory access, and changes in resource usage. These observations provide temporal evidence of infection and may reveal suspicious activity before complete encryption [3], [6], [7], [8].

DL is particularly relevant to runtime data because behavioral traces are high-dimensional, sequential, nonlinear, and often variable in length. DL models can learn hierarchical representations and temporal dependencies from such traces, making them suitable for API-call sequences, system-call streams, and dynamic execution logs. Recent studies using LSTM, BERT, RoBERTa, and related architectures report progress in ransomware detection, family classification, explainability, and early detection [6], [7], [10], [11]. Classical ML remains useful as a baseline and for structured, low-cost features, but it is treated here as methodological context rather than the central computational focus [1], [9].

Despite these developments, ransomware detection research remains fragmented across data sources, feature representations, model architectures, evaluation metrics, and deployment assumptions. Prior reviews have mainly summarized selected studies narratively or systematically, leaving limited quantitative evidence about how DL methods are connected to behavioral and dynamic analysis, which contributors shape the field, and which themes are emerging. This gap is important because progress in DL-driven detection must be assessed together with false-positive control, inference latency, interpretability, dataset diversity, robustness to concept drift, and operational feasibility [2], [5], [15], [17], [18].

Accordingly, this study conducts a bibliometric analysis of ransomware detection research at the intersection of DL and behavioral or dynamic analysis. The search produced 481 Web of Science records, of which 478 eligible records remained after document-type screening. The study maps publication growth, leading sources and contributors, citation patterns, keyword co-occurrence, thematic structure, collaboration, and future research directions. Its specific contribution is to connect the computational perspective of DL with the data and detection paradigms of static, dynamic, and behavioral analysis, thereby providing a quantitative view of how the field is moving toward early, explainable, robust, and deployable ransomware detection [6], [7], [10], [17], [18].

Based on this background, the study addresses the following research questions:

1. RQ1. How has scientific production on machine learning and deep learning-based ransomware detection evolved over time?
2. RQ2. Which authors, institutions, countries, and publication sources have contributed most significantly to ransomware detection research?
3. RQ3. What are the dominant keywords, conceptual structures, and thematic clusters in ransomware detection research?
4. RQ4. How have behavioral analysis, dynamic analysis, API call sequences, system call analysis, and deep learning approaches shaped the intellectual development of the field?
5. RQ5. What emerging research directions can be identified for future ransomware detection studies?

The remainder of this paper is organized as follows. The next section reviews ransomware detection research and clarifies the relationship between detection paradigms and computational methods. The methodology section then describes the Web of Science source, search query, screening criteria, data cleaning, and bibliometric procedures. The results and discussion sections present publication trends, sources, contributors, citation patterns, keyword structures, thematic findings, research gaps, and future directions, followed by the conclusion.

2. RELATED WORKS

Ransomware detection literature can be organized along two independent dimensions. The first is the detection paradigm or data-acquisition strategy: static analysis examines artifacts before execution, whereas dynamic or behavioral analysis observes runtime events such as API calls, system calls, file access, process activity, registry changes, memory usage, and resource consumption. The second is the computational method used to interpret those data, including ML and DL. This distinction prevents ML or DL from being treated as synonyms for behavioral analysis. Recent work increasingly combines dynamic data with learning methods to improve early detection and resilience against obfuscated or previously unseen ransomware variants [1], [3], [6], [7], [8].

Within the computational dimension, DL has become a major direction because it can learn complex sequential and nonlinear representations of ransomware behavior. CNN, LSTM, BiLSTM, GRU, attention, BERT, RoBERTa, and transformer-based models have been applied to execution traces for detection and ransomware-family classification [6], [7], [10], [11]. Other studies address explainability, uncertainty, few-shot learning, early detection, and dataset development, including XRan for explainable dynamic detection [6], fine-tuned language models for family classification [10], Siamese networks for limited-sample classification [12], and RanSMAP for storage and memory access patterns [30]. However, the literature remains fragmented in datasets, feature representation, evaluation metrics, and deployment readiness. This study addresses that fragmentation through bibliometric mapping of the field rather than through a new classifier benchmark.

3. METHOD

This study employed a bibliometric analysis approach to examine the development of ransomware detection research, particularly studies related to machine learning, deep learning, behavioral analysis, dynamic analysis, API call sequences, system calls, and cybersecurity applications. Bibliometric analysis was selected because it enables a systematic mapping of publication trends, scientific productivity, citation impact, collaboration patterns, keyword structures, thematic development, and future research directions. This approach is suitable for identifying how ransomware detection research has evolved and how major themes

ISSN: 2715-6427

such as behavioral analysis, deep learning models, real time detection, and transformer-based approaches have shaped the field [17], [18].

The data were retrieved from the Web of Science Core Collection because it provides standardized bibliographic metadata, cited references, author affiliations, source information, citation counts, and indexing categories suitable for bibliometric analysis [17], [18]. The search used the Topic Search field, which covers titles, abstracts, author keywords, and Keywords Plus. The query was: TS=(ransomware AND (detection OR classification OR identification) AND ("machine learning" OR "deep learning" OR "neural network" OR CNN OR LSTM OR BiLSTM OR BERT OR RoBERTa OR transformer OR "attention mechanism") AND ("behavioral analysis" OR "dynamic analysis" OR "API call" OR "system call" OR "behavioral feature" OR "memory behavior" OR "file activity" OR "process behavior" OR "runtime indicator" OR "runtime behavior")). This formulation aligns the search terms with the behavioral and runtime indicators used in the inclusion criteria.

The initial search produced 481 records. Records were included when they were English-language publications indexed in Web of Science and connected ransomware detection, classification, or family attribution with at least one AI term in the query and at least one static, dynamic, behavioral, or runtime indicator, such as API calls, system calls, memory behavior, file activity, process behavior, or runtime behavior. Records were excluded when they lacked a clear ransomware component, lacked sufficient bibliographic metadata, or were classified as editorial materials or retracted publications. Screening removed one retracted publication and two editorial materials, leaving 478 eligible records for the analytical corpus. Figure 1 summarizes this screening outcome.

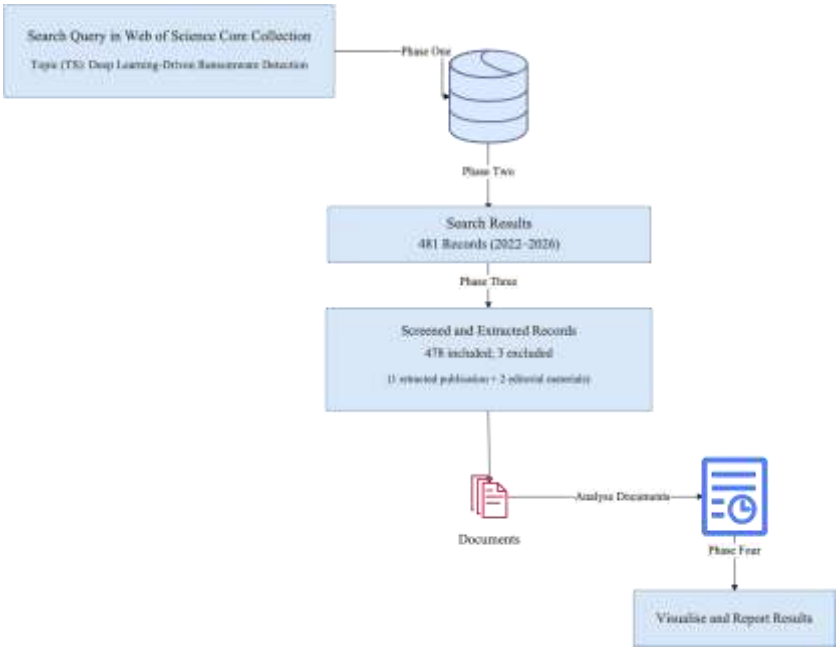

Figure 1. Phases of the research methodology

Before analysis, data cleaning was conducted to improve the consistency and reliability of the bibliometric results. Duplicate records were checked using title, DOI, and Web of Science accession number. Author keywords were standardized by merging similar terms, such as machine learning and ML, deep learning and DL, API call and application programming interface call, system call and system calls, and behavioral analysis and behaviour analysis. Author names, source titles, and institutional names were also reviewed to reduce fragmentation caused by spelling differences or abbreviation variants. Records with missing DOI or author keywords were not automatically removed because they still contained useful bibliographic information for publication trend, citation, and collaboration analysis.

The cleaned dataset was analyzed using Bibliometrix and Biblioshiny in R. The analysis consisted of two main procedures, namely performance analysis and science mapping [17], [18]. Performance analysis was used to examine annual scientific production, document type distribution, most productive sources, influential authors, productive countries, institutional contributions, and globally cited documents. Science mapping was used to examine the intellectual and conceptual structure of the field through keyword co-occurrence analysis, thematic mapping, country collaboration analysis, and conceptual structure analysis. Keyword co-occurrence

was used to identify dominant research topics, while thematic mapping was used to classify themes based on centrality and density into motor themes, basic themes, niche themes, and emerging or declining themes.

The results of the bibliometric analysis were interpreted to answer the research questions concerning publication growth, major contributors, dominant keywords, thematic clusters, collaboration patterns, and future directions. Particular attention was given to the role of behavioral analysis, dynamic analysis, API call sequences, system calls, deep learning, transformer models, explainable artificial intelligence, and real time ransomware detection. Although this study provides a systematic overview of the field, it has several limitations. The dataset was limited to Web of Science, so relevant publications indexed only in Scopus, IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, or Google Scholar may not be included. In addition, bibliometric analysis relies on metadata rather than full text evaluation, so it cannot fully assess the technical validity, dataset quality, algorithmic robustness, or reproducibility of individual ransomware detection models.

4. RESULTS AND DISCUSSION

This section presents and interprets the bibliometric results using figures, tables, performance indicators, and science-mapping outputs.

The Web of Science search retrieved 481 records. After one retracted publication and two editorial materials were excluded, the final analytical corpus contained 478 eligible records. Figure 2 makes this distinction explicit. The annual counts in Figure 2 and Table 1 are reported from the retrieval-stage export and therefore sum to 481; the screening result used for the analytical corpus is 478. Stating both counts avoids conflating the initial retrieval with the eligible research set.

The distribution of document types shows that journal articles dominated the dataset, followed by conference proceedings and review papers. This pattern indicates that ransomware detection research is simultaneously developing through mature journal-based contributions and fast-moving conference-based dissemination. This profile is expected in computer science and cybersecurity, where conference papers often report early experimental models, while journal articles provide more complete methodological validation and discussion [2], [17].

The dominance of journal articles suggests that ransomware detection has become a consolidated academic topic, while the large number of proceedings papers reflects the technical and experimental nature of the field. Since cybersecurity threats evolve rapidly, researchers often publish initial models and prototypes in conferences before extending them into full journal articles. This pattern is also consistent with the development of deep learning and transformer based ransomware detection, where new architectures such as LSTM, BiLSTM, BERT, RoBERTa, and attention based models are frequently tested in experimental studies [6], [7], [10], [11].

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2022:2026
Sources (Journals, Books, etc)	271
Documents	481
Annual Growth Rate %	-9.35
Document Average Age	1.98
Average citations per doc	8.05
References	15855
DOCUMENT CONTENTS	
Keywords Plus (ID)	191
Author's Keywords (DE)	1283
AUTHORS	
Authors	1599
Authors of single-authored docs	17
AUTHORS COLLABORATION	
Single-authored docs	17
Co-Authors per Doc	3.96
International co-authorships %	31.81
DOCUMENT TYPES	
article	301
article; early access	6
article; proceedings paper	1
article; retracted publication	1
editorial material	1
editorial material; early access	1
proceedings paper	150
review	18

Figure 2. Main information of the bibliometric dataset

ISSN: 2715-6427

4.2 Annual Scientific Production

The annual publication trend shows a clear increase from 2022 to 2025. Figure 2 visualizes the year-on-year retrieval trend, while Table 1 reports the corresponding counts and percentages. The number of retrieved publications increased from 77 documents in 2022 to 149 documents in 2025. The number of documents in 2026 was 52, but this figure should be interpreted carefully because 2026 is an incomplete publication year in the dataset. The pattern confirms that ransomware detection has become a rapidly expanding topic within cybersecurity and artificial intelligence research [1], [2], [20], [21].

The increasing trend indicates that ransomware detection research has entered a period of rapid growth. Between 2022 and 2025, the number of publications almost doubled. This growth can be explained by several factors. First, ransomware has increasingly targeted critical infrastructure, healthcare, educational institutions, and enterprise networks, creating a strong demand for intelligent detection systems. Second, traditional signature-based detection is increasingly insufficient against new and obfuscated ransomware variants. Third, recent advances in machine learning, deep learning, and transformer based models have created new opportunities to detect ransomware through behavioral traces such as API call sequences, system calls, memory access patterns, file operations, and process activities [2], [6], [7], [8], [11].

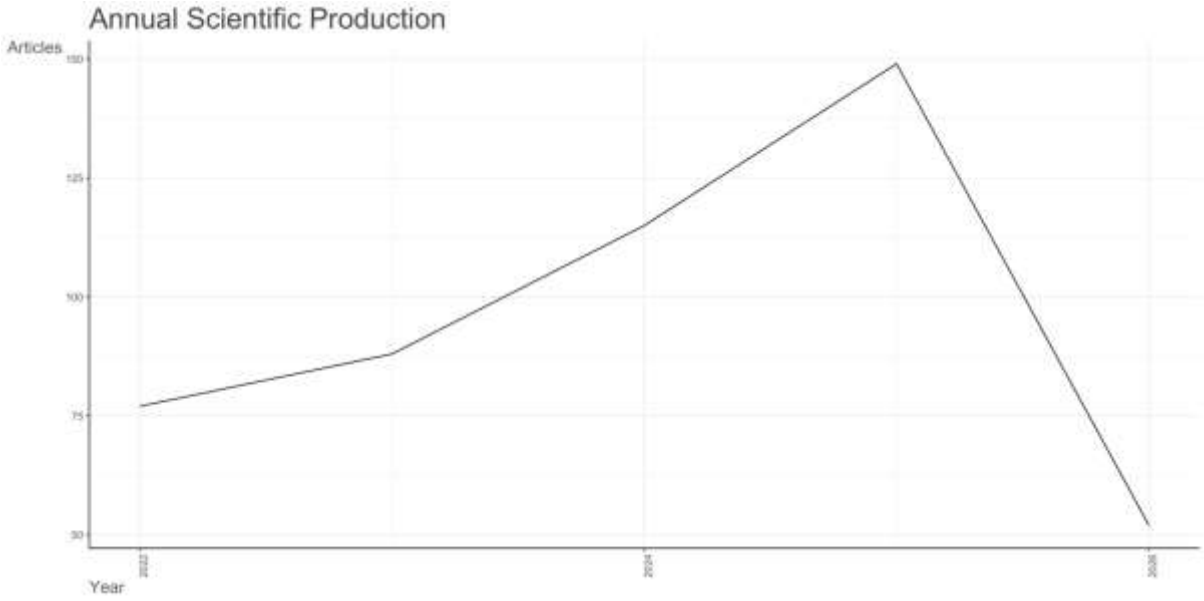

Figure 2. Annual scientific production

Table 1. Annual scientific production from the retrieval-stage export (n = 481)

Year	Number of documents	Percentage
2022	77	16.01%
2023	88	18.30%
2024	115	23.91%
2025	149	30.98%
2026	52	10.81%

The growth pattern also suggests that the field is moving from general malware detection toward more specialized ransomware detection. Earlier studies often treated ransomware as one class within broader malware detection, while recent studies increasingly focus on ransomware specific behavior, family classification, early detection, and real time detection. This transition is important because ransomware has distinctive behavioral characteristics, especially during file enumeration, encryption, system modification, and resource access [7], [8], [10], [11], [22].

4.3 Leading Publication Sources

The analysis of publication sources shows that ransomware detection research is distributed across cybersecurity, artificial intelligence, computer science, information systems, and engineering venues. As presented in Figure 3 and Table 2, the most productive source was IEEE Access with 32 documents, followed

by Computers & Security with 18 documents, Sensors with 14 documents, and International Journal of Advanced Computer Science and Applications and Scientific Reports with 12 documents each. This distribution reflects the multidisciplinary nature of ransomware detection, which combines malware analysis, machine learning, deep learning, behavioral computing, IoT security, and applied cybersecurity [2], [20], [23], [24].

The dominance of IEEE Access indicates that open access multidisciplinary engineering journals play an important role in disseminating ransomware detection research. Meanwhile, the strong presence of Computers & Security and International Journal of Information Security confirms that the topic is highly relevant to specialized cybersecurity venues. The presence of Sensors, Electronics, and Applied Sciences suggests that ransomware detection is increasingly linked with Internet of Things, cyber physical systems, embedded devices, and applied artificial intelligence [19], [20], [23], [25].

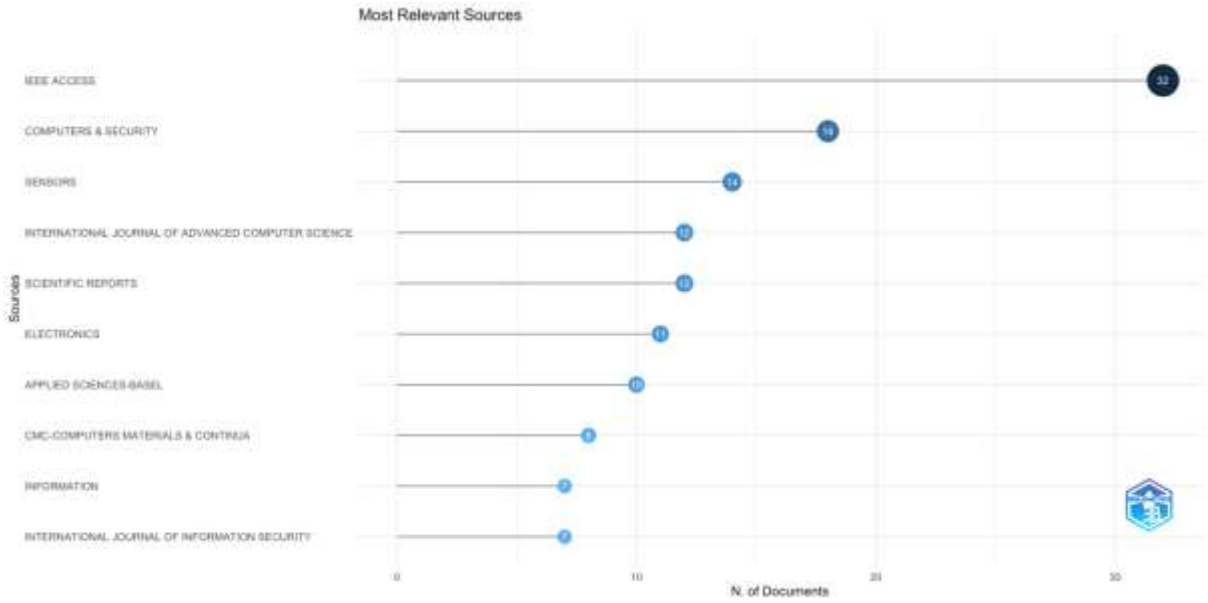

Figure 3. Most productive publication sources

Table 2. Most productive publication sources

Sources	Articles
IEEE ACCESS	32
COMPUTERS & SECURITY	18
SENSORS	14
INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS	12
SCIENTIFIC REPORTS	12
ELECTRONICS	11
APPLIED SCIENCES-BASEL	10
CMC-COMPUTERS MATERIALS & CONTINUA	8
INFORMATION	7
INTERNATIONAL JOURNAL OF INFORMATION SECURITY	7

This source distribution also shows that ransomware detection is no longer a narrow malware analysis topic. It has expanded into several application domains, including healthcare IoT, smart infrastructure, cloud security, encrypted traffic analysis, and enterprise cybersecurity. Therefore, future studies should not only focus on classification accuracy but also consider operational deployment, real time performance, interpretability, and domain specific constraints [6], [8], [11], [15], [22].

ISSN: 2715-6427

4.4 Highly Cited Documents and Intellectual Foundations

The most cited documents show that the intellectual foundation of ransomware detection research is built on three main streams. The first stream consists of broad cybersecurity and malware detection reviews, the second stream consists of ransomware specific surveys and taxonomies, and the third stream consists of empirical studies proposing machine learning, deep learning, or behavioral detection models. The most globally cited documents are presented in Figure 4 and Table 3, showing the influence of studies on cybersecurity vulnerabilities, deep learning based malware detection, dynamic ransomware detection surveys, ransomware taxonomy, crypto ransomware detection, and artificial intelligence based malware mitigation [20], [21], [24], [25], [26], [27], [28], [29].

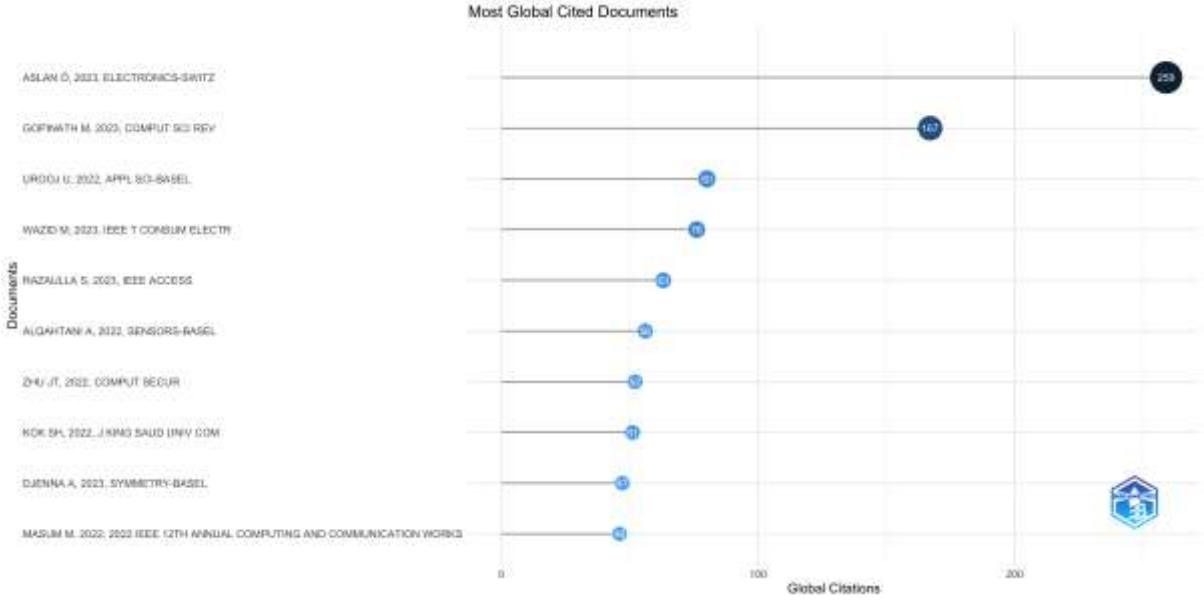

Figure 4. Most globally cited documents

Table 3. Most globally cited documents

Paper	DOI	Total Citations	TC per Year	Normalized TC
Aslan Ö, 2023, Electronics-Switz [20]	10.3390/electronics12061333	259	64,75	14,60
Gopinath M, 2023, Comput Sci Rev [24]	10.1016/j.cosrev.2022.100529	167	41,75	9,41
Urooj U, 2022, Appl Sci-Basel [26]	10.3390/app12010172	80	16,00	5,55
Wazid M, 2023, Ieee T Consum Electr [27]	10.1109/TCE.2022.3208795	76	19,00	4,28
Razaulla S, 2023, IEEE Access [21]	10.1109/ACCESS.2023.3268535	63	15,75	3,55
Alqahtani A, 2022, Sensors-Basel [25]	10.3390/s22051837	56	11,20	3,88
Zhu Jt, 2022, Comput Secur [12]	10.1016/j.cose.2022.102691	52	10,40	3,61
Kok Sh, 2022, J King Saud Univ Com [28]	10.1016/j.jksuci.2020.06.012	51	10,20	3,54
Djenna A, 2023, Symmetry-Basel[29]	10.3390/sym15030677	47	11,75	2,65
Masum M, 2022, IEEE Ccwc [9]	10.1109/CCWC54503.2022.9720869	46	9,20	3,19

The citation structure shows that review papers and survey papers have high influence because they provide conceptual foundations, taxonomies, and research directions. However, empirical model-based studies also receive significant attention, especially when they address early detection, dynamic analysis, or advanced learning models. For example, few shot meta learning using entropy features has become influential because it addresses the scarcity of labeled ransomware samples, while dynamic analysis and pre encryption detection

papers are highly cited because they focus on identifying ransomware before irreversible damage occurs [5], [26], [28].

This citation pattern confirms that the field is driven by both methodological innovation and practical urgency. Researchers are not only interested in improving classification metrics but also in solving operational problems, including how to detect unknown ransomware, how to respond before encryption is completed, and how to design systems that remain effective under changing attack behavior [5], [6], [8], [11], [15].

4.5 Productive Authors, Institutions, and Countries

The author productivity analysis shows that ransomware detection research involves a broad and distributed scholarly community. Table 4 lists the most productive authors, including Sheldon F. T., Bou Harb E., Bovet G., Stiller B., Blasingame Z., Celdran A. H., Gazzan M., Kobayashi R., Liu C., and Mahboubi A. The productivity distribution is relatively dispersed, indicating that the field has not yet been dominated by a small group of authors. This reflects the emerging nature of ransomware detection research and its multidisciplinary connections with artificial intelligence, malware analysis, IoT security, network security, and digital forensics [2], [6], [21], [23].

Table 4. Most productive authors

Authors	Articles	Articles Fractionalized
Sheldon, Frederick T.	7	3,08
Bou-Harb, Elias	6	1,54
Bovet, Gerome	6	0,91
Stiller, Burkhard	6	0,91
Blasingame, Zander	5	1,58
Celdran, Alberto Huertas	5	0,77
Gazzan, Mazen	5	1,88
Kobayashi, Ryotaro	5	2,50
Liu, Chen	5	1,58
Mahboubi, Arash	5	1,04

Country productivity shows that India, the United States, Saudi Arabia, China, Australia, Korea, Pakistan, the United Kingdom, Ireland, and Italy were among the most active contributors. These results are detailed in Table 5, which reports article counts, single-country publications, multiple-country publications, and the percentage of multiple-country contributions. The distribution reflects the global relevance of ransomware defense and the expanding role of artificial intelligence in cybersecurity research [2], [21], [23], [25].

Table 5. Most productive countries based on author affiliations

Country	Articles	Articles %	SCP	MCP	MCP %
India	83	17,3	72	11	13,3
Usa	76	15,8	61	15	19,7
Saudi Arabia	44	9,1	30	14	31,8
China	43	8,9	27	16	37,2
Australia	21	4,4	17	4	19
Korea	18	3,7	6	12	66,7
Pakistan	16	3,3	4	12	75
United Kingdom	11	2,3	7	4	36,4
Ireland	9	1,9	4	5	55,6
Italy	9	1,9	5	4	44,4

The institutional distribution was also diverse. Institutions such as Charles Sturt University, King Khalid University, King Saud University, Universiti Teknologi Malaysia, Najran University, Princess Nourah bint

Abdulrahman University, University of Texas at San Antonio, Kennesaw State University, Air University Islamabad, and Al Jouf University appeared as active contributors. This indicates that ransomware detection research is not concentrated only in elite global institutions but is increasingly being developed across a wide range of universities and cybersecurity research centers.

From a collaboration perspective, the topic has strong potential for international cooperation because ransomware is a global threat that crosses national boundaries. However, the field still requires stronger cross country collaboration, especially in dataset sharing, benchmark standardization, realistic testing environments, and reproducible experimental protocols. These needs are consistent with bibliometric recommendations that collaboration networks can improve knowledge diffusion and methodological consolidation in emerging research areas [17], [18]

4.6 Keyword Co Occurrence and Conceptual Structure

The keyword analysis shows that the dominant terms in the dataset were ransomware, machine learning, deep learning, cybersecurity, ransomware detection, malware, malware detection, artificial intelligence, feature selection, Internet of Things, dynamic analysis, static analysis, classification, blockchain, anomaly detection, encryption, and explainable AI. As shown in Table 6, these keywords indicate that ransomware detection research is structured around machine learning based detection, deep learning and sequence modeling, behavioral and dynamic analysis, and deployment contexts such as IoT, cloud, healthcare, and real time monitoring [2], [6], [7], [10], [23].

The first conceptual area is machine learning-based ransomware detection. This cluster includes terms such as machine learning, feature selection, classification, Random Forest, Support Vector Machine, XGBoost, LightGBM, and anomaly detection. The dominance of machine learning indicates that many studies still rely on engineered features and classical classifiers because these models can provide strong performance with relatively low computational cost, especially when applied to structured tabular data such as entropy values, memory features, file statistics, and process based indicators [1], [3], [4], [9], [23].

Table 6. Most frequent author keywords

Words	Occurrences
ransomware	202
machine learning	185
deep learning	97
cybersecurity	78
ransomware detection	78
malware	67
malware detection	56
security	44
classification	40
artificial intelligence	31

The second conceptual area is deep learning and sequence modeling. This cluster includes deep learning, neural network, LSTM, BiLSTM, CNN, BERT, transformer, attention mechanism, and API call sequences. Studies in this area treat ransomware behavior as sequential data. For example, Lin et al. [7] showed that LSTM and BERT models can distinguish ransomware from benign applications using API call sequences collected from Cuckoo Sandbox, while Hussain et al. [10] showed that transformer based models can support ransomware family classification. This confirms that API call sequences can be treated as structured behavioral narratives that deep learning models can learn from [7], [10], [11].

The third conceptual area is behavioral and dynamic analysis. This cluster includes behavioral analysis, dynamic analysis, API call, system call, DLL, mutual exclusion, file activity, process behavior, and runtime behavior. The prominence of this cluster reflects the shift from static detection toward runtime behavior monitoring. Gulmez et al. [6] proposed XRan, an explainable deep learning based ransomware detection system that combines different dynamic analysis sequences, including API calls, DLLs, and mutual exclusions, while Chew et al. [8] and Davidian et al. [11] demonstrated the relevance of system call and API call features for early and real time detection.

The fourth conceptual area is application environments and deployment contexts. This includes IoT, healthcare IoT, cloud security, blockchain, smart healthcare, encrypted traffic, and real time detection. The increasing appearance of IoT related keywords suggests that ransomware detection is expanding beyond desktop systems into connected devices and cyber physical systems. This shift is important because IoT devices often have limited computing capacity, fragmented standards, and weak security configurations, making ransomware detection more challenging [19], [23], [27].

4.7 Thematic Interpretation

Thematic analysis indicates that ransomware detection research is structured around several dominant and emerging themes. The strongest basic theme is the relationship between ransomware, machine learning, malware detection, and cybersecurity. This theme forms the conceptual foundation of the field because almost all studies are connected to the broader objective of detecting malicious behavior using computational intelligence [2], [17], [18], [24].

A more specialized and rapidly growing theme is deep learning based behavioral detection. This theme includes API call sequence modeling, LSTM, BERT, RoBERTa, CNN, BiLSTM, and attention mechanisms. The emergence of transformer-based models is particularly significant because they can capture contextual dependencies among API calls and behavioral events. This suggests that future ransomware detection research will increasingly move toward contextual sequence representation rather than simple frequency based feature engineering [7], [10], [11].

Another important theme is explainable ransomware detection. Although deep learning models often achieve high performance, they are frequently criticized as black box systems. In cybersecurity operations, interpretability is important because analysts need to understand why a process is classified as malicious, which features contribute to detection, and whether the alert is reliable. The presence of explainable AI in the keyword structure indicates that the field is beginning to address this limitation, particularly through approaches that combine dynamic analysis with model explanation techniques [6], [15].

The dataset also shows the emergence of early and real time ransomware detection. This is a critical theme because ransomware causes damage quickly once encryption begins. Detection after encryption is less valuable than detection during early behavioral stages. Recent studies on system call based detection, API call based early detection, and real time backup integration show that detection models must be evaluated not only by accuracy but also by their ability to operate before irreversible damage occurs [8], [11], [22].

4.8 Discussion of Research Gaps

The results reveal several important gaps in the ransomware detection literature. First, many studies still prioritize model accuracy over operational feasibility. Metrics such as accuracy, precision, recall, and F1 score are frequently reported, but inference time, false positive rate, computational overhead, memory consumption, and deployment constraints are less consistently discussed. This is problematic because a ransomware detector must operate quickly and reliably in real environments [4], [5], [8], [11], [15].

Second, dataset diversity remains a major limitation. Many studies rely on a single dataset, self-collected samples, or controlled sandbox environments. Although these settings are useful for experimental validation, they may not capture the diversity of ransomware behavior in real systems. Recent studies have started to address this limitation through new datasets such as RanSMAP, which focuses on ransomware storage and memory access patterns for deep learning based detectors, and through broader malware benchmarking across ransomware and other malware families [23], [30].

Third, behavioral feature representation remains fragmented. Some studies use API call sequences, while others use system calls, file entropy, memory traces, network traffic, registry behavior, PE headers, or hardware performance counters. Each feature type captures a different aspect of ransomware behavior, but there is still limited consensus on which feature combinations are most effective under real deployment conditions. This reinforces the need for integrated detection systems that combine data acquisition, preprocessing, feature representation, classification, explanation, and decision output [6], [7], [8], [30].

Fourth, interpretability and robustness remain underdeveloped. Deep learning and transformer-based models can identify complex behavioral patterns, but they may be vulnerable to adversarial manipulation, concept drift, and distribution shifts. Ransomware authors can modify API sequences, delay encryption behavior, inject benign calls, or change execution patterns to evade detection. Therefore, future research should integrate explainable AI, uncertainty estimation, adversarial robustness, and concept drift adaptation [5], [6], [15].

Fifth, real time detection needs stronger evaluation. Some studies claim real time applicability, but not all provide detailed evidence on response time, system overhead, or deployment feasibility. Future research should evaluate detection latency, resource consumption, and practical integration with endpoint protection systems, backup systems, and security operation centers [8], [11], [22].

ISSN: 2715-6427

4.9 Future Research Directions

Based on the bibliometric results and thematic interpretation, several future research directions can be proposed. First, future studies should develop integrated ransomware detection systems rather than isolated classification models. Such systems should include behavioral data acquisition, preprocessing, sequence construction, feature extraction, classification, explanation, and response mechanisms [6], [8], [11], [30].

Second, future research should strengthen behavioral feature fusion. Combining API calls, system calls, file activity, process behavior, memory access, entropy change, and resource usage may improve detection robustness. However, feature fusion should be designed carefully to avoid excessive computational overhead and to maintain practical feasibility in real time environments [6], [7], [8], [11], [30].

Third, transformer-based ransomware detection should be explored further. BERT, RoBERTa, and other transformer architectures are promising because they can model contextual relationships among API calls or system events. However, future studies should compare them with lighter models to determine whether their performance gains justify their computational cost [7], [9], [14].

Fourth, explainable and uncertainty aware detection should become a priority. Ransomware detection systems should provide not only a class label but also confidence information, feature explanations, and behavioral evidence. This can improve analyst trust and reduce false positive fatigue.

Fifth, more realistic evaluation protocols are needed. Future studies should use cross dataset validation, temporal validation, unseen ransomware family testing, and zero-day simulation. These evaluation designs would provide stronger evidence of model generalization than single dataset experiments [5], [23], [30].

4.10 Summary of Findings

Overall, the results show that ransomware detection research has grown rapidly from 2022 to 2025, with strong contributions from journals and conferences in computer science, cybersecurity, artificial intelligence, and applied engineering. The field is dominated by machine learning and deep learning approaches, while behavioral analysis and dynamic analysis are becoming increasingly central. Keyword and thematic patterns indicate a shift from static and signature based detection toward API call sequence modeling, system call analysis, transformer based learning, explainable AI, early detection, and real time deployment [6], [7], [8], [10], [11].

The findings also show that the field is still fragmented in terms of datasets, feature representation, evaluation metrics, and deployment assumptions. Therefore, the next stage of ransomware detection research should focus on integrated, interpretable, efficient, and behavior aware systems. This direction is consistent with the broader need to detect ransomware earlier, reduce false positives, adapt to new variants, and support practical cybersecurity operations [5], [15], [23], [30].

Sixth, deployment-oriented metrics should be reported more consistently. In addition to accuracy, precision, recall, and F1 score, studies should report false positive rate, inference latency, CPU usage, memory consumption, throughput, and response time. These metrics are essential for real time ransomware defense and operational deployment [8], [11], [22].

5. CONCLUSION

Analysis of the 478 eligible Web of Science records shows that ransomware detection research expanded rapidly during 2022–2025, while the 2026 count remains incomplete. The field is driven mainly by journal articles and conference proceedings and is distributed across cybersecurity, artificial intelligence, computer science, and engineering sources. Its thematic structure indicates a transition from static and signature-based detection toward DL-supported behavioral and dynamic analysis, including API-call and system-call modeling, early detection, explainable AI, transformer-based learning, and real-time monitoring. The mapping also reveals continuing fragmentation in datasets, behavioral feature representation, false-positive evaluation, inference latency, interpretability, robustness, and deployment feasibility. These findings support future work on integrated and explainable DL detectors that combine diverse runtime signals with temporal validation, concept-drift handling, uncertainty-aware decisions, and realistic deployment testing. Because this study analyzes bibliographic metadata rather than re-evaluating individual models, its conclusions describe field-level research patterns and do not establish the performance superiority of any particular algorithm.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the Institute for Research and Community Service, Universitas Lancang Kuning, for providing financial support for this research through the APBU

research funding scheme. The support has contributed significantly to the completion of this study and the development of this manuscript.

REFERENCES

[1] E. Berrueta, D. Morato, E. Magaña, and M. Izal, "Crypto-ransomware detection using machine learning models in file-sharing network scenarios with encrypted traffic," *Expert Syst. Appl.*, vol. 209, p. 118299, Dec. 2022, doi: 10.1016/j.eswa.2022.118299.

[2] D. W. Fernando, N. Komninos, and T. Chen, "A Study on the Evolution of Ransomware Detection Using Machine Learning and Deep Learning Techniques," *IoT*, vol. 1, no. 2, pp. 551–604, Dec. 2020, doi: 10.3390/iot1020030.

[3] C.-M. Hsu, C.-C. Yang, H.-H. Cheng, P. E. Setiasabda, and J.-S. Leu, "Enhancing File Entropy Analysis to Improve Machine Learning Detection Rate of Ransomware," *IEEE Access*, vol. 9, pp. 138345–138351, 2021, doi: 10.1109/ACCESS.2021.3114148.

[4] S. H. Kok, A. Azween, and N. Jhanjhi, "Evaluation metric for crypto-ransomware detection using machine learning," *J. Inf. Secur. Appl.*, vol. 55, p. 102646, Dec. 2020, doi: 10.1016/j.jisa.2020.102646.

[5] D. W. Fernando and N. Komninos, "FeSAD ransomware detection framework with machine learning using adaption to concept drift," *Comput. Secur.*, vol. 137, p. 103629, Feb. 2024, doi: 10.1016/j.cose.2023.103629.

[6] S. Gulmez, A. Gorgulu Kakisim, and I. Sogukpinar, "XRan: Explainable deep learning-based ransomware detection using dynamic analysis," *Comput. Secur.*, vol. 139, p. 103703, Apr. 2024, doi: 10.1016/j.cose.2024.103703.

[7] T.-L. Lin *et al.*, "Ransomware Detection by Distinguishing API Call Sequences through LSTM and BERT Models," *Comput. J.*, vol. 67, no. 2, pp. 632–641, Feb. 2024, doi: 10.1093/comjnl/bxad005.

[8] C. J. W. Chew, V. Kumar, P. Patros, and R. Malik, "Real-time system call-based ransomware detection," *Int. J. Inf. Secur.*, vol. 23, no. 3, pp. 1839–1858, Jun. 2024, doi: 10.1007/s10207-024-00819-x.

[9] M. Masum *et al.*, "Ransomware Classification and Detection With Machine Learning Algorithms," presented at the 2022 IEEE 12TH ANNUAL COMPUTING AND COMMUNICATION WORKSHOP AND CONFERENCE (CCWC), 2022, pp. 316–322. doi: 10.1109/CCWC54503.2022.9720869.

[10] A. Hussain, A. Saadia, and F. M. Alserhani, "Ransomware detection and family classification using fine-tuned BERT and RoBERTa models," *Egypt. Inform. J.*, vol. 30, p. 100645, Jun. 2025, doi: 10.1016/j.eij.2025.100645.

[11] M. Davidian, M. Kiperberg, and N. Vanetik, "Early Ransomware Detection with Deep Learning Models," *Future Internet*, vol. 16, no. 8, p. 291, Aug. 2024, doi: 10.3390/fi16080291.

[12] J. Zhu, J. Jang-Jaccard, A. Singh, I. Welch, H. Al-Sahaf, and S. Camtepe, "A few-shot meta-learning based siamese neural network using entropy features for ransomware classification," *Comput. Secur.*, vol. 117, Jun. 2022, doi: 10.1016/j.cose.2022.102691.

[13] T. Dam, N. Nguyen, T. Le, T. Le, S. Uwizeyemungu, and T. Le-Dinh, "Visualizing Portable Executable Headers for Ransomware Detection: A Deep Learning-Based Approach," *J. Univers. Comput. Sci.*, vol. 30, no. 2, pp. 262–286, 2024, doi: 10.3897/jucs.104901.

[14] M. Gazzan, B. Alobaywi, M. Almutairi, and F. T. Sheldon, "A Deep Learning Framework for Enhanced Detection of Polymorphic Ransomware," *Future Internet*, vol. 17, no. 7, p. 311, Jul. 2025, doi: 10.3390/fi17070311.

[15] M. Gazzan and F. T. Sheldon, "Novel Ransomware Detection Exploiting Uncertainty and Calibration Quality Measures Using Deep Learning," *Information*, vol. 15, no. 5, p. 262, 2024, doi: 10.3390/info15050262.

[16] J. Lee, J. Kim, H. Jeong, and K. Lee, "A Machine Learning-Based Ransomware Detection Method for Attackers' Neutralization Techniques Using Format-Preserving Encryption," *Sensors*, vol. 25, no. 8, p. 2406, Apr. 2025, doi: 10.3390/s25082406.

[17] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *J. Bus. Res.*, vol. 133, pp. 285–296, Sep. 2021, doi: https://doi.org/10.1016/j.jbusres.2021.04.070.

[18] M. Aria and C. Cuccurullo, "bibliometrix : An R-tool for comprehensive science mapping analysis," *J. Informetr.*, vol. 11, no. 4, pp. 959–975, Nov. 2017, doi: https://doi.org/10.1016/j.joi.2017.08.007.

[19] A. Alqahtani, M. O. Ohemeng, and F. T. Sheldon, "An Intelligent Sensing Framework for Early Ransomware Detection Using MHSA-LSTM Machine Learning," *Sensors*, vol. 26, no. 3, p. 952, Feb. 2026, doi: 10.3390/s26030952.

[20] Ö. Aslan, S. S. Aktuğ, M. Ozkan-Okay, A. A. Yilmaz, and E. Akin, "A Comprehensive Review of Cyber Security Vulnerabilities, Threats, Attacks, and Solutions," *Electronics*, vol. 12, no. 6, p. 1333, Mar. 2023, doi: 10.3390/electronics12061333.

[21] S. Razaulla *et al.*, "The Age of Ransomware: A Survey on the Evolution, Taxonomy, and Research Directions," *IEEE ACCESS*, vol. 11, pp. 40698–40723, 2023, doi: 10.1109/access.2023.3268535.

[22] K. Higuchi and R. Kobayashi, "Real-time open-file backup system with machine-learning detection model for ransomware," *Int. J. Inf. Secur.*, vol. 24, no. 1, p. 54, Feb. 2025, doi: 10.1007/s10207-024-00966-1.

[23] M. Maghanaki, S. Keramati, F. F. Chen, and M. Shahin, "Systematic Evaluation of Machine Learning and Deep Learning Models for IoT Malware Detection Across Ransomware, Rootkit, Spyware, Trojan, Botnet, Worm, Virus, and Keylogger," *Sensors*, vol. 26, no. 6, p. 1750, Mar. 2026, doi: 10.3390/s26061750.

[24] M. Gopinath and S. Sethuraman, "A comprehensive survey on deep learning based malware detection techniques," *Comput. Sci. Rev.*, vol. 47, Feb. 2023, doi: 10.1016/j.cosrev.2022.100529.

[25] A. Alqahtani and F. T. Sheldon, "A Survey of Crypto Ransomware Attack Detection Methodologies: An Evolving Outlook," *Sensors*, vol. 22, no. 5, p. 1837, 2022, doi: 10.3390/s22051837.

ISSN: 2715-6427

[26] U. Urooj, B. Al-rimy, A. Zainal, F. Ghaleb, and M. Rassam, "Ransomware Detection Using the Dynamic Analysis and Machine Learning: A Survey and Research Directions," *Appl. Sci.-BASEL*, vol. 12, no. 1, Jan. 2022, doi: 10.3390/app12010172.

[27] M. Wazid, A. Das, and S. Shetty, "BSFR-SH: Blockchain-Enabled Security Framework Against Ransomware Attacks for Smart Healthcare," *IEEE Trans. Consum. Electron.*, vol. 69, no. 1, pp. 18–28, Feb. 2023, doi: 10.1109/TCE.2022.3208795.

[28] S. Kok, A. Abdullah, and N. Jhanjhi, "Early detection of crypto-ransomware using pre-encryption detection algorithm," *J. KING SAUD Univ. Comput. Inf. Sci.*, vol. 34, no. 5, pp. 1984–1999, May 2022, doi: 10.1016/j.jksuci.2020.06.012.

[29] A. Djenna, A. Bouridane, S. Rubab, and I. Marou, "Artificial Intelligence-Based Malware Detection, Analysis, and Mitigation," *SYMMETRY-BASEL*, vol. 15, no. 3, Mar. 2023, doi: 10.3390/sym15030677.

[30] M. Hirano and R. Kobayashi, "RanSMAP: Open dataset of Ransomware Storage and Memory Access Patterns for creating deep learning based ransomware detectors," *Comput. Secur.*, vol. 150, p. 104202, Mar. 2025, doi: 10.1016/j.cose.2024.104202.